

(19)



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS  
ESPAÑA



(11) Número de publicación: **2 666 301**

(51) Int. Cl.:

**C07K 16/00** (2006.01)  
**C07K 16/08** (2006.01)  
**C07K 16/18** (2006.01)  
**C07K 16/44** (2006.01)

(12)

## TRADUCCIÓN DE PATENTE EUROPEA

T3

(86) Fecha de presentación y número de la solicitud internacional: **09.03.2012 PCT/US2012/028501**

(87) Fecha y número de publicación internacional: **13.09.2012 WO12122484**

(96) Fecha de presentación y número de la solicitud europea: **09.03.2012 E 12711501 (2)**

(97) Fecha y número de publicación de la concesión europea: **17.01.2018 EP 2683736**

---

### (54) Título: **Métodos y reactivos para crear anticuerpos monoclonales**

(30) Prioridad:

**09.03.2011 US 201161450922 P  
15.11.2011 US 201161560006 P  
05.12.2011 US 201161566876 P  
03.02.2012 US 201261594729 P**

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:  
**03.05.2018**

(73) Titular/es:

**CELL SIGNALING TECHNOLOGY, INC. (100.0%)  
3 Trask Lane  
Danvers, MA 01923, US**

(72) Inventor/es:

**POLAKIEWICZ, ROBERTO;  
CHEUNG, WAN CHEUNG;  
RUSH, JOHN EDWARD y  
BEAUSOLEIL, SEAN ANDRE**

(74) Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 666 301 T3**

---

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

## DESCRIPCIÓN

Métodos y reactivos para crear anticuerpos monoclonales

### Campo de la descripción

Esta descripción se refiere a biología, y más específicamente, a biología molecular e inmunología.

**5 Antecedentes de la descripción**

Los anticuerpos son polipéptidos biológicamente y comercialmente significativos que se unen con gran especificidad y afinidad a una molécula diana concreta, denominada antígeno. Los anticuerpos son producidos por células inmunitarias de animales vertebrados, y todos los anticuerpos naturales comparten la misma estructura básica, a saber, dos cadenas pesadas idénticas unidas covalentemente a dos cadenas ligeras idénticas. Las regiones N-terminales de una sola cadena pesada y una sola cadena ligera forman un sitio de unión a antígeno que es particular para cada anticuerpo individual. La región C-terminal de las cadenas pesadas determina el isotipo particular del anticuerpo, y la misma célula productora de anticuerpos puede producir anticuerpos de diferentes isotipos, donde todos los anticuerpos producidos por la célula tienen el mismo sitio de unión a antígeno. Los diferentes isotipos suelen realizar diferentes funciones en el animal. Por ejemplo, los anticuerpos del isotipo E (es decir, anticuerpos IgE) están implicados en la respuesta alérgica, mientras que los anticuerpos del isotipo A (es decir, anticuerpos IgA) pueden encontrarse en la membrana de la mucosa, la saliva y la leche materna. La molécula de anticuerpo de cuatro cadenas puede existir por sí misma (p.ej., un anticuerpo IgG) o con monómeros adicionales para formar dímeros (p.ej., un anticuerpo IgA) o incluso pentámeros (p.ej., un anticuerpo IgM).

Con la estructura básica de un anticuerpo bien entendida, se pueden producir anticuerpos recombinantes manipulando las diferentes regiones de un anticuerpo utilizando técnicas de biología molecular convencionales. Por ejemplo, Las Patentes de Estados Unidos Núm. 6.180.370 y 6.548.640 describen la humanización de un anticuerpo que se produce de forma natural en un animal no humano manipulando diversas regiones del anticuerpo no humano utilizando técnicas de biología molecular. Se describen otros métodos para manipular o generar anticuerpos recombinantes utilizando técnicas de biología molecular convencionales (véanse, p.ej. Publicación PCT Núm. WO91/17271, Publicación PCT Núm. WO92/01047; Patentes de Estados Unidos Núm. 5.969.108, 6.331.415, 7.498.024 y 7.485.291).

Durante una respuesta inmunitaria, un animal generará numerosos anticuerpos diferentes, cada uno con una especificidad de unión a antígeno diferente. Esta población de anticuerpos se denomina población policlonal de anticuerpos. Si la respuesta inmunitaria se dirige hacia un antígeno concreto, la mayoría (pero no todos) de los anticuerpos policlónicos preparados por el animal se unirán específicamente a ese antígeno. Sin embargo, con las diferencias en la afinidad de unión y los sitios de unión en el antígeno, algunos de los anticuerpos policlónicos son más favorecidos que otros anticuerpos policlónicos. En su descubrimiento ganador del Premio Nobel en 1975, Kohler y Milstein descubrieron una forma de aislar e inmortalizar una sola célula productora de anticuerpos, que produce un anticuerpo monoclonal que se une específicamente al antígeno de interés, de un animal productor de anticuerpos policlónicos (Kohler y Milstein, Nature 256: 495-497, 1975). Esta tecnología de inmortalización, que implica fusionar las células productoras de anticuerpos a una célula inmortalizada para producir un hibridoma productor de anticuerpos monoclonales, ha sido el modelo de la industria para producir anticuerpos monoclonales durante los últimos 35 años.

A pesar de su popularidad y su longevidad, el método del hibridoma de Kohller y Milstein tiene numerosos inconvenientes. Por ejemplo, conlleva mucho tiempo y requiere mucha mano de obra. De manera más relevante, dado que conlleva mucho tiempo y requiere mucha mano de obra, solo una pequeña fracción de las células productoras de anticuerpos del animal se inmortaliza y se selecciona por su capacidad de producir un anticuerpo que se une específicamente al antígeno. Finalmente, incluso una vez que se ha aislado un hibridoma con la especificidad de antígeno deseada, la obtención de la secuencia de aminoácidos del anticuerpo para facilitar la manipulación adicional, tal como la humanización, del anticuerpo, es ardua y conlleva mucho tiempo.

Existe la necesidad de métodos mejorados para crear anticuerpos monoclonales que se unan específicamente a un antígeno deseado.

Obermeier et al. 'Matching of oligoclonal immunoglobulin transcriptomes and proteomes of cerebrospinal fluid in multiple sclerosis'; Nature Medicine; 2008, vol. 14, no. 6; páginas, 688-693, describe un método para correlacionar proteomas de inmunoglobulina (Ig) con transcriptomas de células B en muestras de fluido y tejido humanos, utilizando la esclerosis múltiple como paradigma.

De Costa et al. 'Sequencing and Quantifying IgG Fragments and Antigen-Binding Regions by Mass Spectrometry'; Journal of Proteome Research; 2010, vol. 9, no. 6, páginas 2937-2945, describe el aislamiento reproducible y la secuenciación de Fab a partir de sueros donantes sanos mediante espectrometría de masas de alta resolución.

**Compendio de la descripción**

La invención es la que se reivindica en las reivindicaciones 1-14.

Los diversos aspectos y realizaciones de la invención proporcionan métodos para crear de forma rápida y precisa anticuerpos monoclonales que se unen específicamente a un antígeno de interés. En aspectos y realizaciones adicionales, la descripción proporciona reactivos y composiciones para realizar los diversos métodos de la invención, y reactivos y composiciones resultantes de la realización de los diversos métodos de la invención. En algunas realizaciones, los métodos, reactivos y composiciones descritos en la presente memoria son útiles para crear anticuerpos monoclonales a partir de la circulación de un sujeto.

5 En un aspecto, la invención proporciona un método para obtener las secuencias de una inmunoglobulina (o regiones variables de la misma) que se une específicamente a un antígeno que comprende: (a) proporcionar secuencias de ácido nucleico que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) de múltiples inmunoglobulinas de al menos un animal; (b) obtener información de los espectros de masas de fragmentos peptídicos de cadenas pesadas de inmunoglobulina y cadenas ligeras de inmunoglobulina de una población de inmunoglobulinas policlonales que se unen específicamente a un antígeno; (c) correlacionar la información de los

10 espectros de masas de los fragmentos peptídicos con la información pronosticada de los espectros de masas de las secuencias de ácidos nucleicos, dicha información pronosticada de los espectros de masas derivada de secuencias de aminoácidos pronosticadas codificadas por secuencias de nucleótidos de dichas secuencias de ácido nucleico para identificar secuencias de nucleótidos que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) que comprenden los fragmentos peptídicos; y (d) seleccionar entre las secuencias de nucleótidos o

15 secuencias de aminoácidos identificadas de las cadenas de inmunoglobulina (o regiones variables de las mismas) basándose en la cobertura de la secuencia de aminoácidos de las cadenas de inmunoglobulina o fragmentos de las mismas por los fragmentos peptídicos, para obtener secuencias de nucleótidos o secuencias de aminoácidos de cadenas pesadas o ligeras de inmunoglobulinas que se unen específicamente a un antígeno.

20 En algunas realizaciones, una cadena de inmunoglobulina pesada y una cadena ligera de inmunoglobulina (o regiones variables de las mismas) seleccionadas en la etapa (d) se ensamblan para crear una inmunoglobulina (o regiones variables de la misma) que se une específicamente al antígeno. En algunas realizaciones, las secuencias de nucleótidos o las secuencias de aminoácidos de las regiones variables de la cadena de inmunoglobulina obtenidas en la etapa (d) se sintetizan mediante técnicas recombinantes de biología molecular o técnicas de síntesis de genes antes del ensamblaje.

25 En algunas realizaciones, el método comprende adicionalmente: escrutar con un inmunoensayo la inmunoglobulina (o las regiones variables de la misma) creadas para confirmar que dicha inmunoglobulina (o regiones variables de la misma) se une específicamente al antígeno. En algunas realizaciones, el inmunoensayo se selecciona del grupo que consiste en un ensayo de citometría de flujo, un ensayo inmunoabsorbente ligado a enzima (ELISA), un ensayo de transferencia Western, un ensayo inmunohistoquímico, un ensayo de inmunofluorescencia, un radioinmunoensayo, 30 un ensayo de neutralización, un ensayo de unión, un ensayo de afinidad o un ensayo de inmunoprecipitación de proteínas o péptidos.

35 En algunas realizaciones, la selección de cadenas pesadas de inmunoglobulina y cadenas ligeras de inmunoglobulina en la etapa (d) se realiza basándose en la cobertura de la secuencia de aminoácidos de una porción de las cadenas (p.ej., la región variable o una región determinante de complementariedad) por los fragmentos peptídicos.

40 En otras realizaciones, la selección de cadenas pesadas o ligeras de inmunoglobulina en la etapa (d) se realiza basándose en la cobertura de la secuencia de aminoácidos de las cadenas de inmunoglobulina o fragmentos de las mismas por los fragmentos peptídicos, combinada con al menos un parámetro seleccionado del grupo que consiste en el número de péptidos únicos mapeados, la compartición de espectro, el recuento total de péptidos, el recuento de péptidos únicos, la frecuencia de las secuencias 20 de ácidos nucleicos codificantes y la relación clonal.

45 En diversas realizaciones, las secuencias de ácido nucleico y la información derivada de las secuencias de ácido nucleico (incluyendo, por ejemplo, las secuencias de nucleótidos, las secuencias de aminoácidos pronosticadas y los espectros de masas pronosticados) están ubicadas en una base de datos de material genético.

50 En algunas realizaciones, el animal del que se obtienen las secuencias de ácido nucleico es un animal expuesto al antígeno.

55 En algunas realizaciones, las secuencias de aminoácidos pronosticadas codificadas por dichas secuencias de ácido nucleico que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) de inmunoglobulinas múltiples del animal se obtienen: (1) aislando moléculas de ácido nucleico de glóbulos blancos de dicho animal; (2) amplificando la cadena de inmunoglobulina (o su región variable) que codifica las moléculas de ácido nucleico utilizando cebadores específicos para secuencias de polinucleótidos adyacentes a dicha cadena de inmunoglobulina (o región variable de la misma) que codifica las moléculas de ácido nucleico; (3) obteniendo secuencias de nucleótidos de dichas moléculas amplificadas de ácido nucleico que codifican las cadenas de inmunoglobulina (o regiones variables de las mismas) de múltiples inmunoglobulinas del animal; y (4) utilizando el código genético para

traducir las secuencias de nucleótidos a las secuencias de aminoácidos pronosticadas.

En algunas realizaciones, las secuencias de ácido nucleico son secuencias de ácido nucleico expresadas (p.ej., transcritas a ARN y/o traducidas a proteína en células del animal).

5 En algunas realizaciones, las secuencias de aminoácidos pronosticadas codificadas por las moléculas de ácido nucleico que codifican las cadenas de inmunoglobulina (o regiones variables de las mismas) de inmunoglobulinas múltiples del animal se obtienen: (1) aislando moléculas de ácido nucleico de glóbulos blancos de dicho animal; (2) secuenciando la cadena de inmunoglobulina (o su región variable) que codifica las moléculas de ácido nucleico utilizando cebadores específicos para secuencias de polinucleótidos adyacentes a dicha cadena de inmunoglobulina (o región variable de la misma) que codifica 15 moléculas de ácido nucleico para obtener las secuencias de nucleótidos que codifican las cadenas de inmunoglobulina (o regiones variables de las mismas) de múltiples inmunoglobulinas del animal; y (3) utilizando el código genético para traducir las secuencias de ácido nucleico a secuencias de aminoácidos. En algunas realizaciones, las moléculas de ácido nucleico son moléculas de ARN y dicha etapa de amplificación incluye una etapa de transcripción inversa inicial.

10 15 En algunas realizaciones, las secuencias polinucleotídicas adyacentes a las moléculas de ácido nucleico que codifican la cadena de inmunoglobulina (o su región variable) se seleccionan del grupo que consiste en genes de inmunoglobulina flanqueantes de ADN genómico, secuencias de polinucleótidos que codifican regiones constantes de cadenas de inmunoglobulina y secuencias de polinucleótidos que codifican regiones marco de cadena de inmunoglobulina.

20 En algunas realizaciones, la información de los espectros de masas pronosticados se obtiene utilizando un método que comprende las etapas de: (i) realizar un resumen teórico de secuencias de aminoácidos pronosticadas codificadas por las secuencias de nucleótidos de las moléculas de ácido nucleico con una o más proteasas y/o uno o más reactivos químicos de escisión de proteínas para generar fragmentos de péptidos virtuales; y (ii) crear espectros de masas pronosticados de dichos fragmentos peptídicos virtuales.

25 30 35 En algunas realizaciones, la información del espectro de masas observada de los fragmentos peptídicos se obtiene utilizando un método que comprende las etapas de: (i) aislar una población de inmunoglobulinas policlonales que se unen específicamente al antígeno; (ii) digerir la población con una o más proteasas y/o uno o más reactivos de escisión química de proteínas para generar fragmentos; y (iii) obtener información de los espectros de masas de dichos fragmentos peptídicos. En algunas realizaciones, la población de anticuerpos policlónicos se aísla utilizando un método que comprende las etapas de: (1) obtener fluido corporal o una fracción del mismo (p.ej., sangre, suero y/o plasma) de un animal; (2) hacer pasar el fluido corporal o una fracción del mismo sobre el antígeno en condiciones en las que las inmunoglobulinas que se unen específicamente al antígeno se unirán al antígeno; y (3) recoger dichas inmunoglobulinas unidas a dicho antígeno para obtener la población de inmunoglobulinas policlonales que se unen específicamente al antígeno. En algunas realizaciones, el antígeno se une a un soporte sólido (p.ej., el antígeno se une covalente o no covalentemente al soporte sólido). En algunas realizaciones, el soporte sólido puede ser una esfera (p.ej., una esfera de agarosa o magnético), una pared de una columna, o un fondo de una placa (p.ej., una placa de cultivo de tejido).

En algunas realizaciones, el animal es un animal previamente expuesto al antígeno. En algunas realizaciones, el animal expuesto previamente al antígeno es un animal previamente inmunizado con el antígeno.

40 45 50 En otro aspecto, la invención proporciona un método para obtener las secuencias de aminoácidos de la región variable de la cadena de inmunoglobulina de una inmunoglobulina que se une específicamente a un antígeno, que comprende: (a) proporcionar secuencias de ácido nucleico que codifican regiones variables de inmunoglobulina de múltiples inmunoglobulinas de un animal; (b) obtener información de los espectros de masas de los fragmentos peptídicos de las regiones variables de la cadena de inmunoglobulina de una población de inmunoglobulinas policlonales que se unen específicamente a un antígeno; (c) correlacionar la información de los espectros de masas de los fragmentos peptídicos con la información pronosticada de los espectros de masas de las secuencias de ácido nucleico, dicha información pronosticada de los espectros de masas derivada de las secuencias de aminoácidos pronosticadas codificadas por dichas secuencias de ácido nucleico, para obtener secuencias de aminoácidos de regiones variables de cadena de inmunoglobulina que comprende los fragmentos peptídicos; y (d) seleccionar entre las secuencias de nucleótidos identificadas o secuencias de aminoácidos de las regiones variables de cadena de inmunoglobulina basándose en la cobertura de la secuencia de aminoácidos de las regiones variables por los fragmentos peptídicos, para obtener secuencias de nucleótidos o secuencias de aminoácidos de regiones variables de inmunoglobulinas que se unen específicamente a un antígeno

55 En algunas realizaciones, el método comprende adicionalmente la etapa (e) escrutar las secuencias de aminoácidos de dichas regiones variables de la cadena de inmunoglobulina con un inmunoensayo para aislar una región variable de la cadena de inmunoglobulina de una inmunoglobulina que se une específicamente al antígeno. En algunas realizaciones, las secuencias de nucleótidos o secuencias de aminoácidos de las regiones variables de la cadena de inmunoglobulina obtenidas en la etapa (d) se sintetizan mediante técnicas recombinantes de biología molecular o técnicas de síntesis de genes antes de la etapa de escrutinio de la etapa (e). En algunas realizaciones, la región variable de la cadena de inmunoglobulina producida en la etapa (d) se ensambla con una segunda región variable de

5 cadena de inmunoglobulina para crear un dominio de unión a anticuerpo de una inmunoglobulina que se une específicamente al antígeno. En algunas realizaciones, el inmunoensayo se selecciona del grupo que consiste en un ensayo de citometría de flujo, un ensayo inmunoabsorbente ligado a enzimas (ELISA), un ensayo de transferencia Western y un ensayo inmunohistoquímico, un ensayo de inmunofluorescencia, un radioinmunoensayo, un ensayo de neutralización, un ensayo de unión, un ensayo de afinidad o un ensayo de inmunoprecipitación de proteínas o péptidos.

En algunas realizaciones, la región variable de la cadena de inmunoglobulina es una región variable de cadena pesada o una región variable de cadena ligera.

10 En un aspecto adicional, la descripción proporciona un método para crear un dominio de unión a antígeno de una inmunoglobulina que se une específicamente a un antígeno que comprende: (a) proporcionar secuencias de ácido nucleico que codifican regiones variables de cadena pesada de inmunoglobulina y regiones variables de cadena ligera de inmunoglobulinas múltiples de un animal; (b) obtener información de los espectros de masas de fragmentos peptídicos de cadenas pesadas de inmunoglobulina y cadenas ligeras de inmunoglobulina de una población de inmunoglobulinas policlonales que se unen específicamente a un antígeno; (c) correlacionar información de los  
15 espectros de masas de los fragmentos peptídicos con información de los espectros de masas pronosticados de las secuencias de ácidos nucleicos, dicha información pronosticada de los espectros de masas derivada de secuencias de aminoácidos pronosticadas codificadas por dichas secuencias de ácidos nucleicos, para obtener secuencias de nucleótidos o secuencias de aminoácidos de regiones variables de la cadena de inmunoglobulina que comprenden los fragmentos peptídicos; (d) seleccionar entre las secuencias de nucleótidos identificadas o las secuencias de  
20 aminoácidos de las regiones variables de la cadena de inmunoglobulina basándose en la cobertura de la secuencia de aminoácidos de los fragmentos peptídicos para obtener secuencias de nucleótidos o secuencias de aminoácidos de regiones variables de inmunoglobulinas que se unen específicamente a un antígeno; y (e) ensamblar una región variable de cadena pesada de inmunoglobulina seleccionada con una región variable de cadena ligera de inmunoglobulina seleccionada para crear un dominio de unión a antígeno de una inmunoglobulina que se une  
25 específicamente al antígeno.

En diversas realizaciones de todos los aspectos de la invención, el animal es un animal vertebrado. En diversas realizaciones, el animal es un mamífero. En algunas realizaciones, el animal es un ser humano. En algunas realizaciones, el animal es una rata, un conejo o un ratón. En algunas realizaciones, el animal es un ave, un animal domesticado, un animal de compañía, ganado, un roedor o un primate. En algunas realizaciones, el animal es un animal transgénico no humano, por ejemplo, un animal transgénico no humano que expresa secuencias de anticuerpos humanos y/o produce anticuerpos que son al menos parcialmente humanos.

30 En diversos aspectos, la descripción también proporciona una inmunoglobulina (o región variable de la misma), o una región variable de la cadena de inmunoglobulina o un dominio de unión a antígeno de una inmunoglobulina que se une específicamente a un antígeno aislado o creado de acuerdo con las diversas realizaciones no limitantes de la revelación. En diversas realizaciones, la inmunoglobulina (o región variable de la misma), o una región variable de la cadena de inmunoglobulina o un dominio de unión a antígeno de una inmunoglobulina que se une específicamente a un antígeno, están aisladas o son recombinantes. En diversas realizaciones, la descripción también proporciona un vehículo farmacéuticamente aceptable y una inmunoglobulina (o región variable de la misma), o una región variable de la cadena de inmunoglobulina o un dominio de unión a antígeno de una inmunoglobulina que se une  
35 específicamente a un antígeno aislado o creado de acuerdo con las diversos realizaciones no limitantes de la descripción.

40 En un aspecto adicional, la descripción proporciona un método para tratar un animal que tiene o se sospecha que tiene una enfermedad caracterizada por un antígeno de esa enfermedad, donde el método comprende administrar una cantidad eficaz de una composición de acuerdo con diversas realizaciones de la descripción, en donde el antígeno específicamente unido por la inmunoglobulina (o región variable de la misma), o la región variable de la cadena de inmunoglobulina o un dominio de unión a antígeno de la composición y el antígeno de la enfermedad son los mismos. En algunas realizaciones, el animal es un ser humano. En algunas realizaciones, el animal es un roedor, ganado, un animal domesticado, un animal de compañía o un primate.

45 En un aspecto adicional, la descripción proporciona un método para reducir la probabilidad de aparición en un animal de una enfermedad caracterizada por la presencia en el animal de un antígeno de enfermedad, donde el método comprende administrar una cantidad eficaz de una composición de acuerdo con diversas realizaciones de la descripción, en donde el antígeno específicamente unido por la inmunoglobulina (o región variable de la misma), o la  
50 región variable de la cadena de inmunoglobulina o un dominio de unión a antígeno de la composición y el antígeno de la enfermedad son los mismos. En algunas realizaciones, la composición comprende adicionalmente un coadyuvante. En algunas realizaciones, el animal es un ser humano. En algunas realizaciones, el animal es un roedor, ganado, un animal domesticado, un animal de compañía o un primate.

#### Breve descripción de los dibujos

La **Figura 1** es un diagrama esquemático de un anticuerpo que comprende dos cadenas pesadas y dos cadenas ligeras. Las dos cadenas pesadas se unen entre sí mediante dos enlaces disulfuro situados en la región bisagra del

anticuerpo. Cada cadena ligera está unida a una cadena pesada mediante un enlace disulfuro único. El sitio de unión al antígeno se crea en el extremo N terminal de las cadenas pesada y ligera.

La **Figura 2** es un diagrama esquemático que muestra un ejemplo de un método no limitante de diversas realizaciones de la invención. En este ejemplo, las muestras que comprenden linfocitos B (p.ej., una muestra de sangre o una muestra de tejido) y suero sanguíneo y/o plasma se recogen del mismo animal (p.ej., un ser humano, un ratón o un conejo). Las moléculas de ácido nucleico que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) se secuencian y estas secuencias de ácido nucleico se utilizan para generar información del espectro de masas teórica o pronosticada basada en las secuencias de aminoácidos pronosticadas codificadas por las secuencias de ácido nucleico. Los anticuerpos policlonales de los sueros sanguíneos se digieren proteolíticamente o se fragmentan químicamente y los fragmentos peptídicos resultantes se someten a análisis mediante espectrometría de masas. La información de las secuencias de ácido nucleico (p.ej., los espectros de masas s) se compara con la información del espectro de masas de los fragmentos de péptidos para identificar la secuencia de una cadena de inmunoglobulina (o dominio variable de la misma) de un anticuerpo. Este anticuerpo puede generarse de manera recombinante según los métodos convencionales.

La **Figura 3** es un diagrama esquemático que muestra otro ejemplo de un método no limitante de diversas realizaciones de la invención. En este ejemplo, los linfocitos B y el suero sanguíneo y/o el plasma se recogen del mismo animal (en este caso un conejo). A partir de los linfocitos B, el ARNm se extrae y se somete a secuenciación utilizando la máquina Genome Sequencer FLX System disponible comercialmente de 454 Life Sciences utilizando cebadores de secuenciación de genes específicos de inmunoglobulina. Esta información se utiliza para generar espectros de masas teóricos basados en las secuencias de aminoácidos pronosticadas. A partir del suero y/o plasma sanguíneos, los anticuerpos policlonales se aislan y se someten a digestión con proteasas y/o escisión con reactivos de escisión química de proteínas. Los fragmentos peptídicos resultantes se separan mediante cromatografía líquida, seguida de análisis de espectrometría de masas (MS/MS). Los espectros de masas de los fragmentos peptídicos se correlacionan con los espectros de masas teóricos de las secuencias de ácidos nucleicos para obtener las secuencias de aminoácidos de las cadenas de inmunoglobulinas incluyendo los fragmentos peptídicos. A continuación, se pueden ensamblar una cadena pesada y ligera para crear una inmunoglobulina recombinante clonando las secuencias de ácido nucleico que codifican las cadenas de inmunoglobulina en uno o varios vectores de expresión y expresando los vectores de expresión en una célula. La inmunoglobulina recombinante expresada se caracteriza adicionalmente.

La **Figura 4** es un diagrama esquemático que representa otro ejemplo de un método no limitante de varias realizaciones de la invención. En este ejemplo, se recogen del mismo animal (p.ej., un ser humano, un ratón o un conejo) una fuente no limitante de células B (p.ej., esplenocitos) y anticuerpos policlonales. Las moléculas de ácido nucleico se extraen de la fuente de células B y se someten a secuenciación de nueva generación (NGS) utilizando la máquina Roche 454 utilizando cebadores de secuenciación de genes específicos de inmunoglobulina. Esta información, que se puede poner en una base de datos de material genético, se puede utilizar para generar espectros de masas teóricos basados en las secuencias de aminoácidos pronosticadas codificadas por la secuencia de ácido nucleico. También del animal (p.ej., un ser humano, un ratón o un conejo), se cargan anticuerpos policlonales (o fragmentos peptídicos de los mismos) en el espectrómetro de masas para su análisis. Las secuencias de ácido nucleico se analizan utilizando reglas de Kabat para identificar las secuencias de las regiones variables (p.ej., una de las regiones CDR o regiones FR) de la secuencia. Las secuencias de los fragmentos peptídicos de los anticuerpos policlonales analizados se escrutan a continuación para identificar qué péptidos coinciden con la totalidad o parte de la región variable de una secuencia de aminoácidos pronosticada.

La **Figura 5** es una tabla que muestra secuencias de NGS de cadena pesada y ligera (es decir, secuenciación de nueva generación) que tenían una buena correlación de espectrometría de masas y péptido sobre la región variable. Algunos de estos péptidos aparecieron con bastante frecuencia (véase, p.ej., la referencia de cadena ligera Núm. G623FKB01A3GC7) y algunos tenían un alto recuento de frecuencia de secuencia de ácido nucleico (véase, p.ej., la referencia de cadena ligera G623FKB01AXJ1C). Las filas en negrita y cursiva representan las cadenas de inmunoglobulina que, al realizar la prueba, se encontró que contenían secuencias que se unen específicamente al antígeno (véanse los resultados de la prueba en la Figura 6).

La **Figura 6** es una tabla que muestra los resultados de los ensayos ELISA que prueban anticuerpos preparados utilizando un método no limitante de la invención escrutado frente a placas ELISA recubiertas con péptidos p-Erk. Las diferentes cadenas ligeras y pesadas que se muestran en la Fig. 5 se combinaron al azar entre sí. Como se puede observar a partir de la Fig. 6, una serie de emparejamientos dio como resultado anticuerpos que pudieron unirse específicamente a las placas recubiertas con p-ERK (los anticuerpos positivos se muestran sombreados).

La **Figura 7** es una representación fotográfica de un gel de agarosa que muestra los resultados de una reacción de RT-PCR (es decir, reacción en cadena de la polimerasa con transcriptasa inversa) de cadenas pesadas y cadenas ligeras kappa y lambda de ADNc generado a partir de esplenocitos de conejos inmunizados con antígeno p-MET.

La **Figura 8** es una tabla que muestra las secuencias de las cadenas de anticuerpos después de combinar los espectros de masas teóricos (es decir, pronosticados) derivados de las secuencias de ácidos nucleicos con datos de LC-MS/MS del anticuerpo purificado por afinidad. También se mostró la abundancia de la cadena de anticuerpos

basada en la frecuencia de NGS. Las cadenas representadas en cursiva se sintetizaron y ensamblaron en anticuerpo; y las cadenas en negrita y cursiva son aquellas que, tras la prueba con análisis de transferencia Western, se encontró que se unían específicamente al antígeno p-MET.

La Figura 9 es una representación fotográfica que muestra los resultados de un experimento de transferencia Western que analiza productos lisados preparados a partir de células Hela sin tratar (- calles en las tres transferencias) o tratados con factor de crecimiento humano (HGF) (+ calles en las tres transferencias) con dos anticuerpos de conejo diferentes generados utilizando un método no limitante de la invención (transferencias marcadas como 1 y 2) y con un anticuerpo de control (transferencias más a la izquierda). Después de los resultados positivos con la transferencia Western, se identificaron anticuerpos específicos de antígeno (emparejamiento de cadena pesada y ligera). Como se muestra, los anticuerpos identificados en ambas líneas 1 y 2 utilizaban la misma cadena pesada, pero tenían cadenas ligeras diferentes. Las secuencias de aminoácidos de las cadenas pesada y ligera de los dos anticuerpos de conejo se muestran debajo de los resultados de la transferencia Western, con las regiones CDR3 de las cadenas pesada y ligera subrayadas.

**Figuras 10a-e.** Purificación por afinidad de IgG policlonal de conejo específica de receptor de progesterona. (a) Se aisló IgG total del suero del conejo inmunizado con Proteína A y se purificó por afinidad adicionalmente sobre péptidos antigenicos inmovilizados mediante flujo por gravedad. Después de un lavado exhaustivo para reducir la IgG no específica, se usó una elución secuencial con pH progresivamente ácido para fraccionar la IgG policlonal específica de antígeno. Cada fracción se sometió a ensayo para determinar su actividad específica mediante transferencia Western a una concentración de anticuerpo correspondiente (21,5 ng/ml) para detectar PR A/B en productos lisados de células T47D (+). También se analizaron los productos lisados de control negativo de HT1080 (-), (b). La fracción con la actividad específica más alta, pH 1,8, se procesó con cuatro proteasas para análisis de LC-MS/MS, (c). Un espectro de MS/MS emparejado mediante SEQUEST al péptido triptico completo de la región V **GFALWGPGLTVVSSGOPK** (SEC ID NO: 305) que contiene CDRH3 (subrayado) con una XCorr de 5,560 y un ΔM (m/z observada - m/z esperada) de 0,39 ppm. (d). Los espectros de MS/MS se mapearon para a péptidos de la región V mediante SEQUEST y se filtraron a una FDR de ≤ 2%. Los péptidos de alta confianza se volvieron a mapear frente a la base de datos de la región V generada por NGS, teniendo en cuenta la proteasa utilizada para la preparación de muestras y haciendo un seguimiento del número total de péptidos, el número de péptidos únicos, la compartición del espectro y la cobertura de aminoácidos de toda la región V. Se seleccionaron secuencias de la región V de alta cobertura, expresadas como anticuerpos monoclonales, y se escrutaron para determinar la actividad deseada (e). La cobertura de identificación de secuencia de cadena pesada y ligera del clon F9. Las secuencias de la región V representadas, cuando se emparejan, se unen específicamente a PR A/B humana (véanse las Figuras 11a-e). Los aminoácidos mapeados para uno o más péptidos se muestran en negrita. Para maximizar la cobertura de la región V y tener en cuenta la composición de aminoácidos altamente variable, se utilizaron proteasas complementarias (quimotripsina, elastasa, pepsina, tripsina).

**Figuras 11a-e.** Identificación y caracterización de anticuerpos monoclonales funcionales contra el receptor de progesterona A/B. (un). El emparejamiento combinatorio de cadenas pesadas y ligeras produjo 12 clones reactivos mediante ELISA específicos de antígeno indicados en color amarillo. La secuencia de CDR3 se utiliza como un identificador: ✓ indica clones positivos para la Transferencia Western (véase la Fig. 11b), (b). Seis clones (F1 F9, HI, CI, F7 y H9) fueron específicos para la detección del receptor A/B de progesterona mediante transferencia Western. Los clones E6 (negativo para ELISA, negativo para Western) y H7 (positivo para ELISA, negativo para Western) se muestran como controles. +, T47D (positivo para PR A/B); -, MDA-MB-231 (negativo para PR A/B). Todos los anticuerpos se sometieron a ensayo a 21,5 ng/mL. (c). Comparación de la actividad específica del clon F9 con la mezcla polyclonal purificada por afinidad inmunohistoquímica. 0,4 µg/ml de F9 tuvieron específicamente tejido o líneas celulares (T47D y MCF-7) positivos para PR A/B, pero no una línea celular (MDA-MB-231) negativa para PR A/B. Se utilizaron 0,2 µg/ml de anticuerpo polyclonal como control positivo, (d). Análisis de citometría de flujo, Color azul, células T47D (línea celular positiva para receptor de progesterona A/B); Color negro, MDA-MB-231 (línea celular negativa para el receptor de progesterona A/B). Razón señal/ruido del anticuerpo polyclonal, 1,69; concentración, 3,7 µg/mL. Razón señal/ruido del anticuerpo monoclonal F9 = 36,4; concentración 0,5 µg/mL. (e) El análisis de microscopía de inmunofluorescencia confocal mostró un patrón específico de tinción nuclear en la línea celular positiva MCF-7 del receptor de progesterona A/B pero no en las células MDA-MB-231 a 0,46 µg/ml. No se incluyó anticuerpo primario como control de tinción de fondo. Los anticuerpos polyclonales también se utilizaron como comparación a una concentración de 1,85 µg/mL.

**Figuras 12a-d.** La caracterización del clon C3 del anticuerpo monoclonal anti-Lin28A, (a) El emparejamiento combinatorio de cadenas pesadas y ligeras produjo 5 clones reactivos mediante ELISA específicos de antígeno indicados mediante sombreado. V indica clones positivos para la transferencia Western. La secuencia de CDR3 se utiliza como un identificador, (b) El análisis de transferencia Western se realizó utilizando diversos productos lisados celulares positivos para Lin28A, NCCIT, NTERA, mMES e IGROV1. (c) El análisis de inmunofluorescencia confocal se realizó con células (HeLa) negativas para Lin28A y células (NTERA) positivas para Lin28A. (d) Análisis de citometría de flujo de anticuerpo monoclonal. Pico izquierdo, células HeLa (Lin28A -); pico derecho, células NTERA (Lin28A +). Las regiones V\* tenían la misma secuencia de CDR3 pero no secuencias de región V idénticas.

**Figuras 13a-c.** Identificación y caracterización de anticuerpos monoclonales de ratón funcionales contra fosfo-Erk. (a) Purificación de anticuerpos polyclonales de fosfo-Erk a partir de los sueros agrupados de tres ratones. Los sueros

agrupados, la IgG total purificada de proteína G de los sueros agrupados, la fracción no unida de la purificación de proteína G, y las fracciones de elución ácida de pH 3,5, 2,7 y 1,8 se analizaron mediante transferencia Western para determinar la especificidad de unión contra fosfo-Erk en producto lisado de células Jurkat +, Células Jurkat estimuladas con TPA; -, células Jurkat tratadas con U0126. (b) El emparejamiento combinatorio de cadenas pesadas y ligeras produjo 15 clones, indicados mediante sombreado, que son reactivos mediante ELISA de antígeno peptídico. V indica clones positivos para la transferencia Western (Véase (c)). La secuencia CDR3 se utiliza como un identificador. Para las secuencias de cadena pesada, la parte subrayada indica el final de la Región Marco 3. (c) Tres clones (C10, F10 y M3) fueron específicos para la detección de fosfo-Erk mediante transferencia Western. El clon C9 (positivo para ELISA, negativo para Western) se muestra como control. Todos los anticuerpos fueron sometidos a ensayo a 100 ng/mL.

#### **Descripción detallada**

Esta descripción se refiere a métodos y sistemas para obtener de forma rápida y precisa las secuencias de aminoácidos (y secuencias de ácidos nucleicos codificantes) de anticuerpos monoclonales que se unen específicamente a un antígeno de interés. Más específicamente, la presente metodología implica una investigación proteómica directa, basada en espectrometría de masas de anticuerpos policlónicos circulantes del suero de un animal, contra una base de datos de material genético que está compuesta de moléculas de ácido nucleico que codifican cadenas de inmunoglobulina completas o regiones variables. En realizaciones específicas, la base de datos de material genético se genera a partir del repertorio de células B de un animal (p.ej., el mismo animal cuyo suero se recogió para obtener los anticuerpos policlónicos) utilizando tecnologías de secuenciación de ácidos nucleicos. Por lo tanto, el enfoque actual implica esencialmente correlacionar (es decir, realizar comparaciones cruzadas o referencias cruzadas) la información de dos fuentes: información del espectro de masas de los anticuerpos policlónicos circulantes reales de un animal, e información (incluyendo, p.ej., espectros de masas pronosticados) de la base de datos de material genético. A continuación, se puede identificar una lista de secuencias de ADN de cadena pesada y ligera a partir de la base de datos de material genético que corresponde a los anticuerpos reales del suero. Dichas cadenas pesadas y ligeras se pueden expresar por pares para obtener anticuerpos monoclonales funcionales.

En algunas realizaciones, la presente metodología no requiere la inmortalización de células B, la clasificación de células individuales y la clonación molecular, o la presentación en fagos, y no implica el ensamblaje de secuencias de anticuerpos basadas en conjetas. Aprovechando las fortalezas de las tecnologías de espectrometría de masas y de secuenciación de ácidos nucleicos (como Next Generation DNA Sequencing o NGS), el enfoque de esta invención puede reducir significativamente la cantidad de tiempo necesaria para aislar las secuencias de anticuerpos monoclonales específicos de antígenos de una población polyclonal, permitiendo así una transición más rápida a anticuerpos recombinantes tales como anticuerpos completamente humanos o anticuerpos humanizados (p.ej., anticuerpos murinos humanizados) que pueden utilizarse terapéuticamente.

Además, la presente metodología es capaz de identificar anticuerpos raros que las tecnologías existentes probablemente no detectaron. Los autores de la presente invención han descubierto sorprendentemente que los anticuerpos individuales con especificidad muy selectiva (p.ej., un anticuerpo que se une específicamente a un residuo de tirosina fosforilado dentro de un polipéptido) pueden aparecer muy raramente dentro de una población polyclonal. Los métodos que dependen de las frecuencias de los ARNm que codifican anticuerpos y la amplificación de PGR pueden pasar por alto estos anticuerpos porque sus cadenas variables se producen con baja frecuencia. Por el contrario, la presente metodología utiliza, por ejemplo, análisis proteómicos basados en espectrometría de masas de fragmentos de péptidos reales derivados de una población de anticuerpos policlónicos, y por lo tanto no sufre los errores de frecuencia después de la amplificación de PGR.

Además, la presente metodología permite la creación rápida de nuevos anticuerpos específicos de antígeno que pueden no existir en la población polyclonal de partida. Por ejemplo, la molécula de inmunoglobulina creada que tiene las mayores cualidades deseadas (p.ej., mayor afinidad de unión (o KD más baja) para el antígeno o un isotipo deseado (p.ej., IgG2a) puede ser el resultado de una cadena ligera de un primer anticuerpo en la población polyclonal ensamblada con una cadena pesada de un segundo anticuerpo (es decir, diferente del primer anticuerpo) en la población polyclonal.

Los métodos descritos en la presente memoria tienen aplicaciones en inmunología básica y terapéutica. Por ejemplo, los métodos pueden proporcionar la base para comprender cuestiones centrales en el campo de la inmunología, incluyendo diversidad de anticuerpos séricos, dinámica, cinética, clonalidad y migración de las células B después de la exposición al antígeno. Los métodos también pueden utilizarse para buscar anticuerpos monoclonales humanos terapéuticamente relevantes de individuos inmunizados, infectados naturalmente o enfermos.

Como se demuestra en la presente memoria, la presente metodología se ha aplicado con éxito a varios antígenos diferentes en animales de laboratorio y en seres humanos, y ha conducido al aislamiento de anticuerpos monoclonales con actividades específicas de antígeno que recapitulan o superan las de los anticuerpos policlónicos purificados por afinidad originales encontrados en el suero de sujetos inmunizados.

Por consiguiente, esta descripción proporciona adicionalmente anticuerpos monoclonales recombinantes aislados específicos para un antígeno, que incluyen anticuerpos terapéuticos específicos para un antígeno de enfermedad, así como métodos terapéuticos para tratar una enfermedad basados en la administración de anticuerpos monoclonales terapéuticos.

- 5 Los diversos aspectos y realizaciones de la invención se describen con más detalle a continuación. Las patentes, las aplicaciones publicadas y la bibliografía científica a las que se hace referencia en la presente memoria establecen el conocimiento de los expertos en la técnica. Cualquier conflicto entre cualquier referencia citada en la presente memoria y las enseñanzas específicas de esta memoria descriptiva se resolverá a favor de esta última. Del mismo modo, cualquier conflicto entre una definición comprendida en la técnica de una palabra o frase y una definición de la palabra o frase tal como se enseña específicamente en esta memoria descriptiva se resolverá a favor de esta última.
- 10

#### **Definiciones**

Según se utiliza en la presente memoria, los siguientes términos tienen los significados indicados. Según se utiliza en esta memoria descriptiva, las formas singulares "un", "uno", "una", "el" y "la" específicamente también abarcan las formas plurales de los términos a los que se refieren, a menos que el contenido indique claramente lo contrario. El término "aproximadamente" se utiliza en la presente memoria para significar en torno a, en la región de, más o menos, o alrededor de. Cuando el término "aproximadamente" se utiliza junto con un intervalo numérico, modifica ese intervalo extendiendo los límites por encima y por debajo de los valores numéricos establecidos. En general, el término "aproximadamente" se utiliza en la presente memoria para modificar un valor numérico por encima y por debajo del valor indicado mediante una variación de 20%.

- 20 Por "péptido" o "fragmento peptídico" se entiende un polímero corto formado a partir de los residuos de aminoácidos individuales que se unen entre sí, donde la unión entre un residuo de aminoácido y el segundo residuo de aminoácido se denomina un enlace amida o un enlace peptídico. Un péptido comprende al menos dos residuos de aminoácidos. Un péptido se distingue de un polipéptido en que es más corto. Al menos dos péptidos, unidos entre sí por un enlace amida o un enlace peptídico entre el residuo de aminoácido C terminal de un péptido y el residuo de aminoácido N' terminal del segundo péptido, forman un polipéptido de acuerdo con diversas realizaciones de la descripción.
- 25

Por "polipéptido" se entiende un polímero largo formado a partir del residuo de aminoácido individual de unión, donde la unión entre un residuo de aminoácido y el segundo residuo de aminoácido se denomina enlace amida o enlace peptídico. Un polipéptido comprende al menos cuatro residuos de aminoácidos; sin embargo, múltiples polipéptidos se pueden unir entre sí a través de enlaces amida o peptídicos para formar un polipéptido incluso más largo.

- 30 Por "molécula de ácido nucleico" se entiende un polímero formado a partir de la unión de nucleótidos individuales (p.ej., desoxirribonucleótidos o ribonucleótidos) entre sí, donde el enlace entre un nucleótido y el otro nucleótido es un enlace covalente que incluye, por ejemplo, un enlace fosfodiéster. Por lo tanto, el término incluye, sin limitación, ADN, ARN e híbridos de ADN-ARN.
- 35

Por "secuencia de ácido nucleico" se entiende una secuencia de ácido nucleico (o secuencia de nucleótidos complementaria a la misma) que incluye nucleótidos que codifican toda o parte de una cadena de inmunoglobulina (p.ej., una cadena pesada o una cadena ligera). En algunas realizaciones, la secuencia de ácido nucleico es ADN genómico (p.ej., ADN exónico con o sin ADN intrónico). En algunas realizaciones, la secuencia de ácido nucleico es ADNc o alguna forma de ARN (p.ej., ARNhn, ARNm, etc.). En algunas realizaciones, la secuencia de ácido nucleico es una secuencia de ácido nucleico expresada que se transcribirá a una molécula de ácido nucleico (p.ej., ADN transcripto a ARN) o se traducirá a un polipéptido en una célula que contiene esa secuencia de ácido nucleico. Por consiguiente, una molécula de ácido nucleico expresada incluye, sin limitación, secuencias de ARNhn, ARNm, ADNc y exónicas genómicas. Por "complementario" en términos de moléculas de ácido nucleico simplemente significa que dos moléculas de ácido nucleico monocatenario contienen nucleótidos que formarán pares de bases de Watson-Crick convencionales para formar una molécula de ácido nucleico bicatenario, ya sea esa molécula bicatenaria ADN, ARN, o un híbrido de ADN-ARN.

- 40 Según se utiliza en la presente memoria, por "linfocito B" se entiende cualquier glóbulo blanco en el que la recombinación génica (o el reordenamiento del gen) ha comenzado a producirse en un locus que contiene un gen que codifica una cadena de inmunoglobulina. Por ejemplo, los genes de inmunoglobulina humana se producen en el cromosoma 14 (locus de cadena pesada), el cromosoma 2 (locus de cadena ligera kappa) y el cromosoma 22 (locus de cadena ligera lambda). Si un glóbulo blanco humano ha sufrido un evento de reordenación génica en un locus de la cadena de inmunoglobulina (p.ej., en el cromosoma 14, cromosoma 2 o cromosoma 22), esa célula se considera un linfocito B. Por consiguiente, los linfocitos B incluyen, sin limitación, células B, células pre-B, células pro-B incluyendo células pro-B tempranas (p.ej., donde las regiones D y J de los genes de la cadena pesada han experimentado reordenamiento pero el gen de la cadena ligera es de la línea germinal (es decir, no está reordenado)) y células pro-B tardías (p.ej., donde las regiones V, D y J del gen de la cadena pesada se reordenan pero el gen de la cadena ligera todavía es de la línea germinal y donde no se expresan proteínas de inmunoglobulina en la superficie celular), células pre-B incluyendo células pre-B grandes y células pre-B pequeñas, células B
- 45
- 50
- 55

inmaduras, células B activas, células B del centro germinal, células plasmáticas (incluyendo plasmablastos) y células B de memoria.

A lo largo de la memoria descriptiva y las reivindicaciones, los términos "anticuerpo" e "inmunoglobulina" se utilizan indistintamente y se pretende que incluyan moléculas de polipéptido de inmunoglobulina intactas de cualquier isotipo o subisotipo (p.ej., IgG, IgG1, IgG2, IgG2a, IgG2b, IgG3, IgG4, IgM, IgD, IgE, IgE1, IgE2, IgA) de cualquier especie de animal tal como primates (p.ej., seres humanos o chimpancés), roedores (p.ej., ratones o ratas), lagomorfos (p.ej., conejos o liebres), ganado (p.ej., vacas, caballos, cabras, cerdos y ovejas), peces (p.ej., tiburones), aves (p.ej., pollos) o camélidos (p.ej., camellos o llamas) o de animales transgénicos no humanos (p.ej., roedores) modificados genéticamente para producir anticuerpos humanos (véanse, p.ej., Lonberg et al., documento WO93/12227; Patente de Estados Unidos Núm. 5.545.806; Kucherlapati, et al., documento WO91/10741; Patente de Estados Unidos Núm. 6.150.584; documento US 2009/0098134; documento US 2010/0212035; documento US 2011/0236378; documento US 2011/0314563; documento WO2011/123708; documento WO2011/004192; documento WO2011/158009); fragmentos de dominios de unión a antígeno de los mismos, tales como Fab, Fab', F(ab')2; variantes de los mismos tales como scFv, Fv, Fd, dAb, scFv biespecíficos, diacuerpos, anticuerpos lineales (véase la Patente de Estados Unidos Núm. 5.641.870; Zapata et al., Protein Eng. 8 (10): 1057-1062, 1995); moléculas de anticuerpo de hebra sencilla; y anticuerpos multiespecíficos formados a partir de fragmentos de anticuerpos; y cualquier polipéptido que comprenda un dominio de unión que sea, o sea homólogo a, un dominio de unión a anticuerpo (definido en otra parte de la presente memoria). Los anticuerpos no limitantes de diversas realizaciones de la descripción incluyen, pero sin limitación, anticuerpos policlonales, monoclonales, monoespecíficos, poliespecíficos y fragmentos de los mismos, anticuerpos quiméricos que comprenden un dominio de unión a inmunoglobulina fusionado a otro polipéptido y anticuerpos humanizados tales como un anticuerpo no humano (p.ej., un anticuerpo de conejo) cuyos dominios constantes y/o FR se han reemplazado por dominios constantes y/o FR de un anticuerpo humano (véase, por ejemplo, las Patentes de Estados Unidos Núm.: 5.530.101; 5.585.089; 5.693.761; 5.693.762; 6.180.370; y 6.548.640). Los animales transgénicos no humanos modificados genéticamente para producir anticuerpos humanos (p.ej., al menos parcialmente humanos) están disponibles en Harbor Antibodies (Rotterdam, Países Bajos), Ablexis (San Francisco, CA), Kymab Ltd (Cambridge, RU), OMT, Inc. (Palo Alto, CA), Amgen (Thousand Oaks, CA), Medarex (Princeton, NJ) y Regeneron (Tarrytown, NY).

Los anticuerpos intactos de origen natural se componen de dos clases de cadenas polipeptídicas, cadenas ligeras y cadenas pesadas. Un anticuerpo no limitante de diversos aspectos de la descripción puede ser un anticuerpo intacto de cuatro cadenas de inmunoglobulina que comprende dos cadenas pesadas y dos cadenas ligeras. La cadena pesada del anticuerpo puede ser de cualquier isotipo, incluyendo IgM, IgG, IgA o IgD o subisotipo, incluyendo IgG1, IgG2, IgG2a, IgG2b, IgG3, IgG4, IgE1, IgE2, etc. La cadena ligera puede ser una cadena ligera kappa o una cadena ligera lambda. Por ejemplo, un solo anticuerpo IgG de origen natural (o intacto) comprende dos copias idénticas de una cadena ligera y dos copias idénticas de una cadena pesada de IgG. Las cadenas pesadas de todos los anticuerpos de origen natural, donde cada cadena pesada contiene un dominio variable (VH) y un dominio constante (CH, que comprende la región CH1, la región bisagra, la región CH2 y la región CHS), se unen entre sí a través de enlaces disulfuro múltiples dentro de sus dominios constantes para formar el "tronco" del anticuerpo. Las cadenas ligeras de todos los anticuerpos de origen natural, donde cada cadena ligera contiene un dominio variable (VL) y un dominio constante (CL), se unen cada una a través de su dominio constante a un dominio constante de cadena pesada a través de la unión disulfuro. En la Figura 1 se muestra un esquema de un anticuerpo de cuatro cadenas de inmunoglobulina (p.ej., un anticuerpo IgG). En la Figura 1, los tres dominios CH se muestran en color azul claro, el dominio VH único se muestra en color azul oscuro, el único dominio CL se muestra en color rosa claro y el dominio VL único se muestra en color rosa oscuro. Como se muestra en la Figura 1, los dominios VL y VH de las cadenas ligera y pesada, respectivamente, se unen para formar el dominio de unión del anticuerpo.

En algunas realizaciones, una cadena de inmunoglobulina intacta (p.ej., una cadena pesada o una cadena ligera) puede comprender en orden de 5' a 3' (para una secuencia de ácido nucleico que codifica la cadena) o desde el extremo amino al extremo carboxi (para la secuencia de aminoácidos de la cadena): un dominio variable y un dominio constante. El dominio variable puede comprender tres regiones determinantes de complementariedad (CDR, también denominadas regiones hipervariables o HV), con regiones marco intercaladas (FR). Los dominios variables de ambas cadenas ligeras y pesadas contienen tres regiones hipervariables intercaladas entre cuatro regiones marco más conservadas (FR), para una estructura de 5' (o N') - FR1, CDR1, FR2, CDR2, FR3, CDR3, FR4 3' (o C), con la región constante 3' (o C) a la región FR4. Las CDR forman bucles que comprenden la superficie principal de unión al antígeno del anticuerpo (véanse Kabat, E.A. et al., Sequences of Proteins of Immunological Interest, National Institutes of Health, Bethesda, Md., (1987) and Wu, T.T. and Kabat, E.A. (1970) J. Exp. Med. 132: 211-250 (1970)) adoptando en gran parte las cuatro regiones marco una conformación de hoja beta y formando las CDR bucles que conectan, y en algunos casos forman parte de, la estructura de hoja beta. Las CDR en cada cadena se mantienen cerca de las regiones marco y, con las CDR de la otra cadena, contribuyen a la formación del dominio de unión al antígeno.

Por "antígeno" se entiende una molécula diana (p.ej., un polipéptido o un hidrato de carbono) que puede unirse específicamente a un anticuerpo. La porción de un antígeno que está unida específicamente al anticuerpo se denomina "epítopo". Un "epítopo" es la porción más pequeña de una molécula diana capaz de unirse específicamente al dominio de unión a antígeno de un anticuerpo. El tamaño mínimo de un epítopo puede ser de aproximadamente tres a siete aminoácidos (p.ej., cinco o seis aminoácidos). Puede haber múltiples epítopos en un

único antígeno, por lo tanto, un único antígeno puede unirse específicamente a múltiples anticuerpos diferentes, todos los cuales se unen específicamente al antígeno (es decir, todos estos anticuerpos son anticuerpos específicos del antígeno) aunque cada anticuerpo individual se une específicamente a un epítopo diferente en el antígeno.

- 5 Por "antígeno de enfermedad" se entiende un antígeno que aparece en un animal durante un estado de enfermedad.
- 10 Por ejemplo, un antígeno viral (p.ej., un antígeno codificado por una molécula de ácido nucleico del material genético de un virus) es un antígeno de enfermedad en un animal infectado con ese virus. De forma similar, algunas enfermedades (p.ej., cáncer) se caracterizan por translocaciones génicas que producen proteínas químéricas (p.ej., BCR-ABL). Por lo tanto, una proteína BCR-ABL es un antígeno de enfermedad. Debe entenderse que un antígeno de enfermedad no se observa necesariamente solo en un animal que padece esa enfermedad.
- 15 Por "enfermedad" simplemente se entiende cualquier afección anormal que afecte a un animal. Los ejemplos no limitantes de enfermedades incluyen, sin limitación, enfermedades autoinmunitarias (p.ej., artritis reumatoide o diabetes tipo I), cáncer (p.ej., leucemia, cáncer de colon o cáncer de próstata, etc.), infecciones virales (p.ej., SIDA causado por infección del virus del VIH o de la varicela causada por la infección del virus varicela zóster), infección parasitaria (p. ej., esquistosomiasis o sarna) e infección bacteriana (p. ej., tuberculosis o difteria).
- 20 20 Por "se une específicamente" se entiende que una inmunoglobulina o anticuerpo interactúa con su antígeno (es decir, su antígeno específico), donde la interacción depende de la presencia de una estructura concreta (p.ej., un epítopo) sobre el antígeno; en otras palabras, el anticuerpo reconoce y se une a una estructura específica en lugar de a todas las moléculas o estructuras en general. Un anticuerpo que se une específicamente al antígeno se puede denominar "anticuerpo específico de antígeno" o "anticuerpo específico para el antígeno". En algunas realizaciones, un anticuerpo que se une específicamente al antígeno puede immunoprecipitar ese antígeno de una solución que contiene el antígeno así como otras moléculas (p.ej., un producto lisado celular). En algunas realizaciones, un anticuerpo que se une específicamente a su antígeno tiene una  $K_D$  por su antígeno de  $1 \times 10^{-6}$  M o menos. En algunas realizaciones, un anticuerpo que se une específicamente a su antígeno tiene una  $K_D$  por su antígeno de  $1 \times 10^{-7}$  M o menos, o una  $K_D$  de  $1 \times 10^{-8}$  M o menos, o una  $K_D$  de  $1 \times 10^{-9}$  M o menos, o una  $K_D$  de  $1 \times 10^{-10}$  M o menos, 25 de una  $K_D$  de  $1 \times 10^{-11}$  M o menos, de una  $K_D$  de  $1 \times 10^{-12}$  M o menos. En ciertas realizaciones, la  $K_D$  de un anticuerpo que se une específicamente a su antígeno para su antígeno específico está entre 1 pM a 500 pM, o entre 500 pM y 1 (iM, o entre 1 (iM a 100 nM, o entre 100 mM a 10 nM. Según se utiliza en la presente memoria, mediante el término " $K_D$ ", se pretende hacer referencia a la constante de disociación de una interacción entre dos moléculas (p.ej., la constante de disociación entre un anticuerpo y su antígeno específico).
- 30 30 "Región variable de una cadena de inmunoglobulina" o una "región variable de cadena de inmunoglobulina" es un polipéptido que comprende al menos una porción del dominio variable de un dominio pesado (es decir, el dominio VH) o de una cadena ligera (es decir, el dominio VL) de una inmunoglobulina, donde la porción de los dominios VL y VH forma un dominio de unión a antígeno de una inmunoglobulina (véase la Fig. 1). Por lo tanto, la región variable de una inmunoglobulina puede incluir, sin limitación, una sola CDR (p.ej., CDR1), dos CDR intercaladas con una sola FR (p.ej., CDR1, FR2 y CDR2), tres CDR intercaladas con dos FR (p. CDR1, FR2, CDR2, FRS y CDRS), o tres CDR flanqueadas por una o ambas FR1 y FR4 (p.ej., FR1, CDR1, FR2, CDR2, FRS, CDRS, FR4). En algunas realizaciones, la región variable de cadena de inmunoglobulina es la región en una de las cadenas pesada o ligera que, cuando se combina con la región variable de cadena de inmunoglobulina de la otra cadena (es decir, la cadena ligera o pesada) de la inmunoglobulina intacta, forma el dominio de unión al antígeno.
- 35 40 Por "dominio de unión a antígeno" se entiende la región de una única cadena pesada ensamblada con una sola cadena ligera en una inmunoglobulina, que conserva la actividad de unión específica del anticuerpo intacto para su antígeno específico. Por lo tanto, una inmunoglobulina IgG intacta, que comprende dos cadenas pesadas y dos cadenas ligeras, tiene dos dominios de unión a antígeno. Asimismo, la fragmentación de un anticuerpo intacto que conserva un enlace covalente entre la cadena pesada y la cadena ligera también dará como resultado un fragmento de inmunoglobulina que tiene un dominio de unión a antígeno. Por ejemplo, la digestión de una inmunoglobulina con la enzima papaina generará fragmentos F(ab), cada uno de los cuales tiene un único dominio de unión a antígeno. Por supuesto, el F(ab) completo no es el dominio de unión al antígeno; más bien, solo la porción del fragmento F(ab) que conserva la capacidad de unirse específicamente al antígeno es el dominio de unión al antígeno.
- 45 50 Los términos técnicos y científicos utilizados en la presente memoria tienen el significado comúnmente entendido por un experto en la técnica a la que pertenece la presente invención, a menos que se defina lo contrario. En la presente memoria se hace referencia a diversas metodologías y materiales conocidos por los expertos en la técnica. Los trabajos de referencia convencionales que establecen los principios generales de las tecnologías de anticuerpos y/o ADN recombinante incluyen Harlow y Lane, *Antibodies: A Laboratory Manual*, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, New York (1988); Sambrook et al. *Molecular Cloning: A Laboratory Manual*, 2<sup>a</sup> Ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N.Y. (1989); Coligan et al., *Current Protocols in Immunology*, John Wiley & Sons, New York, NY (1991-2010); Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, New York, NY (1987-2010); Kaufman et al., Eds., *Handbook of Molecular and Cellular Methods in Biology in Medicine*, CRC Press, Boca Ratón (1995); McPherson, Ed., *Directed Mutagenesis: A Practical Approach*, IRL Press, Oxford (1991). Los trabajos de referencia convencionales que establecen los principios generales de la farmacología incluyen Goodman & Gilman's *The Pharmacological Basis of Therapeutics*, 11<sup>a</sup> edición, McGraw Hill Companies Inc., Nueva York (2006)).
- 55 60

**Métodos para la obtención de secuencias de inmunoglobulinas específicas de antígeno**

En un aspecto, esta invención se refiere a un método para obtener las secuencias de aminoácidos y/o ácidos nucleicos de cadenas de inmunoglobulina (o regiones variables de las mismas) de una única inmunoglobulina de una población de anticuerpos policlonales.

- 5 De acuerdo con el presente método, se obtiene una población de anticuerpos policlonales de interés a partir de un animal y se fragmenta para generar fragmentos peptídicos que se analizan mediante espectrometría de masas. La información del espectro de masas observada a partir de los fragmentos peptídicos se correlaciona a continuación con la información del espectro de masas pronosticado derivada de una base de datos de material genético compuesta de secuencias de ácido nucleico que codifican cadenas pesadas y/o ligeras de inmunoglobulina
- 10 10 completas (o regiones variables de las mismas). Como resultado de tal correlación, pueden identificarse cadenas pesadas y/o ligeras de inmunoglobulina (o regiones variables de las mismas) a partir de la base de datos de material genético que corresponden a cadenas pesadas y/o ligeras de inmunoglobulina (o regiones variables de las mismas) de moléculas de inmunoglobulina dentro del población de anticuerpos policlonales de partida.

Los diversos aspectos del presente método se describen con más detalle a continuación.

15 **Población inicial de anticuerpos policlonales**

Las inmunoglobulinas que se unen específicamente a un antígeno de interés se pueden recoger de un animal, que incluye cualquier mamífero, tal como un ser humano. Las inmunoglobulinas se pueden recoger de una muestra de fluido corporal del animal incluyendo, por ejemplo, sangre, suero o plasma de la sangre, líquido cefalorraquídeo, líquido sinovial, líquido peritoneal, secreciones de la mucosa, lágrimas, secreciones nasales, saliva, leche y

20 secreciones genitourinarias. .

En algunas realizaciones, las inmunoglobulinas no necesitan provenir de un solo animal individual sino, más bien, pueden ser un cóctel de diferentes anticuerpos (monoclonales o policlónicos) tomados de diferentes individuos. En algunas realizaciones, las inmunoglobulinas se recogen de un animal transgénico no humano, por ejemplo, un animal transgénico no humano que expresa secuencias de anticuerpos humanos y/o produce anticuerpos que son al menos parcialmente humanos.

25 25 En algunas realizaciones, estas inmunoglobulinas son específicas para un antígeno de interés, ya sea porque el animal del que se recogen las inmunoglobulinas se inmunizó previamente con el antígeno, o porque el animal del que se recogen las inmunoglobulinas se expuso previamente a una afección por la cual era probable que el animal generara anticuerpos específicos del antígeno. En un ejemplo del último caso, el animal puede haber sido infectado con un virus (p.ej., virus de Epstein Barr), donde el antígeno de interés es la proteína EBNA1, que está codificada por el genoma del virus de Epstein Barr.

30 30 En diversas realizaciones, el animal cuyas inmunoglobulinas se recogen (es decir, se obtienen) es de la misma especie que el animal cuyas secuencias de ácido nucleico de linfocitos B se recogen para crear la base de datos de referencia. En algunas realizaciones, el animal cuyas inmunoglobulinas se recogen para la base de datos de péptidos y el animal cuyos ácidos nucleicos de linfocitos B se recogen para la base de datos de referencia son el mismo animal.

Como se muestra en la Figura 2, cuando el animal es el mismo animal, la sangre extraída del animal puede proporcionar tanto las secuencias de ácido nucleico (p.ej., de las células en la sangre) como los anticuerpos policlonales (p.ej., del suero o plasma de la sangre).

40 40 Las inmunoglobulinas recogidas del animal forman una población polyclonal de inmunoglobulinas, debido a que diferentes linfocitos B producen miembros de la población. Debe observarse que en dicha población polyclonal, no todos los anticuerpos individuales dentro de esa población polyclonal se unirán específicamente al mismo antígeno. De hecho, cada uno de los anticuerpos dentro de la población puede unirse a un antígeno diferente. Sin embargo, todavía se dice que esta población polyclonal se une específicamente a un antígeno concreto si al menos un anticuerpo individual, preferiblemente anticuerpos múltiples, de la población polyclonal se une a ese antígeno (véase, por ejemplo, el Ejemplo 3 a continuación). En otro ejemplo, algunos anticuerpos en la población polyclonal pueden unirse al antígeno con baja afinidad. Sin embargo, se dice que una población polyclonal se une específicamente a un antígeno si algunos (p.ej., al menos uno o más) de los anticuerpos en esa población se unen específicamente al antígeno.

45 50 Debe señalarse que con la frase "anticuerpo polyclonal (o inmunoglobulina) que se une específicamente a un antígeno" se entiende que dentro de la población polyclonal, al menos un anticuerpo se une específicamente al antígeno, sin embargo, ese anticuerpo no está necesariamente aislado de los otros anticuerpos dentro de la población polyclonal que no se unen específicamente al antígeno. Por supuesto, en algunas realizaciones, se une específicamente al antígeno más de un anticuerpo diferente dentro de la población polyclonal.

55 También debe observarse que las diferentes moléculas de anticuerpo son moléculas de anticuerpo producidas por una célula B diferente. Por ejemplo, después de recoger sueros, se puede aislar una población polyclonal de 1000

moléculas de anticuerpo de los sueros (p.ej., utilizando la adherencia de los anticuerpos a una columna de proteína A para aislar los anticuerpos de los otros componentes del suero). Dentro de esa población de 1000 moléculas de anticuerpo, 900 pueden ser idénticas (es decir, secretadas por la misma célula B) y, por lo tanto, en realidad solo hay 101 anticuerpos diferentes dentro de esa población policlonal. Con respecto a una población policlonal, si las 900

5 moléculas de anticuerpo idénticas se unen específicamente al antígeno, la población policlonal de 1000 moléculas de anticuerpo es un anticuerpo policlonal que se une específicamente al antígeno. De forma similar, si 5 moléculas de anticuerpo diferentes adicionales de las 100 moléculas de anticuerpo diferentes restantes también se unen específicamente al antígeno, la población policlonal de 1000 moléculas de anticuerpo es asimismo un anticuerpo policlonal que se une específicamente al antígeno.

10 No se necesita que la mayoría de las moléculas de anticuerpo dentro de una población policlonal se una específicamente a un antígeno para que se haga referencia a esa población como un "anticuerpo policlonal que se une específicamente al antígeno". Por ejemplo, si dentro de una población policlonal de 1000 moléculas de anticuerpo, incluso si solo 1 molécula de anticuerpo se une específicamente al antígeno y 999 moléculas de anticuerpo no lo hacen, esa población de 1000 moléculas de anticuerpo sigue siendo un "anticuerpo policlonal que 15 se une específicamente al antígeno" según se utiliza el término en la presente memoria. Obsérvese también que no se necesita que todos los anticuerpos en una población de anticuerpos policlonales se unan al mismo epítopo en el antígeno. Por ejemplo, una población policlonal puede ser específica para el antígeno cuando cada anticuerpo diferente dentro de la población se une específicamente a un epitopo diferente en el antígeno.

20 En diversas realizaciones de los métodos no limitantes de la invención, la población de inmunoglobulinas policlonales puede tener, por ejemplo, al menos dos inmunoglobulinas diferentes dentro de la población, o al menos tres, o al menos cinco, o al menos diez, o a al menos veinte, o al menos cincuenta, o al menos cien o al menos quinientas inmunoglobulinas diferentes dentro de la población.

25 La invención también contempla la recogida de una población policlonal de inmunoglobulinas a partir de los sobrenadantes del cultivo tisular de células B cultivadas in vitro (p.ej., cuando las secuencias de ácido nucleico se recogen de las propias células B). Por ejemplo, se puede recolectar una población de células B de un animal que ha sido sometido al virus de Epstein Barr. La población puede expandirse, por ejemplo, para enriquecer linfocitos B en la población en comparación con otros glóbulos blancos. A partir de este medio cultivado de estas células (en el que las células secretan los anticuerpos policlonales), se puede aislar la población policlonal de anticuerpos.

30 La población policlonal de inmunoglobulinas recogidas, ya sea de un animal o animales o de sobrenadantes de cultivo tisular de células B, se puede purificar primero antes de la digestión a fragmentos peptídicos. Por ejemplo, los anticuerpos policlonales recogidos pueden someterse a una columna de sefarosa con proteína A o proteína G, que puede separar anticuerpos de otras proteínas del suero sanguíneo, por ejemplo. Véanse, por ejemplo, la Figura 2 y la Figura 3. Alternativamente o adicionalmente, los anticuerpos policlonales recogidos se someten a purificación por afinidad de antígeno para enriquecer los anticuerpos con alta actividad específica. Aunque no es del todo necesario, una etapa de purificación, especialmente la purificación por afinidad de antígenos, puede reducir la complejidad de una 20 mezcla policlonal y finalmente reducir el número de posibles inmunoglobulinas candidatas falsas positivas o negativas. Los anticuerpos policlonales recogidos pueden concentrarse o someterse a intercambio de tampón o ambos, antes o después de la purificación.

40 En una realización ilustrativa, para recoger inmunoglobulinas que se unen específicamente a un antígeno de interés de un animal, se extrae sangre periférica del animal, y los anticuerpos de suero y/o plasma se recogen de acuerdo con métodos convencionales (p.ej., adherencia de los anticuerpos a proteína A). Los anticuerpos en suero y/o plasma se purifican o escrutinan para enriquecer las inmunoglobulinas que se unen específicamente al antígeno. Este escrutinio puede ser, por ejemplo, recubriendo una superficie en fase sólida (p.ej., una esfera de sefarosa o un fondo 45 de un pozo de plástico) con antígeno y haciendo pasar el suero y/o plasma sobre la superficie recubierta de antígeno en condiciones en las que se unirán las inmunoglobulinas que se unen específicamente al antígeno. Los anticuerpos unidos pueden tratarse con una proteasa (p.ej., papaína) o un reactivo de escisión química de proteína que corta específicamente cerca de la región bisagra de la inmunoglobulina para eliminar las porciones de Fc no adherentes. Después de enjuagar las proteínas séricas y/o plasmáticas que no se unen (incluyendo inmunoglobulinas no 50 específicas), se pueden recoger las inmunoglobulinas específicas de antígeno y enriquecer así sus cantidades en comparación con los anticuerpos que no se unen específicamente al antígeno.

#### ***Espectros de masas observados a partir de los anticuerpos policlonales recogidos***

Para obtener espectros de masas observados (es decir, reales), los anticuerpos policlonales recogidos (o fragmentos de los mismos) se analizan mediante métodos de análisis de proteínas (p.ej., espectrometría de masas, cromatografía líquida, etc.).

55 En algunas realizaciones, la información del espectro de masas observado se obtiene a partir de fragmentos de péptidos que se generan a partir de los anticuerpos policlonales. Los anticuerpos policlonales se pueden fragmentar, por ejemplo, con una o más proteasas, y/o un reactivo de escisión química de proteínas, tal como bromuro de cianógeno.

Se sabe que ciertas proteasas dividen sus sustratos en sitios específicos. La Tabla 1 proporciona una lista no exhaustiva de proteasas utilizadas comúnmente y sus sitios de escisión (en código de aminoácidos de 3 letras).

**Tabla 1**

Proteasa	Sitio de escisión
Tripsina	escinde después de (es decir, en el lado carboxilo de) Arg o Lys, a menos que estén seguidos de Pro
Quimotripsina	escinde después de Phe, Trp o Tyr, a menos que estén seguidos de Pro
Elastasa	escinde después de Ala, Gly, Ser o Val, a menos que estén seguidos de Pro.
Endoproteinasa Lys-C	escinde después de Lys
Pepsina	escinde después de Phe o Leu.
Termolisina	escinde antes de Ile, Met, Phe, Trp, Tyr o Val, a menos que estén precedidos de Pro.
Endopeptidasa V8 (sobrenombre Glu-C)	escinde después de Glu.

- 5 En las Tablas 11.1.1 y 11.1.3 de Riviere y Tempst se proporciona una lista más completa de proteasas que pueden utilizarse para digerir proteínas a fragmentos más pequeños (Riviere LR, Tempst P. Enzymatic digestion of proteins in solution. Curr Protoc Protein Sci. mayo de 2001; Capítulo 11:Unidad 11.1. PubMed PMID: 18429101).

10 En realizaciones específicas, se utilizan proteasas múltiples (es decir, dos o más) (p.ej., independientemente o juntas) para digerir los anticuerpos policlonales para maximizar la cobertura de la región V y dar cuenta de 15 composiciones de aminoácidos altamente variables de inmunoglobulinas. Por ejemplo, se puede utilizar una combinación de quimotripsina, elastasa, pepsina y tripsina, como se ilustra en el Ejemplo 7 de esta memoria. En algunas realizaciones, una proteasa o proteasas se eligen sobre la base de que no escinden dentro de regiones de CDR3 pronosticadas basándose en el análisis de las moléculas de ácido nucleico en la base de datos de material genético.

15 Las proteínas pueden digerirse a fragmentos más pequeños que son susceptibles de espectrometría de masas por medio de tratamiento con reactivos de escisión química de proteínas concretas en lugar de enzimas proteolíticas. Véase por ejemplo el capítulo 3 de Sequencing of Proteins and Peptides, Laboratory Techniques in Biochemistry and Molecular Biology, Vol. 9. Elsevier 1989 de G. Allen. Dichos reactivos de escisión química de proteínas incluyen, sin limitación, bromuro de cianógeno, BNPS-skatole, ácido o-yodosobenzoico, ácido diluido (p.ej., HQ diluido), etcétera. 20 Por ejemplo, las proteínas se pueden escindir a residuos de Met con bromuro de cianógeno, a residuos de Cys después de la cianilación, a residuos de Trp con BNPS-skatole o ácido o-yodoobenzoico, etc. También se pueden generar fragmentos de proteína por exposición a ácido diluido, por ejemplo, HCl. Un ejemplo del uso de la hidrólisis ácida parcial para determinar las secuencias de proteínas mediante espectrometría de masas es proporcionada por 25 Zhong et al. (Zhong H, et al., J. Am. Soc. Mass Spectrom. 16(4):471-81, 2005. PubMed PMID: 15792716). Zhong et al., más arriba utilizaron hidrólisis ácida asistida por microondas con ácido trifluoroacético al 25% en agua para fragmentar la bacteriorodopsina para la secuenciación mediante espectrometría de masas. Véase también Wang N, y Li L., J. Am. Soc. Mass. Spectrom. 21(9): 1573-87, 2010. PubMed PMID: 20547072.

30 Las proteínas pueden fragmentarse para hacerlas más aptas para la espectrometría de masas mediante tratamiento con una proteasa, mediante tratamiento con más de una proteasa combinada, mediante tratamiento con un reactivo de escisión química, mediante tratamiento con más de un reactivo de escisión química combinado, o mediante tratamiento con una combinación de proteasas y reactivos de escisión química. Las reacciones se pueden producir a temperaturas elevadas o presiones elevadas. Véase por ejemplo Lopez-Ferrer D, et al., J. Proteome. Res. 7(8):3276-81, 2008. PubMed PMID: 18605748; PubMed Central PMCID: PMC2744211. Se puede permitir que la fragmentación se complete hasta que la proteína se escinda en todos los enlaces que el reactivo de digestión es 35 capaz de escindir; o las condiciones de digestión pueden ajustarse para que la fragmentación no llegue a completarse deliberadamente, para producir fragmentos más grandes que puedan ser particularmente útiles para descifrar las secuencias de la región variable del anticuerpo; o las condiciones de digestión pueden ajustarse para que la proteína se digiera parcialmente a dominios, por ejemplo, como se hace con ADN polimerasa I de *E. coli* para preparar el fragmento de Klenow. Las condiciones que pueden variarse para modular el nivel de digestión incluyen 40 duración, temperatura, presión, pH, ausencia o presencia de reactivo desnaturizante de proteínas, el

desnaturalizante de proteínas específico (p.ej., urea, guanidina HCl, detergente, detergente escindible con ácido, metanol, acetonitrilo, otros disolventes orgánicos), la concentración de desnaturalizante, la cantidad o concentración de reactivo de escisión o su razón en peso relativa a la proteína que se vaya a digerir, entre otras cosas.

- 5 En algunas realizaciones, el reactivo (es decir, la proteasa o los reactivos de escisión química de proteínas) utilizado para escindir las proteínas es un reactivo completamente no específico. Utilizando dicho reactivo, no se pueden realizar limitaciones en el extremo N del péptido, el extremo C del péptido o ambos extremos N y C. Por ejemplo, en los diversos métodos descritos en la presente memoria se puede utilizar una secuencia parcialmente proteolizada que está limitada a tener un sitio de escisión tríptico en un extremo de la secuencia peptídica o en el otro, pero no en ambos.
- 10 Los fragmentos peptídicos resultantes se pueden detectar y analizar utilizando una HPLC acoplada a un espectrómetro de masas a partir del cual se generan los espectros de masas observados. Este método se puede denominar enfoque de proteómica de "bottom up (de abajo hacia arriba)", donde los componentes del proteoma se separan e identifican después de reducir las proteínas a péptidos relativamente pequeños, por ejemplo, de 3 a 45 residuos de longitud.
- 15 En otras realizaciones, se puede emplear un enfoque proteómico alternativo "top down (de arriba hacia abajo)" para obtener los espectros de masas observados, lo que implica el análisis de espectrometría de masas de proteínas intactas o fragmentos de proteínas grandes o dominios de proteínas o polipéptidos grandes. Por ejemplo, para identificar las partes de las regiones variables del anticuerpo que confieren reconocimiento de antígeno específico a una molécula de anticuerpo policlonal concreta, es útil secuenciar grandes porciones de las regiones variables para identificar sus CDR, mediante el análisis directo de fragmentos lo suficientemente grandes como para que las CDR permanezcan unidas entre sí.

Para una revisión que describa las estrategias "de abajo hacia arriba" y "de arriba hacia abajo", véase Han X, Aslanian A, Yates JR 3<sup>a</sup>. Mass spectrometry for proteomics. *Curr Opin Chem Biol.* 2008 Oct;12(5):483-90. Review. PubMed PMID: 18718552; PubMed Central PMCID: PMC2642903.. Para una revisión reciente de la proteómica de arriba hacia abajo aplicada para determinar las secuencias de anticuerpos, véase Zhang Z. et al., *Mass Spectrom Rev.* 2009 Jan-Feb;28(1): 147-76. Review. PubMed PMID: 18720354. Para un artículo reciente que muestra la secuenciación exhaustiva de un anticuerpo monoclonal mediante proteómica de arriba hacia abajo, véase Tsybin et al., *Anal Chem.* 2011, 21 de oct. PubMed PMID: 22017162.

- 30 En algunas realizaciones del método no limitante anterior, aunque las inmunoglobulinas específicas de antígeno se unen a la superficie recubierta con antígeno, las inmunoglobulinas pueden digerirse con papaína o pepsina para generar fragmentos F(ab) y F(ab)2, respectivamente. Dado que la totalidad de una región variable de cadena de inmunoglobulina se localiza en una cadena de un fragmento F(ab), este pretratamiento con papaína y/o pepsina enriquecerá las regiones variables de la cadena de inmunoglobulina. Después de enjuagar las partes no unidas de las inmunoglobulinas, se pueden recoger las regiones variables de la cadena de inmunoglobulina.
- 35 Después del paso de los fragmentos de inmunoglobulina a través del espectrómetro de masas, se generarán numerosos espectros de masas observados. Sin embargo, dado el número potencialmente grande de diferentes inmunoglobulinas dentro de una población policlonal, cada una con una secuencia de aminoácidos diferente, que se analizan con el espectrómetro de masas, los espectros de masas observados resultantes serán difíciles de volver a ensamblar en una región variable de la cadena de inmunoglobulina funcional. En los métodos de diversas realizaciones de la invención, debido a que está disponible la secuencia de ácido nucleico subyacente, no hay necesidad de ensamblar los datos de los espectros de masas observados. En cambio, el espectro de masas observado de un único fragmento peptídico puede correlacionarse con los espectros de masas pronosticados de la secuencia de ácido nucleico para obtener la secuencia de aminoácidos (y nucleótidos subyacentes) de toda la cadena de inmunoglobulinas (o región variable de la misma) de una inmunoglobulina que se une específicamente a un antígeno de una población de inmunoglobulina policlonal de partida. Esta etapa de correlación se describe adicionalmente a continuación.

50 Además de la información del espectro de masas s, la información adicional derivada de los fragmentos peptídicos de los anticuerpos policloniales es útil en diversas realizaciones de la invención. Esta información incluye, sin limitación, la masa de cada péptido, la longitud (en residuos de aminoácidos) de cada péptido, el espectro de masas observado de cada péptido (p.ej., a partir de la espectrometría de masas en tandem tal como el espectro MS2 o MS3), la razón masa a carga de cada péptido, la carga iónica de cada péptido, el perfil cromatográfico de cada péptido y la secuencia de aminoácidos de cada péptido.

#### **Análisis de espectrometría de masas**

- 55 En los métodos de esta invención, la información del espectro de masas puede obtenerse mediante análisis de espectrometría de masas de inmunoglobulinas recogidas o fragmentos generados a partir de ellas. Un espectrómetro de masas es un aparato capaz de medir la razón masa a carga ( $m/z$ ) de moléculas individuales ionizadas, lo que permite a los investigadores identificar compuestos desconocidos, cuantificar compuestos conocidos y dilucidar la estructura y las propiedades químicas de las moléculas. En algunas realizaciones, se

comienza el análisis de espectrometría de masas aislando y cargando una muestra en el aparato. Una vez cargada, la muestra se vaporiza y a continuación se ioniza. Posteriormente, los iones se separan de acuerdo con su razón masa a carga a través de la exposición a un campo magnético. En algunas realizaciones, se utiliza un aparato de sector, y los iones se cuantifican de acuerdo con la magnitud de la desviación de la trayectoria del ión mientras pasa a través del campo electromagnético del aparato, que está directamente correlacionado con la razón masa-carga de iones. En otras realizaciones, las razones masa a carga de iones se miden cuando los iones pasan a través de cuadripolos, o en base a su movimiento en trampas de iones tridimensionales o lineales u Orbitrap, o en el campo magnético de un espe ́ctrómetro de masa de resonancia ión-ciclotrón con transformada de Fourier. El aparato registra la abundancia relativa de cada ion, que se utiliza para determinar la composición química, molecular y/o isotópica de la muestra original. En algunas realizaciones, se utiliza un aparato de tiempo de vuelo, y se utiliza un campo eléctrico para acelerar iones a través del mismo potencial, y mide el tiempo que le lleva a cada ion alcanzar el detector. Este enfoque depende de que la carga de cada ion sea uniforme para que la energía cinética de cada ion sea idéntica. La única variable que influye en la velocidad en este escenario es la masa, viajando los iones más ligeros a mayores velocidades y alcanzando el detector más rápidamente en consecuencia. Los datos resultantes se representan en un espectro de masas o un histograma, razón de intensidad frente a masa-carga, con picos que representan compuestos o fragmentos ionizados.

Para obtener datos de espectros de masas de una muestra de proteína, la muestra se carga en el aparato y se ioniza. La ionización puede realizarse mediante, p.ej., ionización por electropulverización y desorción/ionización por láser asistida por matriz ("MALDI"). Véase, p.ej., Zenobi, "Ion Formation in MALDI Mass Spectrometry", 17 Mass Spectrometry Review, 337 (1998). La caracterización de proteínas se puede realizar de una de dos maneras, de arriba hacia abajo o de abajo hacia arriba. El enfoque de arriba hacia abajo implica la ionización de proteínas intactas o fragmentos de proteína más grandes. Véase, p.ej., Allison 15 Doerr, "Top-down Mass Spectrometry", 5 Nature Methods, 24 (2008). El enfoque de abajo hacia arriba implica digerir enzimática o químicamente la proteína hasta sus péptidos constituyentes utilizando una proteasa. Véase Biran Chait, "Mass Spectrometry: Bottom-Up or Top- Down?", 6 Science 65 (2006). Los péptidos resultantes se introducen en el aparato y finalmente se identifican mediante huella peptídica o espectrometría de masas en tandem.

En algunas realizaciones, el análisis de espectrometría de masas puede combinarse con un fraccionamiento cromatográfico (p.ej., cromatografía líquida).

Los datos de los espectros de masas útiles en esta invención se pueden obtener mediante huella peptídica. La huella peptídica implica introducir la masa observada de un espectro de la mezcla de péptidos generados por digestión proteolítica en una base de datos y correlacionar las masas observadas con las masas pronosticadas de fragmentos que surgen de las digestiones de proteínas conocidas in silico. Las masas conocidas correspondientes a masas de muestra proporcionan evidencia de que la proteína conocida está presente en la muestra analizada.

Los datos de los espectros de masas se pueden obtener mediante espectrometría de masas en tandem. En algunas realizaciones, la espectrometría de masas en tandem típicamente utiliza la disociación inducida por colisión, que hace que los iones peptídicos colisionen con el gas y se fragmenten (p.ej., debido a la energía vibratoria impartida por la colisión). El proceso de fragmentación produce productos de escisión que se rompen en los enlaces peptídicos en varios sitios a lo largo de la proteína. Las masas de los fragmentos observados pueden combinarse con una base de datos de masas pronosticadas para una de muchas secuencias peptídicas dadas, y se puede pronosticar la presencia de una proteína. Véase, p.ej., Eng, 5 An Approach to Correlate Tandem Mass Spectral Data of Peptides with Amino Acid Sequences in a Protein Database, JASMS, 976 (1994).

En otra realización, la espectrometría de masas en tandem se realiza mediante disociación inducida por colisión de alta energía (HCD), que en algunos espe ́ctrómetros de masas muestra fragmentos de iones producto más cerca de los extremos peptídicos que la disociación inducida por colisión. Véase Olsen JV, Macek B, Lange O, Makarov A, Homing S, Mann M. Higher-energy C-trap dissociation for peptide modification analysis. Nat. Methods. 2007 Sep;4(9):709-12. Epub 2007, 26 de agosto. PubMed PMID: 17721543.

En otra realización, la espectrometría de masas en tandem se realiza mediante disociación por transferencia de electrones (ETD), que se basa en reacciones ion-ion donde un ion químico reactivo distinto dona un radical a un ion peptídico, que a continuación se fragmenta rápidamente para formar iones producto. Véase Mikesh LM, Ueberheide B, Chi A, Coon JJ, Syka JE, Shabanowitz J, Hunt DF. La utilidad de la espectrometría de masas ETD en el análisis proteómico. Biochim Biophys Acta. 2006 Dic; 1764(12): 1811-22. Epub 2006, 30 de octubre. Revisión. PubMed PMID: 17118725; PubMed Central PMCID: PMC1853258. Ciertos métodos de fragmentación, tales como ETD, son particularmente adecuados para estrategias proteómicas "de arriba hacia abajo". Otros mecanismos de fragmentación son específicos de ciertos mecanismos de ionización, por ejemplo, tal como la descomposición posterior a la fuente (PSD) es compatible con la ionización por desorción láser asistida por matriz (MALDI) y también es adecuada para estrategias proteómicas "de arriba hacia abajo".

#### **Base de datos de material genético**

De acuerdo con la presente invención, la información de espectros de masas observada a partir de la población de inmunoglobulinas policlonales de partida se correlaciona con la información de espectros de masas pronosticada

derivada de una base de datos de material genético, para obtener las secuencias de aminoácidos (y nucleótidos subyacentes) de cadenas de inmunoglobulinas (o regiones variables de la misma) de inmunoglobulinas de la población de inmunoglobulina policlonal de partida.

5 Según se utiliza en la presente memoria, una base de datos de material genético incluye secuencias de ácido nucleico que codifican una pluralidad de cadenas de inmunoglobulina (o regiones variables de las mismas). Por lo tanto, la información que puede obtenerse o derivarse de dicha base de datos de material genético incluye, por ejemplo, la información de la secuencia de nucleótidos de cada molécula de ácido nucleico, la longitud (en nucleótidos) de cada molécula de ácido nucleico, la información de secuencia de aminoácidos de los polipéptidos o péptidos codificados por cada molécula de ácido nucleico, la masa de un polipéptido o péptido codificados por cada molécula de ácido nucleico, la longitud (en residuos de aminoácidos) de un polipéptido o péptido codificados por cada molécula de ácido nucleico, la información de los espectros de masas de polipéptidos o péptidos codificados por cada molécula de ácido nucleico (p.ej., una información de espectros de masas pronosticada basada en la secuencia de aminoácidos del polipéptido o péptido), y la secuencia de aminoácidos de un polipéptido o péptido codificados por cada molécula de ácido nucleico.

10 15 En algunas realizaciones de la invención, la base de datos de material genético contiene información genética de secuencias de ácido nucleico que codifican cadenas de inmunoglobulina completas (y no solo las regiones variables de las mismas). En algunas realizaciones, las secuencias de ácido nucleico son expresadas (es decir, transcritas a ARN y/o traducidas a proteína) por la célula de la que derivan dichas secuencias. En realizaciones específicas, la base de datos de material genético incluye secuencias de ácido nucleico expresadas que codifican regiones variables de cadena de inmunoglobulina de múltiples inmunoglobulinas de un animal. En algunas realizaciones, la base de datos de material genético contiene al menos cien secuencias de ácido nucleico expresadas diferentes. En otras realizaciones, la base de datos de material genético contiene al menos mil secuencias de ácido nucleico expresadas diferentes.

20 25 Las moléculas de ácido nucleico que codifican cadenas de inmunoglobulina (o las regiones variables de las mismas) se pueden obtener fácilmente a partir de una población de células (p.ej., glóbulos blancos periféricos) que contienen linfocitos B. En algunas realizaciones, las moléculas de ácido nucleico se obtienen a partir de esplenocitos o células mononucleares, tales como células mononucleares de sangre periférica (PBMC). En algunas realizaciones, los linfocitos B son de un animal no tratado (p.ej., un animal que no ha sido expuesto al antígeno al que se busca un anticuerpo específico de antígeno). En algunas realizaciones, el animal no tratado se ha expuesto a muy pocos antígenos (p.ej., un animal criado en un ambiente estéril o libre de patógenos). En algunas realizaciones, el animal no tratado es un animal típico que ha sido expuesto a antígenos típicos, pero que no ha sido expuesto al antígeno de elección.

30 35 En algunas realizaciones, el animal del que se obtienen las moléculas de ácido nucleico que codifican las cadenas de inmunoglobulina (o las regiones variables de las mismas) es un animal que se ha expuesto previamente al antígeno. Por ejemplo, el animal puede ser un animal inmunizado con el antígeno (p.ej., el antígeno mezclado con un coadyuvante o un antígeno acoplado a un vehículo inmunogénico tal como la hemocianina de lapa californiana (KLH)), puede ser un animal infectado con un patógeno que comprende el antígeno (p.ej., un animal infectado con el virus VIH cuando el antígeno de elección es el antígeno p24 del VIH), o puede de otra manera haber sido expuesto previamente al antígeno. En algunas realizaciones, el animal es un ave (p.ej., un pollo o pavo) o un mamífero, tal como un primate (p.ej., un ser humano o un chimpancé), un roedor (p.ej., un ratón, un hámster o una rata), un lagomorfo (p.ej., un conejo o liebre), un camélido (p.ej., un camello o una llama), o un mamífero domesticado, tal como un animal de compañía (p.ej., un gato, un perro o un caballo), o ganado (p.ej., una cabra, oveja o una vaca).

40 45 Se debe entender que no se necesita que las secuencias de ácido nucleico de los diversos aspectos y realizaciones de la invención provengan de un solo animal. Por ejemplo, algunas de las secuencias de ácido nucleico de diversas realizaciones de la invención pueden provenir de un animal previamente expuesto a un antígeno, y algunas de las secuencias de ácido nucleico pueden provenir de animales sin tratamiento previo. En algunas realizaciones de la invención, las secuencias de ácido nucleico son de animales de una sola especie. Por ejemplo, cuando hay múltiples animales de los que se obtienen secuencias de ácidos nucleicos, todos esos animales pueden ser de la misma especie (p.ej., todos son conejos o todos son seres humanos). En algunas realizaciones, las secuencias de ácido nucleico se obtienen de animales de una sola especie. En otras realizaciones, pueden obtenerse secuencias de ácido nucleico de más de una especie de animal. Por ejemplo, las secuencias de ácido nucleico se pueden obtener de ratones y ratas, y los espectros de masas pronosticados basados en estas secuencias se pueden utilizar para correlacionar o comparar con la información del espectro de masas real del fragmento peptídico de anticuerpos policloniales para crear una inmunoglobulina (o región variable, dominio de unión a antígeno, o cadena de los mismos) que se une específicamente al antígeno. En algunas realizaciones, las secuencias de ácido nucleico se obtienen de animales de un único género (p.ej., todos los animales son hembras).

50 55 60 El animal del que se recogen los anticuerpos policloniales y el animal del que se recogen las secuencias de ácido nucleico pueden ser el mismo animal, o la misma especie de animal, o animales singénicos (p.ej., ambos son ratones Balb/c), o de animales del mismo sexo (p. ej., ambos son hembras). Los espectros de MS2 de los componentes de unión a antígeno de los anticuerpos policloniales se pueden correlacionar así con los espectros teóricos de MS2 derivados de las secuencias de ácidos nucleicos obtenidas de un animal, con el fin de identificar las

secuencias de ácidos nucleicos que codifican anticuerpos de unión a antígeno.

También se debe entender que las secuencias de ácido nucleico y los anticuerpos policlonales se pueden recoger de células de un animal en el que las células se cultivaron *in vitro* después de la eliminación del animal y antes de la recolección de los anticuerpos policlonales (p.ej., a partir del sobrenadante o medio de cultivo de las células cultivadas) y la recolección de las secuencias de ácido nucleico de las células. Esta etapa de cultivo es útil, por ejemplo, para expandir o enriquecer linfocitos B en comparación con otras células sanguíneas o tisulares (p.ej., para enriquecer linfocitos B sobre glóbulos rojos o células epiteliales). El número de secuencias de ácido nucleico individuales utilizadas para crear espectros de masas teóricos en las diversas realizaciones de la invención es ilimitado. Por ejemplo, se pueden obtener cinco y diez o cincuenta, o cien, o mil, o un millón, o un billón, o un billón o más de secuencias diferentes de ácido nucleico y utilizarse para crear espectros de masas teóricos. Las secuencias de ácido nucleico pueden provenir de cualquier fuente, y pueden ser de una combinación de fuentes. Por ejemplo, las secuencias de ácido nucleico pueden obtenerse secuenciando moléculas de ácido nucleico expresadas que codifican regiones variables de cadena de inmunoglobulina (o toda la cadena completa de inmunoglobulina que incluye las regiones variables y la región constante) como se describe en la presente memoria. Las secuencias de ácidos nucleicos también pueden obtenerse a partir de ADN genómico que puede o no haber experimentado una recombinación V(D)J completa. Las secuencias de ácido nucleico también pueden obtenerse a partir de fuentes disponibles públicamente. Por ejemplo, se conocen numerosas secuencias de aminoácidos (y secuencias de nucleótidos) de regiones variables de cadena de inmunoglobulina (y polinucleótidos que codifican la misma) de múltiples especies de animales (véanse, por ejemplo, las siguientes publicaciones de patentes PCT y de Estados Unidos: US 20100086538; WO 2010/097435; US 20100104573; US 7.887.805; US 7.887.801; US 7.846.691; US 7.833.755; US 7.829.092).

Los linfocitos B de los que se obtienen secuencias de ácido nucleico pueden ser de cualquier fuente sanguínea o tisular incluyendo, sin limitación, médula ósea, sangre fetal, hígado fetal, sitios de inflamación (p.ej., fluido sinovial que rodea articulaciones inflamadas en pacientes reumatoideos), tumores (p.ej., linfocitos infiltrantes de tumor), sangre periférica, en los ganglios linfáticos, en las placas de Peyer, en las amígdalas y en el bazo o en cualquier órgano linfoide. En algunas realizaciones, se puede procesar el tejido completo (p.ej., médula ósea o ganglio linfático) (p.ej., células separadas entre sí y lisadas), eliminar el material genético y secuenciar las moléculas de ácido nucleico que codifican cadenas de inmunoglobulina (o regiones variables de las mismas).

En algunas realizaciones, se enriquecen en linfocitos B tejidos o una población de células (p.ej., sangre periférica) que los contienen antes de aislar el material genético de los linfocitos B. De acuerdo con diversas realizaciones de la invención, los métodos para enriquecer en linfocitos B un animal son bien conocidos. Los linfocitos B se pueden encontrar en muchos órganos y zonas del cuerpo incluyendo, sin limitación, médula ósea, sangre fetal, hígado fetal, sitios de inflamación (p.ej., fluido sinovial que rodea articulaciones inflamadas en pacientes reumatoideos), tumores (p.ej., linfocitos infiltrantes de tumores), sangre periférica, en los ganglios linfáticos y en el bazo. A partir de estas muestras de tejido (p.ej., sangre periférica o el bazo de un animal), los glóbulos blancos pueden aislarse de acuerdo con métodos convencionales (p.ej., utilizando los reactivos Ficoll-Paque PLUS o Ficoll-Paque PREMIUM disponibles comercialmente de GE Healthcare, Piscataway, NJ, de acuerdo con las instrucciones del fabricante). Los propios linfocitos B se pueden aislar posteriormente a partir de otros glóbulos blancos utilizando, por ejemplo, marcadores de superficie celular encontrados en los linfocitos B. Los marcadores de superficie de linfocitos B incluyen, sin limitación, cadenas de inmunoglobulina expresadas en la superficie celular (p.ej., cadena ligera lambda, cadena ligera kappa y cadena pesada tal como IgM o IgG). Los marcadores de superficie celular de linfocitos B adicionales incluyen, sin limitación, CD21, CD27, CD138, CD20, CD19, CD22, CD72 y CD79A. Otros marcadores de superficie de células B linfocitarias adicionales incluyen, sin limitación, CD38, CD78, CD80, CD83, DPP4, FCER2, IL2RA, TNFRSF8, CD24, CD37, CD40, CD74, CD79B, CR2, IL1R2, ITGA2, ITGA3, MS4A1, ST6GAL1, CD1C, CD138 y CHST10.

Estos marcadores de superficie de linfocitos B pueden utilizarse secuencialmente para el enriquecimiento en linfocitos B. Por ejemplo, los anticuerpos específicos para los marcadores de superficie de un linfocito B (p.ej., CD19) pueden ser acoplados a esferas magnéticas (p.ej., Dynabeads comercialmente disponibles de Invitrogen Corp., Carlsbad, CA) y las células que se adhieren a las esferas (p.ej., células CD19 positivas) aisladas a partir de células que no expresan CD19. Las células CD19 positivas pueden enriquecerse adicionalmente en linfocitos B mediante, por ejemplo, clasificación mediante citometría de flujo de células que expresan cadenas de inmunoglobulina en su superficie celular. Estos linfocitos B de enriquecimiento se pueden aislar de este modo para su uso en los métodos de diversas realizaciones de la invención.

Los linfocitos B específicos de antígeno también pueden purificarse directamente utilizando el antígeno deseado como cebo para aislar células B que expresan el receptor de células B específico de antígeno (inmunoglobulina de membrana). Por ejemplo, las células B se pueden agregar a una columna a la que se adhiere el antígeno deseado. Las células B específicas de antígeno fluirán a través de la columna más lentamente que las células B no específicas u otras células (p.ej., glóbulos rojos, macrófagos, etc.). De este modo se puede producir un enriquecimiento en células B específicas de antígeno utilizando este método.

Los linfocitos B que han experimentado enriquecimiento o no de un animal (p.ej., enriquecimiento mediante diversos métodos) también se pueden someter a cultivo celular *in vitro* durante 1 o 2 o 3 o 4 o más días antes de la extracción

del ácido nucleico. Tal cultivo *in vitro* puede expandir el número de linfocitos B y así enriquecerlos sobre células de linfocitos no B. En un ejemplo no limitante, los linfocitos B humanos aislados con CD27 pueden someterse a diversos cócteles de citoquina y moléculas extracelulares (tales como, pero sin limitación, medio acondicionado de células T activado o cualquier combinación de crecimiento de células B y/o factores de diferenciación) antes de la extracción del ácido nucleico para estimular el crecimiento y/o la diferenciación de los linfocitos B antes de la extracción del ácido nucleico de los linfocitos B. También se pueden añadir otras moléculas biológicas a los medios de cultivo tisular durante el cultivo *in vitro* para ayudar al crecimiento, la diferenciación y/o la inmunización *in vitro*, y/o cualquier combinación de los anteriores.

A partir de estos linfocitos B aislados, sometidos a enriquecimiento o estimulados, se pueden extraer las secuencias de ácidos nucleicos (p.ej., ADN genómico, ARNn, ARNm, etc.) utilizando métodos convencionales (p.ej., extracción con fenol: cloroformo, véase Ausubel et al., más arriba). Este ácido nucleico puede someterse a análisis de secuenciación utilizando una variedad de métodos para la secuenciación.

En algunas realizaciones, las secuencias de ácido nucleico pueden secuenciarse directamente a partir del material biológico (es decir, sin amplificarse antes de la secuenciación). Los servicios y reactivos para la secuenciación directa a partir de secuencias de ácido nucleico están disponibles comercialmente, por ejemplo, de Helicos BioSciences Corp. (Cambridge, MA). Por ejemplo, Helicos' True Single Molecule Sequencing permite la secuenciación directa de ADN, ADNc y ARN. Véanse también las Patentes de Estados Unidos Núm. 7.645.596; 7.037.687, 7.169.560; y las publicaciones Harris et al., *Science* 320: 106-109, 2008; Bowers et al., *Nat. Methods* 6: 493-494, 2009; y Thompson y Milos, *Genome Biology* 12: 217, 2011.

En otras realizaciones, las secuencias nucleicas se amplifican (p.ej., mediante reacción en cadena de la polimerasa (PCR)) antes de obtener la información de la secuencia.

En un ejemplo no limitante, se utiliza un cebador de PCR oligo dT para la RT-PCR. En otro ejemplo no limitante, la RT-PCR específica de gen se realiza utilizando los cebadores de PCR descritos en la presente memoria, tales como los cebadores de ratón de fusión específica 454, los cebadores de fusión de cadena de inmunoglobulina de conejo 454 o los cebadores de la región ligera variable y pesada variable. En otro ejemplo, los cebadores de PCR contra poblaciones de cadena pesada y cadena ligera en un ratón tienen secuencias establecidas en la Publicación PCT Núm. WO2010/097435.

Con o sin enriquecimiento en células B, los materiales genéticos purificados (ADN o ARNm) se pueden amplificar (p.ej., mediante PCR o RT-PCR) siguiendo procedimientos convencionales (véase, p.ej., Ausubel et al., más arriba) para preparar una biblioteca antes de la secuenciación NGS.

Los linfocitos B aislados mencionados anteriormente por diversos medios también pueden someterse a encapsulación de células individuales utilizando un método en la técnica tal como encapsulación en emulsión de aceite o mediante un aparato comercial tal como tecnología RainDance (RainDance Technologies, Inc., Lexington, MA). Estos linfocitos B encapsulados se pueden fusionar a continuación con un reactivo de RT-PCR de células individuales apropiado (p.ej., el reactivo comercializado por Qiagen, como Núm. Cat. 210210) con los cebadores de amplificación apropiados para generar productos de PCR de cadena Pesada y Ligera unidos de cada célula B individual. La PCR de ligación o solapamiento se conoce en el campo y se pone en práctica rutinariamente para diversas aplicaciones de biología molecular para unir 2 piezas de ADN en una (véanse, por ejemplo, Meijer P.J. et al., *J. Mol. Biol.* 358(3): 764-72, 2006 para la PCR de solapamiento). Este enfoque permite la preservación e identificación del emparejamiento análogo durante la secuenciación.

#### **Métodos de secuenciación de ADN**

Los métodos para la secuenciación de ADN que son bien conocidos y están disponibles generalmente en la técnica se pueden utilizar para obtener las secuencias de ácidos nucleicos de las diversas realizaciones de la invención. Los métodos pueden emplear enzimas tales como el fragmento Klenow de la ADN polimerasa I, SEQUENASE® (US Biochemical Corp, Cleveland, Ohio), polimerasa Taq (Invitrogen), polimerasa de T7 termoestable (Amersham, Chicago, 111), ADN ligasa (p.ej., de T4) o combinaciones de polimerasas recombinantes y exonucleasas de corrección de pruebas tales como el Sistema de Amplificación ELONGASE comercializado por Gibco BRL (Gaithersburg, MD). El procedimiento puede automatizarse con máquinas tales como Hamilton Micro Lab 2200 (Hamilton, Reno, Nev.), Peltier Thermal Cycler (PTC200; MJ Research, Watertown, Mass.) y los secuenciadores de ADN ABI 377 (Applied Biosystems).

Los métodos no limitantes para secuenciar moléculas de ácido nucleico y así generar secuencias de ácido nucleico (p.ej., para poblar una base de datos de material genético) de diversas realizaciones de la invención incluyen el método de Sanger (véanse, p.ej. Sanger et al., *Nature* 24: 687-695, 1977), el método de Maxam-Gilbert (véase, por ejemplo, Maxam y Gilbert, *Proc. Natl. Acad. Sci. USA* 74: 560-564, 1977), y pirosecuenciación (véase, por ejemplo, Ronaghi et al., *Science* 281 (5375): 363, 1998 y Ronaghi et al., *Analytical Biochemistry* 242 (1): 84, 1996). La pirosecuenciación, otro método de secuenciación no limitante que se puede utilizar para obtener secuencias de polinucleótidos, utiliza luciferasa para generar luz para la detección de los nucleótidos individuales (dATP, dTTP, dGTP o dCTP, colectivamente "dNTP") agregados al ADN naciente, y los datos combinados se utilizan para generar

lecturas de secuencia.

En algunas realizaciones, las secuencias de ácido nucleico se obtienen utilizando secuenciación profunda o secuenciación de nueva generación. Un paso limitante de la velocidad en la secuenciación de ADN convencional surge de la necesidad de separar los polímeros de ADN terminados aleatoriamente mediante electroforesis en gel.

5 Los dispositivos de secuenciación de nueva generación evitan esta limitación, p.ej., colocando físicamente moléculas de ADN en superficies sólidas y determinando la secuencia de ADN *in situ*, sin necesidad de separación en gel. Estas técnicas de secuenciación de alto rendimiento permiten secuenciar numerosas moléculas de ácido nucleico en paralelo.

10 Por lo tanto, se pueden secuenciar simultáneamente miles o millones de diferentes moléculas de ácido nucleico (véanse Church, G.M., *Sci. Am.* 294 (1): 46-54, 2006; Hall, N., *J. Exp. Biol.* 210 (Pt. 9): 1518-1525, 2007; Schuster et al. *Nature Methods* 5(1): 16-18, 2008; y MacLean et al. *Nature Reviews Microbiology* 7: 287-296, 2009) Existe una variedad de diferentes métodos y máquinas para realizar la secuenciación de nueva generación, cualquiera de los cuales puede utilizarse para generar secuencias de ácido nucleico. Véase Lin et al., *Recent Patents on Biomedical Engineering* 1:60-67, 2008 para una descripción general de numerosas tecnologías de secuenciación de nueva generación.

15 Por ejemplo, Shendure, J. et al., *Science* 309(5741): 1728-32, 2005 y la Publicación de Patente de Estados Unidos Núm. 20070087362, describe el método de secuenciación de nueva generación de colonias que utiliza un método de secuenciación basado en ligación (véase también la Patente de Estados Unidos Núm. 5750341). La tecnología SOLiD disponible comercialmente de Applied Biosystems (una compañía de LifeTechnolgies Corp., Carlsbad, CA) emplea secuenciación por ligación. Utilizando la tecnología SOLiD, una biblioteca de fragmentos de ADN a secuenciar se amplifica mediante PCR en emulsión, y de los múltiples fragmentos de la biblioteca, se unirá una única especie de fragmento a una única esfera magnética (las denominadas esferas clonales). Los fragmentos unidos a las esferas magnéticas tendrán una secuencia adaptadora PI universal unida de modo que la secuencia inicial de cada fragmento sea conocida e idéntica. A continuación se seleccionan los cebadores que hibridan con la secuencia adaptadora PI dentro del molde de la biblioteca. Un conjunto de cuatro sondas di-base marcadas fluorescentemente compiten por la unión al cebador de secuenciación. La especificidad de la sonda di-base se logra interrogando cada 1<sup>a</sup> y 2<sup>a</sup> base en cada reacción de ligación.

20 Otro método de secuenciación de nueva generación es el de Margulies et al., *Nature* 437: 376-380, 2005 y Patentes de Estados Unidos Núm. 7.211.390; 7.244.559; y 7.264.929, que describen una versión paralelizada de 30 pirosecuenciación que amplifica el ADN dentro de las gotas de agua en una solución de aceite (PCR en emulsión), conteniendo cada gota un único molde de ADN unido a una sola esfera recubierta con cebador. Utilizando la máquina de secuenciación (la máquina Genome Sequencer FLX System disponible comercialmente de 454 Life Sciences, una compañía Roche, Branford, CT), los adaptadores de oligonucleótidos se ligan a moléculas de ácido nucleico fragmentadas y a continuación se inmovilizan a la superficie de microesferas microscópicas antes de la 35 amplificación por PCR en una emulsión de gotas de aceite. A continuación se aíslan las esferas en múltiples pocillos de volumen de picolitros, cada uno con una sola esfera, enzimas de secuenciación y dNTP. La incorporación de un dNTP a la cadena complementaria libera pirofosfato, que produce ATP, que a su vez genera luz que a continuación puede registrarse como una imagen para su análisis.

40 La Patente de Estados Unidos Núm. 7.115.400 describe otra técnica para la amplificación en fase sólida de moléculas de ácido nucleico. Esto permite que una gran cantidad de diferentes secuencias de ácido nucleico se agrupen y amplifiquen simultáneamente. Esta tecnología se incorpora al sistema Genome Analyzer disponible comercialmente en Solexa (Illumina, Inc.). En esta tecnología, las moléculas de ADN se unen primero a los cebadores en un portaobjetos y se amplifican de modo que se formen colonias clonales locales (amplificación puente). Se añaden cuatro tipos de ddNTP y los nucleótidos no incorporados se eliminan mediante lavado. A 45 diferencia de la pirosecuenciación, el ADN solo puede extenderse un nucleótido a la vez. Una cámara toma imágenes de los nucleótidos marcados fluorescentemente, a continuación el colorante junto con el bloqueador terminal 3' se elimina químicamente del ADN, lo que permite un próximo ciclo.

50 Las secuencias polinucleotídicas que codifican las regiones variables de la cadena de inmunoglobulina pueden extenderse utilizando una secuencia parcial de nucleótidos y empleando diversos métodos conocidos en la técnica para detectar secuencias aguas arriba tales como promotores y elementos reguladores. Por ejemplo, un método que puede emplearse, PCR de "sitio de restricción", utiliza cebadores universales para recuperar secuencias desconocidas adyacentes a un locus conocido (Sarkar, G., *PCR Methods Applic.* 2: 318-322 (1993)). En particular, el ADN genómico se amplifica primero en presencia de cebador para la secuencia conectora y un cebador específico para la región conocida. Los cebadores ilustrativos son los descritos en el Ejemplo 4 de la presente memoria. Las secuencias amplificadas se someten a continuación a una segunda ronda de PCR con el mismo cebador conectar y otro cebador específico interno al primero. Los productos de cada ronda de PCR se transcriben con una ARN polimerasa apropiada y se secuencian utilizando transcriptasa inversa.

55 La PCR inversa también se puede utilizar para amplificar o extender secuencias utilizando cebadores divergentes basados en una región conocida (Triglia et al. *Nucleic Acids Res.* 16: 8186 (1988)). Los cebadores pueden diseñarse utilizando el soporte lógico OLIGO 4.06 Primer Analysis (National Biosciences Inc., Plymouth, Minnesota) u otro

programa apropiado, que tenga 22-30 nucleótidos de longitud, para tener un contenido de GC de 50% o más, y para la reasociación con la secuencia diana a temperaturas de aproximadamente 68-72°C. El método utiliza varias enzimas de restricción para generar un fragmento adecuado en la región conocida de un gen. El fragmento se circulariza a continuación mediante ligadura intramolecular y se utiliza como molde de la PCR.

- 5      Otro método que puede utilizarse es la PCR de captura que implica la amplificación mediante PCR de fragmentos de ADN adyacentes a una secuencia conocida en ADN cromosómico artificial humano y de levadura (Lagerstrom et al., PCR Methods Applic. 1: 111 - 119 (1991)). En este método, también pueden utilizarse digestiones y ligaciones con enzimas de restricción múltiples para colocar una secuencia de doble hebra modificada genéticamente en una porción desconocida de la molécula de ADN antes de realizar la PCR. Otro método que puede utilizarse para recuperar secuencias desconocidas es el descrito por Parker et al. Nucleic Acids Res. 19: 3055 - 3060 (1991)). Además, se pueden utilizar PCR, cebadores anidados y bibliotecas PROMOTERFINDER® para pasear en el ADN genómico (Clontech, Palo Alto, California). Este proceso evita la necesidad de examinar bibliotecas y es útil para encontrar uniones intrón/exón.

- 10     Se debe entender que el ácido nucleico de los linfocitos B puede escrutar adicionalmente para detectar las moléculas de ácido nucleico que codifican inmunoglobulinas antes de la secuenciación. Para hacer esto, pueden emplearse cebadores específicos para moléculas de ácido nucleico que codifican inmunoglobulina (o específicas para regiones adyacentes a la misma).

- 15     Según se utiliza en la presente memoria, por "cebador" se entiende una secuencia de ácido nucleico que puede tener al menos aproximadamente 15 nucleótidos, o al menos aproximadamente 20 nucleótidos, o al menos aproximadamente 30 nucleótidos, o al menos aproximadamente 40 nucleótidos de longitud. Se pretende que un cebador específico para una molécula de ácido nucleico concreto incluya un cebador que hibride con una porción de la molécula de ácido nucleico bajo condiciones de reasociación mediante PCR (p.ej., 60°C durante treinta segundos). En algunas realizaciones, un cebador específico para una molécula concreta de ácido nucleico es aquel que es complementario a esa molécula de ácido nucleico.

- 20     25    Los cebadores utilizados para secuenciar la secuencia de ácido nucleico pueden denominarse cebadores de secuenciación. Los cebadores utilizados para la amplificación de una secuencia de ácido nucleico diana mediante la reacción en cadena de la polimerasa (PCR) también pueden denominarse cebadores de PCR o cebadores de amplificación (véase la descripción de PCR, por ejemplo, de Sambrook et al., más arriba y Ausubel et al., más arriba).

- 30     35    40    45    En un ejemplo no limitante para obtener secuencias de ácido nucleico de acuerdo con diversas realizaciones de la invención, el ácido nucleico total de los linfocitos B puede convertirse en monocatenario (p.ej., calentando el ácido nucleico a 94-980°C durante al menos un minuto. El ácido nucleico de hebra sencilla se puede pasar a continuación sobre un soporte sólido (p.ej., una columna o gel) al que se adhieren cebadores de hebra sencilla que son específicos para regiones no variantes de moléculas de ácido nucleico que codifican inmunoglobulinas o regiones no codificantes adyacente al mismo (p.ej., promotores, potenciadores y/o intrones del gen de inmunoglobulina). Algunos ejemplos no limitantes para estas regiones no variantes de inmunoglobulinas incluyen la región constante de la cadena pesada, y la región constante de las cadenas ligeras, y el régión FR1 de la cadena pesada o de la cadena ligera. Se permite que el ácido nucleico hibride con los cebadores unidos a soporte en fase sólida, y se elimina el ácido nucleico que no se hibrida. Despues de la eliminación el ácido nucleico hibrido (que está enriquecido en moléculas de ácido nucleico que codifican inmunoglobulinas) se libera de los cebadores mediante, por ejemplo, adición de calor o aumento de la concentración de EDTA en el tampón.

- 50     55    En otra realización de la invención, independientemente de si el ácido nucleico de los linfocitos B está enriquecido en moléculas de ácido nucleico que codifican inmunoglobulina, las moléculas de ácido nucleico que codifican inmunoglobulina pueden amplificarse para aumentar su número de copias. Esta amplificación puede realizarse, por ejemplo, mediante amplificación por PCR utilizando cebadores específicos para regiones no variantes de moléculas de ácido nucleico que codifican inmunoglobulina o regiones no codificantes adyacentes a las mismas.

- En todos los métodos anteriores para obtener secuencias de ácido nucleico de acuerdo con las diversas realizaciones de la invención, se entenderá que los cebadores (p.ej., cebadores de secuenciación o de PCR) utilizados para generar las secuencias de ácido nucleico que codifican la región variable de la cadena de inmunoglobulina pueden ser universales (p.ej., cola poliA) o pueden ser específicos para las secuencias que codifican inmunoglobulinas.

- En algunas realizaciones, el material de partida a partir del cual se obtiene la información de la secuencia de ácido nucleico que codifica el gen de inmunoglobulina es ADN genómico. Por ejemplo, si las regiones variables de la cadena de inmunoglobulina son de seres humanos, los cebadores (p.ej., cebadores de secuenciación y/o cebadores de PCR) se pueden seleccionar para que sean idénticos o hibriden a un promotor de gen de cadena de inmunoglobulina. Por ejemplo, la secuencia del genoma humano es conocida. Dado que el gen que codifica la cadena pesada se encuentra en el cromosoma 14 y el gen que codifica la cadena ligera en el cromosoma 22 (cadena ligera lambda) y 2 (cadena ligera kappa), sería habitual que el biólogo experto en la técnica diseñara cebadores que hibriden con elementos reguladores del gen que codifica la cadena pesada y el gen que codifica la

cadena ligera. Dichos elementos reguladores incluyen, sin limitación, promotores, potenciadores e intrones.

Los cebadores específicos de la región variable de inmunoglobulina también se pueden determinar fácilmente para las inmunoglobulinas de ratones ya que se sabe que el gen de la cadena ligera kappa murina está localizado en el cromosoma 6 y se sabe que el gen de la cadena pesada murina está localizado en el cromosoma 12.

- 5 En otra realización no limitante, el material de partida a partir del cual se obtiene la información de la secuencia de ácido nucleico que codifica el gen de inmunoglobulina es ARNm o ADNc obtenido mediante traducción inversa a partir del ARNm. En este ejemplo, para obtener secuencias de ácido nucleico que codifican una región variable de inmunoglobulina, los cebadores pueden seleccionarse para que sean idénticos o hibridar con la cola de poliA de un ARNm o con la secuencia de ARNm rica en TTTT (SEC ID NO: 306) complementaria de los ADNc de los ARNm correspondientes. Alternativamente, o además, los cebadores también se pueden seleccionar para que sean idénticos o hibriden con las secuencias de ácido nucleico que codifican FR1. Alternativamente, o además, los cebadores también pueden seleccionarse para que sean idénticos o hibriden con una porción de (o todas) las regiones CH (es decir, CH1, CH2 o CHS) y/o las secuencias de ácido nucleico que codifican la región VH.
- 10

15 Pueden surgir errores de secuenciación a partir del uso de cebadores degenerados universales para secuenciar moléculas de ácido nucleico que codifican inmunoglobulinas a partir de hibridomas. Por ejemplo, Essono et al., Protein Engineering, Design and Selection, pág. 1-8, 2009 describen un método que combina la secuenciación con la toma de huellas dactilares por espectrometría de masas de péptidos de la cadena de Ig correspondiente para determinar la secuencia correcta de un anticuerpo monoclonal producido por un clon de hibridoma. Sin embargo, en los métodos no limitantes de diversas realizaciones de la invención, la presencia de errores de secuenciación simplemente aumentará el número de diferentes secuencias de ácidos nucleicos. A diferencia de Essono et al., más arriba, dado que los métodos de diversas realizaciones de la invención permiten la creación de un único anticuerpo (tanto cadenas pesadas como ligeras o regiones variables del mismo) a partir de una población policlonal inicial de anticuerpos (donde el anticuerpo creado puede no aparecer realmente dentro de la población policlonal inicial de anticuerpos), tener un gran número de secuencias en la base de datos de material genético con la que correlacionar los datos de espectros de masas observados de la base de datos de péptidos es un activo.

20

25

#### **Información de espectros de masas pronosticados de la base de datos de material genético**

De acuerdo con diversas realizaciones de la invención, una vez que se generan las secuencias de nucleótidos de las moléculas de ácido nucleico, puede generarse información adicional basándose únicamente en la información de la secuencia de nucleótidos. Por ejemplo, la información de la secuencia de nucleótidos puede traducirse a secuencias de aminoácidos pronosticadas utilizando el código genético. Aunque el experto en la técnica puede traducir fácilmente secuencias de nucleótidos a secuencias de aminoácidos utilizando el código genético, se pueden utilizar diversas herramientas de traducción automatizadas (que están disponibles públicamente), tal como la herramienta de traducción ExPASy del Instituto Suizo de Bioinformática o la herramienta de traducción EMBOSS Transeq de EMBL-EBI.

30

35 De forma similar, la información del espectro de masas pronosticada de las secuencias de aminoácidos pronosticadas codificadas por las secuencias de ácido nucleico puede ser determinada fácilmente por el experto en la técnica. Por ejemplo, después de la digestión virtual (es decir, *in silico*) de los polipéptidos pronosticados codificados por las secuencias de ácido nucleico, los espectros de masas pronosticados de los fragmentos peptídicos pueden generarse utilizando herramientas de algoritmo de soporte lógico convencionales disponibles públicamente que incluyen, sin limitación, el soporte lógico Sequest (de Thermo Fisher Scientific, Inc., West Palm Beach, FL), el soporte lógico Sequest 3G (de Sage-N Research, Inc., Milpitas, CA), el soporte lógico Mascot (de Matrix Science, Inc., Boston, MA; véanse también Electrophoresis, 15 20(18) 3551-67 (1999)), y el soporte lógico XITandem (código abierto de The Global Proteome Machine Organization, cuyo uso es descrito por Baerenfaller K. et al., Science 320: 938-41, 2008).

40

45 Según se utiliza en la presente memoria, las palabras "pronosticado", "teórico" y "virtual" se utilizan indistintamente para referirse a secuencias de nucleótidos, secuencias de aminoácidos o espectros de masas que se obtienen a partir de la transcripción *in silico* (es decir, en una computadora) y/o traducción (para las secuencias pronosticadas de nucleótidos y aminoácidos) o digestión *in silico* y/o análisis de espectrometría de masas (para los espectros de masas pronosticados) de información de las secuencias de ácidos nucleicos. Por ejemplo, las secuencias de ácido nucleico se obtienen de moléculas de ácido nucleico genómicas obtenidas de linfocitos B como se describe en la presente memoria. La secuencia de nucleótidos de, por ejemplo, ARNm derivado de ADN genómico se pronostica después de la traducción *in silico* del ADN genómico. Este ARNm (o ADNc) pronosticado puede traducirse *in silico* para producir secuencias de aminoácidos pronosticadas. Las secuencias de aminoácidos pronosticadas pueden a continuación digerirse *in silico* con proteasas (p.ej., tripsina) y/o reactivos de escisión química de proteínas (p.ej., bromuro de cianógeno) para producir fragmentos peptídicos pronosticados (o teóricos o virtuales). Los fragmentos de péptidos virtuales se pueden analizar *in silico* para producir información de espectros de masas pronosticados. Por lo tanto, la información del espectro de masas pronosticado, los fragmentos peptídicos pronosticados, las secuencias de aminoácidos pronosticadas y las secuencias de ARNm o ADNc pronosticadas se pueden obtener todas a partir de las secuencias de ácido nucleico recogidas de los linfocitos B (p.ej., de un animal).

50

55

En ciertas realizaciones, la proteasa o proteasas y/o los reactivos químicos utilizados para digerir los polipéptidos pronosticados para generar fragmentos de péptidos pronosticados y por último los espectros de masas pronosticados son las mismas proteasas y/o reactivos utilizados para digerir la población inicial de anticuerpos policlonales, como se describió anteriormente.

5    ***Correlación de los espectros de masas observados con los espectros de masas pronosticados***

Como se describió anteriormente, el paso de los fragmentos derivados de la población policlonal inicial de anticuerpos a través de un espectrómetro de masas genera numerosos espectros de masas observados. Dado el número potencialmente grande de diferentes inmunoglobulinas dentro de una población policlonal, cada una con una secuencia de aminoácidos diferente, que se analizan con el espectrómetro de masas, los espectros de masas observados resultantes serán difíciles de volver a ensamblar en una región variable de la cadena de inmunoglobulina funcional. En los métodos de diversas realizaciones de la invención, debido a que están disponibles las secuencias de ácidos nucleicos codificantes, no hay necesidad de ensamblar los datos de los espectros de masas observados. En cambio, los espectros de masas observados se correlacionan con los espectros de masas pronosticados derivados de las secuencias de ácidos nucleicos de la base de datos de materiales genéticos para obtener las secuencias de aminoácidos (y nucleótidos subyacentes) de cadenas de inmunoglobulina completas (o regiones variables de las mismas) de una inmunoglobulina que se une específicamente a un antígeno de una población de inmunoglobulina policlonal de partida.

También como se describió anteriormente, la base de datos de material genético puede obtenerse a partir de moléculas de ácido nucleico aisladas del repertorio de células B de un animal inmunizado, incluyendo moléculas de ácido nucleico que codifican cadenas pesadas y ligeras de inmunoglobulina completas y regiones variables de las mismas. Los intentos de identificar ácidos nucleicos que codifican inmunoglobulinas específicas de antígeno basándose únicamente en la información de la base de datos de material genético (p.ej., clasificaciones de frecuencia de secuencias de la región variable) pueden pasar por alto aquellas inmunoglobulinas que se producen con bajas frecuencias pero manifiestan actividades superiores específicas de antígeno. Sin embargo, de acuerdo con las diversas realizaciones de esta invención, correlacionando la información de los espectros de masas pronosticados de la base de datos de material genético con la información de espectro de masas observada de los anticuerpos policlónicos circulantes reales como se describe en la presente memoria, pueden seleccionarse las cadenas de inmunoglobulina (o regiones variables de las mismas) en la base de datos de material genético que corresponden a las inmunoglobulinas dentro de los anticuerpos policlónicos circulantes.

30    Por "correlación" se entiende que la información de los espectros de masas observados derivada de los anticuerpos policlónicos de partida y la información de los espectros de masas pronosticados derivada de la base de datos de material genético se cotejan y comparan entre sí, de manera que las cadenas pesadas y/o ligeras de inmunoglobulina (o regiones variables de la misma) pueden identificarse o seleccionarse a partir de la base de datos de material genético que corresponde a cadenas pesadas y/o ligeras de inmunoglobulina (o regiones variables de las mismas) de inmunoglobulinas específicas de antígeno en la población policlonal de partida.

En realizaciones específicas, el proceso de correlación implica comparar la información del espectro de masas observado con la información del espectro pronosticado para identificar emparejamientos. Por ejemplo, cada uno de los espectros observados puede buscarse contra la colección de espectros de masas pronosticados derivados de la base de datos de material genético, identificándose cada espectro pronosticado con una secuencia peptídica de la base de datos de material genético. Una vez que se encuentra un emparejamiento, es decir, un espectro de masas observado se corresponde con un espectro de masas pronosticado, puesto que cada espectro de masas pronosticado se asocia de forma identificable con una secuencia peptídica en la base de datos de material genético, se dice que el espectro de masas observado ha encontrado su secuencia peptídica de emparejamiento - dicho emparejamiento también se denomina en la presente memoria "par espectro-peptido" o "PSM". Debido a la gran cantidad de espectros que se deben buscar y emparejar, este procedimiento de búsqueda y emparejamiento puede realizarse mediante funciones y programas ejecutados por computadora, tal como el algoritmo SEQUEST (Sage-N Research, Inc., Milpitas, CA).

50    En algunas realizaciones, la búsqueda y el emparejamiento se dirigen a dominios funcionales o fragmentos de inmunoglobulinas, tales como secuencias de región variable, secuencias de región constante, y/o una o más secuencias de CDR. Por ejemplo, los espectros observados solo se buscan contra espectros de masas pronosticados derivados de regiones V (y/o secuencias CDR3) de inmunoglobulinas para identificar PSM de región V (y/o CDR3). En otras realizaciones, la búsqueda y el emparejamiento se dirigen a secuencias de cadena pesada o ligera de inmunoglobulina completa.

55    Una vez completados la búsqueda y el emparejamiento, las cadenas pesadas o ligeras de inmunoglobulina en la base de datos de material genético se analizan y seleccionan basándose en uno o más de los siguientes parámetros: número de péptidos únicos, compartición de espectro, cobertura de la secuencia de aminoácidos, recuento de péptidos (recuento de péptidos totales o recuento de péptidos únicos), frecuencia de las secuencias de ácidos nucleicos codificantes y relación clonal.

El término "cobertura" al referirse a una secuencia o región (p.ej., una secuencia de cadena pesada o ligera, una

secuencia de región V o una secuencia de CDR) se define como el número total de aminoácidos dentro de la secuencia que se han identificado en péptidos que se mapean para la secuencia o región y que tienen un espectro observado coincidente, dividido por el número de aminoácidos en la secuencia o región. Cuanto mayor sea la cobertura, más probable es que aparezca la secuencia o región en la población policlonal real.

- 5 Por "número de péptidos únicos" se entiende el número de péptidos distintos que se observan mapeando una única secuencia de proteína (p.ej., una sola cadena pesada o ligera de inmunoglobulina o una región variable de la misma). Cuanto mayor sea el número, más probable es que la cadena de inmunoglobulina esté presente en la población policlonal. En realizaciones específicas, la selección de una cadena de inmunoglobulina se realiza en base a un número de péptidos únicos de al menos 5, 6, 7, 8, 9, 10, 11, 12 o más en la cadena de inmunoglobulina o su
- 10 región variable.

La "compartición de espectro" se determina dividiendo el número total de péptidos asignados a la secuencia por el número total de PSM seguros asignados a la base de datos genéticos completa. La compartición de espectro proporciona un recuento de péptidos legible por seres humanos expresado como el porcentaje de PSM que se correlacionan con una secuencia de región V específica.

- 15 El término "recuento de péptidos" al referirse a una secuencia de proteína (p.ej., una región CDR3 o una región variable) significa el número de veces que se identifica un péptido a partir de los espectros de masas observados que coinciden con la secuencia de proteína. Por ejemplo, el recuento de una región CDR3 significa el número de veces que se identifica un péptido a partir de los espectros de masas observados que coinciden con la región CDR3. El recuento de una región variable significa el número de veces que se identifica un péptido a partir de los espectros de masas observados que coinciden con la región variable. El "recuento total de péptidos" al referirse a una
- 20 secuencia proteica significa el número de veces que se identifica cualquier péptido (único o no único) a partir de los espectros de masas observados que coinciden con la secuencia de la proteína. "Recuento de péptidos únicos" significa la cantidad de veces que se identifica un péptido único a partir de los espectros de masas observados que coinciden con la secuencia de la proteína. Si el mismo péptido se ha identificado múltiples veces a partir de los
- 25 espectros de masas observados, se considerará el número total de veces que se observa este péptido para determinar el recuento de péptidos total, aunque este péptido se contará solo una vez para determinar el recuento de péptidos único.

En realizaciones específicas, se selecciona una cadena ligera o pesada de inmunoglobulina basándose en la cobertura de la secuencia. En otras realizaciones, la selección se basa en una combinación de cobertura de secuencia con uno o más parámetros diferentes, que incluyen el número de péptidos únicos, compartición de espectro, recuento de péptidos total, recuento de péptidos únicos, frecuencia de la secuencia de ácidos nucleicos codificante o relación clonal.

30 Los parámetros anteriores se pueden determinar independientemente con respecto a una cadena pesada o ligera completa, o con respecto a una o más porciones de una cadena pesada o ligera de inmunoglobulina, por ejemplo, la región variable y una CDR (p.ej., CDR1, CDR2, o CDR3, especialmente CDR3). En ciertas realizaciones, la selección de cadenas de inmunoglobulina (o regiones variables de las mismas) se realiza en base a la cobertura de la región V y/o la cobertura de CDR (p.ej., cobertura de CDR3).

35 La selección de cadenas ligeras o pesadas de inmunoglobulina (o regiones variables de las mismas) se puede realizar basándose en el valor absoluto de uno o más parámetros, o en base a la clasificación de valores absolutos para un parámetro relevante. Cuando se considera la clasificación para un parámetro concreto, las secuencias 10, 20, 30, 40, 50, 60, 70, 80, 90, 100 o más mejor clasificadas se pueden seleccionar independientemente de los valores absolutos de ese parámetro. Cuando se considera el valor de un parámetro, por ejemplo, el porcentaje de cobertura de secuencia, en algunas realizaciones, la selección de cadenas de inmunoglobulina se realiza en base a una cobertura de CDR (tal como la cobertura de CDR3) de al menos 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90%, 95%, 98% o más; adicional o alternativamente, en base a una cobertura de la región V de al menos 5%, 10%, 15%, 20%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 80%, 85%, 90% o más.

40 En algunas realizaciones, se realiza un análisis filogenético para determinar la relación clonal de la región variable de la cadena pesada, la región variable de la cadena ligera o una o más CDR (p.ej., CDRH3 o CDRL3). Los cambios o mutaciones de la secuencia nucleica de cadenas pesadas y ligeras en comparación con la secuencia de la línea germinal pueden proporcionar evidencia de la maduración por afinidad de los anticuerpos después de la exposición al antígeno. La relación clonal puede utilizarse como un factor en la selección de secuencias de anticuerpos. Se puede realizar un análisis filogenético por medio de métodos conocidos en la técnica, por ejemplo, los descritos por Dereeper et al., 2008, Nucl. Acids Res., 36(Web Server Issue): W456-459; Dereeper et al., 2010, BMC Evol. Biol., 10:8, y disponible en línea en [www.phylogeny.fr/version2.cgi/index.cgi](http://www.phylogeny.fr/version2.cgi/index.cgi). En algunas realizaciones, las regiones variables de la cadena pesada o ligera se agrupan por homología, a continuación se agrupan por homología con CDR (p.ej., CDR3).

45

50

55

Las secuencias de cadena pesada y ligera seleccionadas se pueden expresar a continuación por pares para ensamblar anticuerpos monoclonales que se analizan para confirmar la funcionalidad específica del antígeno. El

emparejamiento de secuencias de cadena pesada y ligera seleccionadas puede ser completamente aleatorio, o puede tomar en consideración uno o más parámetros descritos anteriormente, incluyendo cobertura de secuencia, número único de péptidos, compartición de espectro, recuento de péptidos total y recuento de péptidos únicos.

- 5 En algunas realizaciones, la abundancia de una población de anticuerpos que tienen una secuencia peptídica concreta se puede determinar utilizando un péptido marcado con isótopo pesado (p.ej., AQUA). Véase, p.ej., el documento WO 03/016861 y Gerber et al., 2003, 100:6940-45. Estos métodos emplean la introducción de una cantidad conocida de al menos un patrón peptídico marcado con isótopos pesados (que tiene una firma única detectable mediante cromatografía LC-SRM) en una muestra biológica digerida con el fin de determinar, en comparación con el patrón peptídico, la cantidad absoluta de un péptido con la misma secuencia y modificación de proteína en la muestra biológica. El péptido puede ser único para una especie de anticuerpo o encontrarse en múltiples (p.ej., relacionados con la clonalidad) anticuerpos. En algunas realizaciones, el péptido puede incluir al menos una porción de una CDR (p.ej., CDR3). La cuantificación de la abundancia de poblaciones de anticuerpos puede ser útil en métodos de verificación de la composición de anticuerpos en suero, por ejemplo, después de la vacunación de un sujeto.
- 10 15 Debe observarse que no se necesita que la inmunoglobulina que se une específicamente al antígeno cuya secuencia de aminoácidos (o secuencia de ácido nucleico) creada utilizando los métodos no limitantes de diversas realizaciones de la invención esté realmente presente dentro de la población policlonal inicial de inmunoglobulinas. Por el contrario, los métodos no limitantes de diversas realizaciones de la invención simplemente permiten la creación rápida de una inmunoglobulina que se une específicamente al antígeno independientemente de si la inmunoglobulina realmente existía en la población policlonal de partida. Por ejemplo, la inmunoglobulina creada que tiene las cualidades deseadas más altas (p.ej., la mayor afinidad de unión (o KD más baja) para el antígeno o un isótipo deseado (p.ej., IgG2a) puede ser el resultado de una cadena ligera de un primer anticuerpo en la población policlonal ensamblada con una cadena pesada de un segundo anticuerpo (es decir, diferente del primer anticuerpo) en la población policlonal. La inmunoglobulina creada resultante se puede caracterizar adicionalmente (p.ej., afinidad de unión por el antígeno o isótipo) de acuerdo con métodos convencionales.
- 20 25

#### **Método de preparación de anticuerpos recombinantes**

Una vez que se dilucida la secuencia de nucleótidos de una cadena de inmunoglobulina (o región variable de la misma) de un anticuerpo que se une específicamente al antígeno, se puede generar una molécula de ácido nucleico que comprende esa secuencia.

- 30 Por ejemplo, si la población de partida a partir de la cual se obtienen las moléculas de ácido nucleico que codifican la cadena de inmunoglobulina (o región variable de la misma) es una biblioteca de ADNc, la molécula de ácido nucleico que comprende la secuencia elucidada puede obtenerse fácilmente a partir de la biblioteca (p.ej., escrutando la biblioteca con un cebador idéntico o capaz de hibridar con una porción de la secuencia elucidada) o mediante PGR que amplifica la molécula de ácido nucleico de la biblioteca utilizando cebadores diseñados para amplificar la secuencia de ácido nucleico elucidada.
- 35

Alternativamente (o además), las moléculas de ácido nucleico que comprenden la secuencia de nucleótidos elucidada se pueden generar simplemente generando artificialmente la molécula de ácido nucleico utilizando una máquina de síntesis de ADN convencional. Numerosas máquinas de síntesis de ADN están disponibles comercialmente, incluyendo, sin limitación, la serie de sintetizadores MerMade (p.ej., MerMade 4, MerMade 6, MerMade 384, etc.) disponible de BioAutomation, Piano, TX; los diversos sintetizadores de ADN/ARN comercializados por Applied Biosystems (ahora parte de Life Technologies, Corp., Carlsbad, CA). Varias compañías también ofrecen servicios de síntesis de ADN (p.ej., BioPioneer, Bio S & R, Biomatik, Epoch BioLabs, etc.)

- 40 Se conocen métodos para expresar ácido nucleico que codifica cadenas pesadas y ligeras de una inmunoglobulina para producir inmunoglobulinas recombinantes (véase, p.ej. Publicación de Patente de Estados Unidos Núm. 6.331.415; 5.969.108; 7.485.291; US 2011-0045534; y PCT Publ. Núm. WO 2011/022077). Las inmunoglobulinas recombinantes pueden prepararse en una variedad de células que incluyen, sin limitación, células de insecto (p.ej., células SF9), células de hámster (p.ej., células CHO), células murinas (p.ej., células NIH-3T3), células de primate (p.ej., células COS), células humanas (p.ej., células Hela) y células procariotas (p.ej., células de *E. coli*). En algunas realizaciones, las células que expresan las inmunoglobulinas recombinantes de diversas realizaciones de la descripción son capaces de añadir modificaciones secundarias (p.ej., glicosilación) a la inmunoglobulina recombinante de una manera similar a la de la especie de la que se obtuvo originalmente la inmunoglobulina. Por ejemplo, cuando se recolecta la población de anticuerpos policlonales cuyos fragmentos se utilizaron para generar los datos del espectro de masas observados de un ser humano, se pueden utilizar células humanas (o células que glicosilan proteínas de manera similar o idéntica a las células humanas).
- 45 50

- 55 Para obtener la expresión de las secuencias de ácido nucleico de una inmunoglobulina recombinante (o fragmento de unión a antígeno de la misma) que se une específicamente al antígeno en una célula, las secuencias de ácido nucleico pueden ligarse en un vector (p.ej., un plásmido o un vector retroviral) que contiene secuencias reguladoras apropiadas de manera que las secuencias de ácido nucleico insertadas se expresan en la célula en la que se introducen las secuencias de ácido nucleico. Dichas secuencias reguladoras incluyen, por ejemplo, promotores,

potenciadores, elementos aceptores de intrón, sitios de poliadenilación, etc. Se puede emplear cualquier método para introducir las secuencias de ácido nucleico de una inmunoglobulina recombinante (o vector que las contiene) en una célula incluyendo, sin limitación, electroporación, transfección por medios químicos (p.ej., CaPO<sub>4</sub>, 20 DEAE-dextrano, polietilenimina), infección, transducción, fusión de liposomas, etc. (véanse los métodos, por ejemplo, en Ausubel et al., más arriba).

De acuerdo con algunas realizaciones de la invención, la cadena pesada de inmunoglobulina y la cadena ligera de inmunoglobulina se seleccionan aleatoriamente para ensamblarse en una inmunoglobulina (o región variable o dominio de unión a anticuerpo de la misma). Por ejemplo, la correlación de los espectros de masas reales de un fragmento peptídico de los anticuerpos policlonales con los espectros de masas pronosticados de un péptido pronosticado codificado por las secuencias de ácidos nucleicos se utilizará para obtener la secuencia de nucleótidos o la secuencia de aminoácidos pronosticada de una cadena de inmunoglobulina que comprende el fragmento peptídico. La secuencia de nucleótidos obtenida de la cadena de inmunoglobulina puede a continuación co-expresarse aleatoriamente y/o con una segunda secuencia de nucleótidos obtenida similarmente de una cadena de inmunoglobulina, donde la segunda secuencia de nucleótidos codifica la otra cadena de un anticuerpo intacto en condiciones en que las dos inmunoglobulinas codificadas las cadenas se ensamblarán en un anticuerpo intacto.

Se conocen las condiciones para coexpresar dos secuencias de nucleótidos (p.ej., en células) que codifican cada una una cadena de inmunoglobulina de manera que se ensamble una inmunoglobulina intacta (véase, p.ej., las Patentes de Estados Unidos Núm. 5.969.108; 6.331.415; 7.498.024; 7.485.291; y la Publicación de Patente de Estados Unidos Núm. 20110045534). Debido al número de secuencias de nucleótidos diferentes que se pueden obtener utilizando los métodos descritos en la presente memoria, la invención contempla el uso de métodos robóticos y de alto rendimiento para seleccionar las inmunoglobulinas codificadas para crear una inmunoglobulina que se une específicamente al antígeno.

Según se utiliza en la presente memoria, por "ensamblado" o "ensamblaje" se entiende que una cadena ligera de un anticuerpo (o un fragmento de la misma) y una cadena pesada de un anticuerpo (o un fragmento de la misma) se combinan entre sí de una manera en la que las dos cadenas se unen para crear un anticuerpo (o un fragmento del mismo). En algunas realizaciones, en el anticuerpo ensamblado (o fragmento del mismo), los residuos de aminoácidos tanto de la cadena pesada como de la cadena ligera contribuyen al dominio de unión al antígeno del anticuerpo ensamblado (o fragmento del mismo). En algunas realizaciones, el anticuerpo ensamblado (o fragmento del mismo) comprende una cadena ligera (o fragmento de la misma) unida covalentemente a una cadena pesada (o fragmento de la misma). En algunas realizaciones, el anticuerpo ensamblado (o fragmento del mismo) comprende una cadena ligera (o fragmento de la misma) unida no covalentemente a una cadena pesada (o fragmento de la misma).

En algunas realizaciones, las secuencias de nucleótidos o secuencias de aminoácidos de las cadenas de inmunoglobulina (o regiones variables de las mismas) identificadas en el análisis proteómico descrito anteriormente se sintetizan mediante técnicas recombinantes de biología molecular o técnicas de síntesis de genes antes del ensamblaje de anticuerpos recombinantes. Por ejemplo, las secuencias de nucleótidos o aminoácidos se pueden sintetizar en una máquina de síntesis de nucleótidos o péptidos antes del ensamblaje. O, las secuencias de nucleótidos o aminoácidos pueden expresarse de forma recombinante clonando las secuencias de nucleótidos en un vector de expresión (p.ej., PCDNA3.1 de Invitrogen, Carlsbad, CA) y expresando el polipéptido codificado en una célula (p.ej., Células HeLa, células CHO, células COS, etc.) transfectadas con el vector de expresión. En algunas realizaciones, la etapa de ensamblaje ocurre en la célula transfectada (p.ej., una sola célula se transfecta con uno o más vectores de expresión que comprenden secuencias de ácido nucleico que codifican una cadena pesada y una cadena ligera, donde la cadena pesada y ligera se expresará como polipéptidos en la célula transfectada).

En diversas realizaciones de la descripción, los anticuerpos recombinantes se aislan. Según se utiliza en la presente memoria, por "aislado" (o "purificado") se entiende que un anticuerpo está sustancialmente libre de otro material biológico con el que está asociado de forma natural, o libre de otros materiales biológicos derivados, por ejemplo, de una célula que ha sido genéticamente modificada para expresar el anticuerpo de la descripción. Por ejemplo, un anticuerpo recombinante aislado es uno que está físicamente separado de otros componentes de la célula anfitriona (p.ej., el retículo endoplásmico o las proteínas citoplásmicas y el ARN). Asimismo, un anticuerpo purificado a partir de suero y/o plasma sanguíneo es un anticuerpo que se aísla de otros componentes de suero o plasma (p.ej., albúmina o células) (utilizando, por ejemplo, la adherencia de los anticuerpos a la proteína A, donde los componentes del suero que no son el anticuerpo no se adherirán a la proteína A). Por lo tanto, un anticuerpo aislado (o inmunoglobulina aislada) de la presente descripción incluye un anticuerpo que es puro en al menos 70-100%, es decir, un anticuerpo que está presente en una composición en la que el anticuerpo constituye 70-100% en peso de la composición total. En algunas realizaciones, el anticuerpo aislado de la presente descripción es puro en 75%-99% en peso, puro en 80%-99% en peso, puro en 90-99% en peso, o puro en 95% a 99% en peso. El grado relativo de pureza de un anticuerpo en diversas realizaciones no limitantes de la descripción se determina fácilmente mediante métodos bien conocidos.

En algunas realizaciones, los anticuerpos recombinantes (o las regiones variables de los mismos) se escrutan o analizan adicionalmente en un inmunoensayo para confirmar que los anticuerpos se unen específicamente al antígeno. En algunas realizaciones, el inmunoensayo es un inmunoensayo convencional tal como un ensayo de

citometría de flujo (p.ej., un FACScan), un ensayo inmunoabsorbente ligado a enzima (ELISA), un ensayo de transferencia Western, un ensayo de inmunohistoquímica, un ensayo de inmunofluorescencia, un radioinmunoensayo, un ensayo de neutralización, un ensayo de unión, un ensayo de afinidad o un ensayo de inmunoprecipitación de proteínas o péptidos. Todos estos inmunoensayos son ensayos convencionales bien conocidos y se han descrito bien en los libros de métodos convencionales (véase, p.ej., Ausubel et al., más arriba, Coligan et al., más arriba, Harlow y Lane, más arriba).

### **Anticuerpos Terapéuticos**

Las diversas realizaciones y métodos no limitantes de la invención son útiles, por ejemplo, para aislar anticuerpos que tienen un valor terapéutico. Por ejemplo, en el transcurso de una respuesta inmunitaria normal en un animal a un patógeno, los anticuerpos con la mayor especificidad para un antígeno del patógeno pueden tardar semanas en aparecer. Esto se debe a que el linfocito B que produce el anticuerpo debe ser estimulado primero por el linfocito T apropiado que también reconoce el antígeno presentado en una célula presentadora de antígeno en el contexto del complejo mayor de histocompatibilidad expresado por cada célula nucleada de un animal. Los linfocitos B que inicialmente responden al antígeno producen anticuerpos que se unen específicamente al antígeno. Sin embargo, los anticuerpos de mayor afinidad son aquellos producidos por linfocitos B que se han unido a su antígeno (a través de la inmunoglobulina expresada en superficie celular complejada con otros antígenos de superficie celular para formar el receptor de célula B) y, tras la estimulación a través del receptor de célula B y otras células (incluyendo linfocitos T) experimentan una maduración de afinidad para producir anticuerpos con alta afinidad por su antígeno específico. Dicho linfocito B que ha experimentado una maduración de afinidad (o su progenie con la misma especificidad de anticuerpo) está disponible en el animal para producir rápidamente anticuerpos de alta afinidad si el animal vuelve a encontrar el mismo patógeno.

Esta estricta regulación de linfocitos T y linfocitos B que responden a un antígeno la primera vez que se observa el antígeno (p.ej., la primera vez que el animal es infectado con un patógeno concreto) es necesaria para prevenir la respuesta inmunitaria autoinmune o inapropiada. Sin embargo, un inconveniente es que en el momento en que un linfocito B específico de antígeno secreta un anticuerpo de la mayor afinidad y especificidad por el antígeno, un patógeno de crecimiento rápido puede haber crecido dentro del animal en la medida en que ya no puede eliminarse fácilmente. En algunas realizaciones de la invención, los métodos permiten el rápido desarrollo de un anticuerpo específico de antígeno que omite el procedimiento que lleva tiempo de aislar en primer lugar un linfocito B específico de antígeno que secreta el anticuerpo e inmortaliza ese linfocito.

Por lo tanto, en otro aspecto, la descripción proporciona una composición terapéutica que comprende un anticuerpo recombinante con un vehículo farmacéuticamente aceptable.

Según se utiliza en la presente memoria, "portador farmacéuticamente aceptable" incluye cualquier material que, cuando se combina con un ingrediente activo (p.ej., un anticuerpo recombinante preparado de acuerdo con diversas realizaciones de la invención), permite que el ingrediente conserve la actividad biológica y no sea reactivo con el sistema inmunitario del sujeto y no tóxico para el sujeto cuando se suministra. Los ejemplos incluyen, pero sin limitación, cualquiera de los vehículos farmacéuticos convencionales tales como una solución salina tamponada con fosfato, agua, emulsiones tales como emulsión de aceite/agua y diversos tipos de agentes humectantes. Los ejemplos no limitantes de diluyentes para administración en aerosol o parenteral son solución salina tamponada con fosfato, solución salina normal (0,9%), solución de Ringer y solución de dextrosa. El pH de la solución puede ser de aproximadamente 5 a aproximadamente 8, o de aproximadamente 7 a aproximadamente 7,5. Otros vehículos incluyen preparaciones de liberación sostenida tales como matrices semipermeables de polímeros hidrófobos sólidos que contienen el anticuerpo, cuyas matrices están en forma de artículos conformados, por ejemplo, películas, liposomas o micropartículas. Resultará evidente para los expertos en la técnica que ciertos vehículos pueden ser más preferibles dependiendo de, por ejemplo, la vía administración y la concentración del anticuerpo que se administra. Las composiciones que comprenden tales vehículos se formulan mediante métodos convencionales bien conocidos (véanse, por ejemplo, Remington's Pharmaceutical Sciences, 18<sup>a</sup> edición, A. Gennaro, ed., Mack Publishing Co., Easton, Pa., 1990; y Remington, The Science and Practice of Pharmacy, 20<sup>a</sup> ed. Mack Publishing, 2000).

Si bien se puede emplear cualquier vehículo adecuado conocido por los expertos en la técnica en las composiciones farmacéuticas de esta descripción, el tipo de vehículo variará dependiendo del modo de administración. En diversas realizaciones de la descripción, son bien conocidas en la técnica numerosos mecanismos de administración para las composiciones farmacéuticas no limitantes descritas en la presente memoria (p.ej., que contiene un agente de unión o un polinucleótido que codifica un agente de unión), tales como los descritos por Rolland, 1998, Crit. Rev. Therap. Drug Carrier Systems 15: 143-198, y referencias allí citadas.

### **Métodos de tratamiento**

En otro aspecto, la descripción proporciona un método para tratar un animal que tiene o se sospecha que tiene una enfermedad caracterizada por un antígeno de enfermedad, comprendiendo el método administrar una cantidad eficaz de una composición terapéutica que comprende una inmunoglobulina que se une específicamente a un antígeno fabricado de acuerdo con los métodos de diversas realizaciones de la invención, en donde el antígeno

específicamente unido a la inmunoglobulina de la composición terapéutica y el antígeno de la enfermedad son los mismos.

En algunas realizaciones, el animal es un ser humano o un animal domesticado (p.ej., un perro, gato, vaca, cabra, oveja, pollo, pavo, llama, emú, elefante o aveSTRUZ).

5 Según se utiliza en la presente memoria, la expresión "caracterizado por" con respecto a una enfermedad y antígeno de enfermedad indicado (p.ej., un antígeno HIVgp 120 de SIDA) significa una enfermedad en la que el antígeno de enfermedad indicado está presente en un animal con esa enfermedad. En algunas realizaciones, el antígeno de enfermedad está codificado por ácido nucleico del agente etiológico de la enfermedad (p.ej., un virus). En algunas realizaciones, el antígeno de la enfermedad está codificado por el genoma del animal (p.ej., el antígeno de la enfermedad de fusión BCR-ABL codificado por el cromosoma Filadelfia en pacientes con leucemia mielógena crónica (LMC)).

10 15 Por "tratar" se entiende detener, retardar o inhibir la progresión de una enfermedad o prevenir el desarrollo de una enfermedad en un animal. Se conocen métodos para detectar si el tratamiento es satisfactorio. Por ejemplo, cuando la enfermedad es un tumor sólido, la progresión de la enfermedad se inhibe, se detiene o se retrasa si hay una regresión del tumor, reducción de metástasis, reducción del tamaño del tumor y/o reducción en el recuento de células tumorales después de la administración de la cantidad eficaz de una composición terapéutica que comprende una inmunoglobulina recombinante producida utilizando los métodos de diversas realizaciones de la invención.

20 25 Según se utiliza en la presente memoria, una "cantidad eficaz" es una cantidad o dosificación suficiente para producir resultados beneficiosos o deseados que incluyen detener, ralentizar, detener, retardar o inhibir la progresión de una enfermedad en un animal o prevenir el desarrollo de una enfermedad en un animal. Una cantidad eficaz variará dependiendo de, por ejemplo, la edad y el peso corporal de un sujeto al que se va a administrar la composición terapéutica que comprende la inmunoglobulina recombinante, la gravedad de los síntomas y la vía de administración, y de este modo se determina la administración de manera individual. En general, la dosis diaria de adultos para la administración oral es de aproximadamente 0,1 a 1000 mg, administrada como una dosis única o en dosis divididas. Para la administración intravenosa continua, las composiciones pueden administrarse en el intervalo de 0,01 ug/kg/min a 1,0 ug/kg/min, deseablemente de 0,025 ug/kg/min a 0,1 ug/kg/min.

30 35 Se puede administrar una cantidad eficaz en una o más administraciones. A modo de ejemplo, una cantidad eficaz de una inmunoglobulina recombinante producida utilizando los métodos de diversas realizaciones de la invención, es una cantidad suficiente para mejorar, detener, estabilizar, revertir, ralentizar y/o retrasar la progresión de una enfermedad (p.ej., un cáncer ) en un animal o es una cantidad suficiente para mejorar, detener, estabilizar, revertir, ralentizar y/o retrasar el crecimiento de una célula enferma (p.ej., una célula cancerosa biodegradada) in vitro. Como se entiende en la técnica, una cantidad eficaz de un anticuerpo recombinante de diversas realizaciones de la descripción puede variar, dependiendo, entre otros, del historial médico del animal, así como de otros factores, tales como el isotipo (y/o la dosificación) del anticuerpo recombinante.

40 45 50 Las cantidades eficaces y las programaciones para administrar las composiciones que comprenden un anticuerpo recombinante no limitante de diversas realizaciones de la descripción se pueden determinar empíricamente, y hacer tales determinaciones está dentro del conocimiento práctico de la técnica. Los expertos en la técnica comprenderán que la dosificación que debe administrarse variará dependiendo de, por ejemplo, el animal que recibirá las composiciones de diversas realizaciones de la descripción, la vía de administración, el tipo concreto de composiciones utilizadas (p.ej., el isotipo del anticuerpo recombinante dentro de la composición) y otros fármacos que se administran al animal. Cuando al animal (p.ej., un paciente humano) se le administra una composición que comprende un anticuerpo, se encuentra orientación en la selección de dosis apropiadas para el anticuerpo en la bibliografía sobre usos terapéuticos de anticuerpos, p.ej., *Handbook of Monoclonal Antibodies*, Ferrone et al., Eds., Noges Publications, Park Ridge, N.J., 1985, cap. 22 y págs. 303-357; Smith et al., *Antibodies in Human Diagnosis and Therapy*, Haber et al., Eds., Raven Press, Nueva York, 1977, págs. 365-389..

55 60 Una dosificación diaria típica de una cantidad eficaz de un anticuerpo utilizado solo podría oscilar entre aproximadamente 1 ug/kg y hasta 100 mg/kg de peso corporal o más por día, dependiendo de los factores mencionados anteriormente. Generalmente, se puede utilizar cualquiera de las siguientes dosis: una dosis de al menos aproximadamente 50 mg/kg de peso corporal; al menos aproximadamente 10 mg/kg de peso corporal; al menos aproximadamente 3 mg/kg de peso corporal; al menos aproximadamente 1 mg/kg de peso corporal; al menos aproximadamente 750 ug/kg de peso corporal; al menos aproximadamente 500 ug/kg de peso corporal; al menos aproximadamente 250 ug/kg de peso corporal; al menos aproximadamente 100 ug/kg de peso corporal; al menos aproximadamente 50 ug/kg de peso corporal; al menos aproximadamente 10 ug/kg de peso corporal; se administra al menos aproximadamente 1 ug/kg de peso corporal, o más. En algunas realizaciones, una dosis de un agente de unión (p.ej., anticuerpo) proporcionado en la presente memoria está entre aproximadamente 0,01 mg/kg y aproximadamente 50 mg/kg, entre aproximadamente 0,05 mg/kg y aproximadamente 40 mg/kg, entre aproximadamente 0,1 mg y aproximadamente 30 mg/kg, entre aproximadamente 0,1 mg y aproximadamente 20 mg/kg, entre aproximadamente 0,5 mg y aproximadamente 15 mg, o entre aproximadamente 1 mg y 10 mg. En algunas realizaciones, la dosis está entre aproximadamente 1 mg y 5 mg. En algunas realizaciones alternativas, la

dosis está entre aproximadamente 5 mg y 10 mg.

Los métodos descritos en la presente memoria (incluyendo los métodos terapéuticos) se pueden realizar mediante una sola inyección directa en un punto temporal único o en múltiples puntos temporales en un sitio único o múltiple.

5 La administración también puede ser casi simultánea a múltiples sitios. La frecuencia de administración puede determinarse y ajustarse a lo largo de la terapia y se basa en lograr los resultados deseados. En algunos casos, pueden ser apropiadas las formulaciones de liberación continua sostenida de las inmunoglobulinas recombinantes de diversas realizaciones de la descripción. Diversas formulaciones y dispositivos para lograr la liberación sostenida son conocidos en la técnica.

10 Las composiciones que comprenden los anticuerpos recombinantes de la presente descripción pueden formularse para cualquier forma de administración apropiada, incluyendo, por ejemplo, administración sistémica, tópica, oral, nasal, intravenosa, intracraneal, intraperitoneal, subcutánea o intramuscular, o por otros métodos, tales como infusión, que aseguran su entrega al torrente sanguíneo en una forma efectiva. La composición también se puede administrar mediante técnicas de perfusión aisladas, tales como perfusión tisular aislada, para ejercer efectos terapéuticos locales. Para administración parenteral, tal como inyección subcutánea, el vehículo preferiblemente

15 comprende agua, solución salina, alcohol, una grasa, una cera o un tampón. Para la administración oral, puede emplearse cualquiera de los vehículos anteriores o un vehículo sólido, tal como manitol, lactosa, almidón, estearato de magnesio, sacarina de sodio, talco, celulosa, glucosa, sacarosa y carbonato de magnesio. En algunas realizaciones, para la administración oral, la formulación de las composiciones es resistente a la descomposición en el tracto digestivo, por ejemplo, en forma de microcápsulas que encapsulan la inmunoglobulina recombinante de

20 diversas realizaciones de la descripción dentro de los liposomas. También se pueden emplear microesferas biodegradables (p.ej., polilactato poliglicolato) como vehículos para las composiciones terapéuticas de esta descripción. Se describen microesferas biodegradables adecuadas, por ejemplo, en las Patentes de Estados Unidos Núm. 4.897.268 y 5.075.109.

25 En algunas realizaciones de la descripción, las composiciones también pueden comprender tampones (p.ej., solución salina tamponada neutra o solución salina tamponada con fosfato), carbohidratos (p.ej., glucosa, manosa, sacarosa o dextrano), manitol, proteínas, polipéptidos o aminoácidos tales como glicina, antioxidantes, agentes quelantes tales como EDTA o glutatión, coadyuvantes (p.ej., hidróxido de aluminio) y/o conservantes. Alternativamente, las composiciones no limitantes de diversas realizaciones de la presente descripción se pueden formular como un producto liofilizado.

30 En algunas realizaciones de la descripción, las inmunoglobulinas recombinantes también pueden atraparse en microcápsulas preparadas, por ejemplo, mediante técnicas de coacervación o mediante polimerización interfacial (p.ej., hidroximetilcelulosa o microcápsulas de gelatina y microcápsulas de poli(metacrilato de metilo), respectivamente), en sistemas coloidales de suministro de fármaco (p.ej., liposomas, microesferas de albúmina, microemulsiones, nanopartículas y nanocápsulas), o en macroemulsiones. Tales técnicas se describen en Remington's Pharmaceutical Sciences, 18<sup>a</sup> edición, A. Gennaro, ed., Mack Publishing Co., Easton, Pa., 1990; y Remington, The Science and Practice of Pharmacy 20<sup>a</sup> Ed. Mack Publishing, 2000. Para aumentar la semivida en suero de la inmunoglobulina recombinante de diversas realizaciones de la descripción, se puede incorporar un epítopo de unión al receptor de rescate al anticuerpo (especialmente un fragmento de anticuerpo) como se describe en Patente de Estados Unidos Núm. 5.739.277, por ejemplo. Según se utiliza en la presente memoria, el término "epítopo de unión al receptor de rescate" se refiere a un epítopo de la región Fc de una molécula de IgG (p.ej., IgG1, IgG2, IgG3 e IgG4) que es responsable del aumento de la vida media en suero in vivo de la molécula de IgG.

45 En algunas realizaciones de la descripción, las inmunoglobulinas recombinantes también se pueden formular en forma de liposomas. Los liposomas que contienen las inmunoglobulinas recombinantes se preparan por medio de métodos conocidos en la técnica, tal como describen Epstein et al., 1985, Proc. Natl. Acad. Sci. USA 82:3688; Hwang et al., 1980, Proc. Natl Acad. Sci. USA 77:4030; y Patentes de Estados Unidos Núm. 4.485.045 y 4.544.545. Los liposomas con mejora del tiempo de circulación se describen en la Patente de Estados Unidos Núm. 5.013.556. Se pueden generar liposomas particularmente útiles mediante el método de evaporación de fase inversa con una composición lípídica que comprende fosfatidilcolina, colesterol y fosfatidiletanolamina derivatizada con PEG (PEG-PE). Los liposomas se extruyen a través de filtros de tamaño de poro definido para producir liposomas con el diámetro deseado. Además, los anticuerpos de diversas realizaciones de la descripción (incluyendo fragmentos del dominio de unión a antígeno tales como fragmentos Fab') se pueden conjugar con los liposomas como describen Martin et al., 1982, J. Biol. Chem. 257:286-288, a través de una reacción de intercambio de disulfuro. La administración de los anticuerpos recombinantes de diversas realizaciones de la descripción incluye administración local o sistémica, incluyendo inyección, administración oral, pistola de partículas o administración mediante cateterización, y administración tópica. Un experto en la técnica está familiarizado con la administración de vectores de expresión para obtener la expresión de una proteína exógena in vivo. Véanse, p.ej., las Patentes de Estados Unidos Núm. 6.436.908, 6.413.942 y 6.376.471.

60 En otro aspecto, la descripción proporciona un método para reducir la probabilidad de aparición en un animal de una enfermedad caracterizada por la presencia en el animal de un antígeno de enfermedad, en el que el método comprende administrar una cantidad eficaz de una composición terapéutica que comprende una inmunoglobulina recombinante de varias realizaciones de la descripción, en donde el antígeno específicamente unido a la

inmunoglobulina de la composición terapéutica y el antígeno de la enfermedad son los mismos.

La preparación de la vacuna se describe generalmente en Vaccine Design ("The subunit and adjuvant approach" (eds Powell M. F. y Newman M. J., (1995) Plenum Press New York). En otro aspecto, la descripción proporciona un kit para determinar la secuencia de aminoácidos de un anticuerpo de un animal que comprende (a) un medio para

5 obtener secuencias de ácido nucleico que codifican regiones variables de cadenas de inmunoglobulina de inmunoglobulinas múltiples de un animal, y (b) instrucciones para correlacionar la información de los espectros de masas de un anticuerpo analizado por espectrometría de masas con la información del espectro de masas pronosticado derivado de las secuencias de ácidos nucleicos para determinar la secuencia de aminoácidos del anticuerpo.

10 Los métodos descritos en la presente memoria se pueden utilizar para controlar los anticuerpos circulantes a lo largo del tiempo, p.ej., en un sujeto inmunizado con un antígeno. En estas realizaciones, se pueden tomar muestras del sujeto en una pluralidad de puntos temporales (p.ej., antes y después de la inmunización) y los métodos descritos en la presente memoria se utilizan para identificar anticuerpos circulantes en cada punto temporal. La composición de los anticuerpos circulantes se puede comparar en la pluralidad de puntos temporales para determinar la eficacia y/o 15 el transcurso del tiempo de la vacunación. Esto puede ser útil para controlar respuestas inmunitarias en sujetos individuales y también en el desarrollo de vacunas.

Los siguientes ejemplos se proporcionan para ilustrar, pero no para limitar, los diversos aspectos y realizaciones de la invención.

#### Ejemplo 1

20 **Identificación de cadenas pesadas individuales de anticuerpos de una población policlonal de anticuerpos que se unen específicamente a un antígeno.**

En este ejemplo, se obtuvieron múltiples anticuerpos monoclonales a partir de una población policlonal de 25 anticuerpos que se unen específicamente a un antígeno. utilizando los métodos de diversas realizaciones de la invención, la información de la base de datos de material genético generada a partir de moléculas de ácido nucleico del animal cuyos sueros comprendía la población policlonal de partida se comparó con información de la base de datos de péptidos del análisis de los anticuerpos monoclonales.

30 Las secuencias de ácido nucleico se obtuvieron de esplenocitos de un animal inmunizado con el antígeno de acuerdo con los métodos descritos en la presente memoria utilizando cebadores específicos para secuencias que codifican cadenas de inmunoglobulina de conejo (véanse, por ejemplo, las secuencias de cebadores en el Ejemplo 6 a continuación). Las regiones CDR3 de las cadenas pesadas de los anticuerpos policloniales se clasificaron en función del número de veces que aparecieron en la base de datos y el porcentaje de veces que apareció cada CDR3 entre todas las regiones CDR3 en la base de datos. La Tabla 2 muestra las 25 regiones superiores de CDR3 y sus 35 frecuencias. Estos resultados muestran que las mismas secuencias de CDR3 se encontraron en muchos anticuerpos diferentes en la mezcla polyclonal. Esta información muestra que los anticuerpos que se unen específicamente al mismo antígeno comparten a menudo secuencias en sus regiones CDR3 (y presumiblemente en las otras regiones CDR). Esta información muestra que los métodos descritos en la presente memoria podrán identificar y aislar aquellas cadenas de inmunoglobulina (o fragmentos de las mismas) que se unirán específicamente al antígeno.

Tabla 2

SEQ ID NO:	CDR3	Recuento	Porcentaje
29	GVKF	582	7,90%
30	GVSTNV	530	7,20%
31	DPYDDPTYRGYGMPL	372	5,05%
32	NPAVNTYAS	345	4,69%
	GGL	198	2,69%
33	HLFLHF	196	2,66%
34	HLFLNL	172	2,34%

SEQ ID NO:	CDR3	Recuento	Porcentaje
	GNV	169	2,30%
	GNI	143	1,94%
35	HLFLNF	129	1,75%
36	GLGYVGSSVYIVKYINL	126	1,71%
37	DLIRVAGDTFYDGFNL	113	1,53%
38	GRYNGWGYSNDL	113	1,53%
39	GGGTTLYTYFDL	111	1,51%
40	GLGYVGSDVYIVKYINL	105	1,43%
41	GGYGYGYGNTDFNL	93	1,26%
42	DDGGVRVDFDL	87	1,18%
43	VDDSGWMPFKL	85	1,15%
44	NVGSSHYNLN	76	1,03%
45	DGTDHGFNIDL	72	0,98%
46	STFRNSYARLAL	69	0,94%
47	IPYGWYSGGAAAPYFDL	65	0,88%
48	NAAIL	62	0,84%
49	AVSDNGYGMYWFL	61	0,83%
50	ELAGYDVGVEF	59	0,80%

Para la creación de la base de datos de péptidos, se utilizaron los siguientes métodos.

#### Digestión proteolítica de anticuerpos

- 5 Se concentraron aproximadamente 10 ug de la población policonal de anticuerpos y se cambió el tampón mediante ultrafiltración (0,5 ml 10K Amicon: Millipore). El volumen inicial se concentró primero, a continuación se cambió mediante la adición de 400 ul de Hepes 200 mM a pH 8. Las muestras se desnaturalizaron resuspendiéndolas en 80 ul de urea 8 M en Hepes pH 8 durante 15 minutos a temperatura ambiente. Los anticuerpos se redujeron en DTT 10 mM a temperatura ambiente durante 40 min.
- 10 La alquilación se realizó durante 1 hora con IAA 20 mM. La concentración de urea se redujo a una concentración final de 2M. Las muestras se dividieron igualmente por cinco y se digirieron por separado durante la noche a 37°C con tripsina, Lys-C, Glu-C, pepsina o quimotripsina, respectivamente. Para los productos digeridos con Pepsina, las muestras se concentraron y se cambiaron con ácido acético 3M y se digirieron a RT durante 1 hora. Las digestiones se inactivaron mediante la adición de TFA al 20% y se purificaron utilizando cartuchos Sep-Pack (Waters). Las muestras limpias se liofilizaron y resuspendieron para su análisis en un espectrómetro de masas LTQ Orbitrap Velos.
- 15 **Espectrometría de masas**

Las mezclas de péptidos producidas digiriendo la fracción de anticuerpo con las proteasas Lys-C, tripsina, quimotripsina, pepsina o Glu-C (es decir, los péptidos se produjeron digiriendo la fracción de anticuerpo con cada una de estas proteasas individualmente) se analizaron mediante LC-MS/MS utilizando individualmente el espectrómetro de masas híbrido LTQ Orbitrap Velos (Thermo-Fisher). Las muestras se cargaron durante 15 min

5 utilizando un automuestreador Famos (LC Packings) en una columna capilar de sílice fundida vertida a mano (125 um de diámetro interno, 18 cm) cargada con resina MagicCISaQ (5 m, 200 Å) utilizando una bomba binaria Agilent serie 1100 con un divisor de flujo en línea. La cromatografía se desarrolló utilizando un gradiente binario a 400 nl/min de 8-30% de disolvente B durante 35 minutos (Disolvente A, ácido fórmico (AF) al 0,25%, Disolvente B, AF al 0,1%, acetonaítrilo al 97%). A medida que los péptidos eluían de la columna de cromatografía líquida en el espectrómetro

10 de masas, se ionizaron y se midieron las razones masa-carga de ión peptídico para generar un espectro MS 1. El espectrómetro de masas a continuación seleccionó los 20 iones peptídicos más abundantes que eluyeron en ese momento y que no habían sido sometidos a la adquisición del espectro MS2 en los últimos 35 segundos, a continuación se aislaron y fragmentaron, a su vez, cada uno de esos 20 iones peptídicos para producir 20 espectros

15 de ion de producto MS2. Se realizó un ciclo completo de adquisición de un espectro MS 1 de iones precursores seguido de la adquisición de 20 espectros de iones producto MS2 de una manera dependiente de los datos en aproximadamente 1,6 segundos, y a continuación se repitió continuamente a medida que los péptidos se eluían de la columna de cromatografía líquida. El examen del estado de carga se usó para rechazar especies cargadas individualmente, y se requirió un umbral de 500 esferas para desencadenar un espectro MS/MS. Cuando fue posible, el LTQ y el Orbitrap se hicieron 30 funcionar en modo de procesamiento paralelo.

## 20 Búsqueda de base de datos y procesamiento de datos.

Los espectros de MS/MS se buscaron utilizando el algoritmo SEQUEST contra una base de datos genéticos. Los parámetros de búsqueda incluyeron especificidad enzimática completa para quimotripsina, Glu-C, Lys-C y tripsina, y ninguna especificidad enzimática para pepsina con una tolerancia de masa parental de 50 p.p.m., una modificación estática de 57,0214 sobre cisteína y modificaciones dinámicas de 15,9949 sobre metionina. Se realizaron búsquedas

25 en los espectros de HCD con una tolerancia de iones fragmento de  $\pm$  0,02 Da, mientras que los espectros de CID se buscaron con una tolerancia de iones fragmento de  $\pm$  1 Da. Los péptidos se filtraron a una FDR de péptido de 1% mediante el enfoque diana-señuelo, utilizando una función discriminante lineal para puntuar cada péptido basándose en parámetros tales como Xcorr, ACN y el error de masa del precursor.

## Resultados

30 La Figura 4 representa esquemáticamente el método seguido en este ejemplo. Las secuencias de ácido nucleico se analizaron utilizando las reglas de Kabat (véanse Kabat, E.A. et al., Sequences of Proteins of Immunological Interest, National Institutes of Health, Bethesda, Md., (1987) y Wu, T.T. y Kabat, E.A. J. Exp. Med. 132: 211-250 (1970)) para determinar dónde estaban ubicadas la región variable y CDR3 (y sus secuencias) dentro de las secuencias. A continuación, se elucidó el porcentaje de cobertura de las regiones CDR3 de la cadena pesada de múltiples

35 anticuerpos monoclonales identificados mediante espectrometría de masas. Como se muestra a continuación en la Tabla 3, se identificaron diecisésis secuencias de péptidos diferentes de la mezcla de anticuerpos policlonales analizada por MS, donde cada uno de los diecisésis péptidos comprendía la totalidad (es decir, 100%) de la Región CDR3 de la secuencia correspondiente a partir de las secuencias de ácido nucleico recogidas del animal.

Tabla 3

SEQ ID NO:	CDR3	% Cobertura CDR3
	GNL	100
	GNV	100
29	GVKF	100
30	GVSTNV	100
51	SRSTSYYINL	100
45	DGTDHGFNIDL	100

SEQ ID NO:	CDR3	% Cobertura CDR3
52	DGSDHGFNIDL	100
53	GADSIYRIYFDL	100
54	NVGSSYYNLNL	100
55	GGDAGYGYFDAFGP	100
56	GGDAGYGSFDAFGP	100
57	GLGYVGSSVYISKYINL	100
58	VPWTGGSGDARLTRLDL	100
36	GLGYVGSSVYIVKYINL	100
59	DLGYASYIGYGYPYYFKL	100
60	DLGYASYRGYGYPSYYFKL	100

- De los péptidos enumerados en la Tabla 3, cinco de los péptidos observados más frecuentes por espectrometría de masas también se observaron como espectros de masas teóricos derivados de la información de las secuencias de ácidos nucleicos. Por lo tanto, este experimento demostró que al comparar y correlacionar los espectros de masas pronosticados (y las secuencias subyacentes) derivados de las secuencias de ácido nucleico con los espectros de masas observados a partir de los fragmentos peptídicos reales de los anticuerpos policlonales, se obtuvieron fácilmente las secuencias de anticuerpos monoclonales múltiples (o al menos las cadenas pesadas de los mismos).

### Ejemplo 2

#### Desarrollo de un anticuerpo humano recombinante específico del antígeno de la gripe

- 10 Durante el invierno de 2009-2010, una cepa del virus de la gripe H1N1 infectó a un gran número de seres humanos, causando la muerte y lesiones permanentes. Utilizando los métodos no limitantes de diversas realizaciones de la invención, los anticuerpos neutralizantes pueden clonarse a partir de seres humanos expuestos previamente a una cepa de virus similar, y utilizarse como una composición para tratar pacientes humanos que actualmente padecen la enfermedad.
- 15 En consecuencia, las personas ancianas que se sabía que habían estado expuestas al virus de la influenza durante la epidemia de influenza de 1918 se someten a pruebas de detección de la presencia de anticuerpos séricos que pueden neutralizar el virus de 1918. Para hacer esto, se sigue el método descrito por Yu et al., Nature 455: 532-536, 2008.

20 Se identifican los pacientes cuyo suero y/o plasma sanguíneo contienen anticuerpos neutralizantes del virus, y se extrae sangre de estos pacientes y se separa en células y suero y/o plasma.

25 A partir de las células sanguíneas, se aislan los linfocitos B de acuerdo con métodos convencionales (véanse, por ejemplo, los métodos descritos en la presente memoria) y se obtienen moléculas de ácido nucleico de los linfocitos B. Las moléculas de ácido nucleico que codifican la cadena de inmunoglobulina se aislan de estas células mediante amplificación por PCR de ADN genómico utilizando cebadores que hibridan con regiones aguas arriba y aguas abajo de los genes de la región variable de cadena pesada (VH) y ligera (VL) de inmunoglobulina humana. Los métodos para preparar dichos cebadores son convencionales en el campo de la inmunología (véanse, por ejemplo, los métodos descritos en Marks y Bradbury, "PCR Cloning of Human Immunoglobulin Genes" en Antibody Engineering: Methods and Protocols, 248: 117-134, 2003).

Estas moléculas de ácido nucleico obtenidas mediante el uso de estos cebadores para la amplificación mediante PCR se utilizan para poblar la base de datos de material genético. Dentro de la base de datos genéticos, las secuencias de ácido nucleico se manipulan adicionalmente utilizando paquetes de soporte lógico convencionales para determinar la secuencia de aminoácidos del polipéptido codificado por cada secuencia de ácido nucleico, y los polipéptidos codificados se digieren virtualmente con tripsina, donde los péptidos resultantes generados a partir de tales productos digeridos se utilizan para generar los espectros de masas pronosticados.

A partir de la sangre de los pacientes, se recoge el suero y/o el plasma. Los anticuerpos presentes en el suero y/o plasma se aislan por medio de métodos convencionales. Por ejemplo, las proteínas séricas se pasan a través de una columna de proteína A-sefarosa, a la que se adhieren las inmunoglobulinas y las proteínas que no son inmunoglobulinas no. Debido a que las personas cuya sangre se recoge no están recién expuestas al virus de la influenza 1918, sus anticuerpos séricos se enriquecen adicionalmente en anticuerpos que se unen específicamente a un antígeno viral 1918 al pasar los anticuerpos séricos en una segunda columna recubierta con virus 1918 (p.ej., virus atenuado o fragmentos del mismo). A continuación, los anticuerpos unidos se tratan con una proteasa (p.ej., papaína) o un reactivo de escisión química de proteína que corta específicamente cerca de la región bisagra de la inmunoglobulina, y se retiran las porciones Fc no adherentes. Finalmente, los fragmentos Fab o Fab2 unidos se tratan con tripsina para generar fragmentos peptídicos, y todos los fragmentos se fraccionan a continuación utilizando cromatografía líquida, y a continuación los fragmentos se analizan mediante espectrometría de masas. utilizando un algoritmo tal como el programa Sequest, los espectros de masas en tandem observados de los péptidos se correlacionan con los espectros de masas pronosticados a partir de las secuencias de ácido nucleico extraídas de los linfocitos B de los pacientes. Utilizando este procedimiento, se puede identificar al menos un péptido encontrado dentro de la secuencia de aminoácidos pronosticada de una cadena de inmunoglobulina única de la base de datos de material genético. La secuencia de ácido nucleico que codifica esta cadena de inmunoglobulina (o región variable de la misma) se recupera después de la base de datos genética y se sintetiza utilizando métodos de síntesis de ADN convencionales. Las secuencias de ADN sintetizadas se subclonian a continuación en vectores de expresión que después se transfecan a células CHO. Los anticuerpos recombinantes producidos por las células se aíslan a continuación y se prueba su capacidad para unirse al virus 1918 (o fragmentos del mismo).

Los anticuerpos recombinantes producidos utilizando este método se combinan a continuación con un vehículo farmacéuticamente aceptable y se administran a pacientes que padecen la infección por el virus H1N1. Debido a que estos anticuerpos recombinantes son de origen completamente humano, no se espera que sean rechazados por el sistema inmunológico de los pacientes.

### Ejemplo 3

#### Obtención de secuencias de ácidos nucleicos

Este protocolo utiliza la secuenciación de nueva generación (NGS) y se basa en la plataforma 454 NGS (FLX +, FLX o junior; disponible comercialmente en 454 Life Sciences, una compañía de Roche, Branford, CT). Se necesitarán ligeras modificaciones para otras plataformas NGS de alto rendimiento y se basarán en las instrucciones de fabricación de NGS.

Los ratones se inmunizan con el antígeno de interés (péptido o péptidos, proteínas recombinantes, virus, toxina, etc.) con protocolos de inmunización convencionales (véase, por ejemplo, Coligan et al., más arriba). Las respuestas inmunitarias se controlan mediante el título de inmunoglobulinas en plasma frente al antígeno específico. La sangre, el bazo, la médula ósea, los ganglios linfáticos o cualquier órgano linfoide se pueden recolectar y procesar para aislar las células B de acuerdo con los métodos convencionales. Este procedimiento de aislamiento también se puede reducir si el material es limitado y se reemplaza con un procedimiento de RT-PCR directo utilizando cebadores de PCR específicos de dominio variable de inmunoglobulina contra poblaciones de cadenas pesadas y ligeras del animal.

Por supuesto, en algunas realizaciones, las secuencias de ácido nucleico pueden secuenciarse directamente a partir del material biológico (es decir, sin amplificarse antes de la secuenciación). Los servicios y reactivos para la secuenciación directa a partir de secuencias de ácido nucleico están disponibles comercialmente, por ejemplo, de Helicos BioSciences Corp. (Cambridge, MA). Por ejemplo, la secuenciación de molécula única verdadera de Helicos permite la secuenciación directa de ADN, ADNc y ARN. Véase también las Patentes de Estados Unidos Núm. 7.645.596; 7.037.687, 7.169.560; y publicaciones Harris et al., Science 320: 106-109, 2008; Bowers et al., Nat. Methods 6: 493-494, 2009; y Thompson y Milos, Genome Biology 12: 217, 2011.

En algunas realizaciones, las secuencias nucleicas se amplifican (p.ej., mediante reacción en cadena de la polimerasa) antes de obtener la información de la secuencia.

En un ejemplo no limitante, se utiliza un cebador de PCR oligo dT para RT-PCR. En otro ejemplo no limitante, se realiza RT-PCR específica de gen utilizando los cebadores de PCR que se describen a continuación. En otro ejemplo, los cebadores de PCR contra poblaciones de cadena pesada y cadena ligera en un ratón tienen secuencias establecidas en Publicación PCT Núm. WO2010/097435.

Con o sin enriquecimiento de células B, los materiales genéticos purificados (ADN o ARNm) se someterán a

continuación a RT-PCR siguiendo procedimientos convencionales (véase, por ejemplo, Ausubel et al., más arriba). Esta es la fase de preparación de la biblioteca de los materiales genéticos antes de ejecutar la secuenciación NGS. La reacción de transcripción inversa (RT) puede aplicar oligo dT o cebadores específicos de inmunoglobulina para generar ADNc. El procedimiento de reacción en cadena de la polimerasa aplicará cebadores específicos de inmunoglobulina para amplificar la región variable de cadenas ligeras y pesadas (reorganizadas o/y expresadas) de la muestra.

5

Estos métodos se describen con más detalles a continuación.

#### **Preparación de la biblioteca**

##### **Ejemplo de preparación de muestra:**

- 10 La sangre, el bazo, la médula ósea o los ganglios linfáticos se aíslan después de que los ratones reciban un refuerzo final con antígeno. Las células mononucleares se aíslan mediante separación con Ficoll como se describió anteriormente. A continuación, las células marcadas con Ficol se lavan con PBS, se cuentan y se congelan instantáneamente para la preparación de ARN total. El ARN total se aisló de las células utilizando el kit Qiagen RNeasy (disponible comercialmente de Qiagen Inc., Hilden, Alemania) de acuerdo con las instrucciones del fabricante, y el ARN total se almacena a -80°C.
- 15 Para la RT-PCR específica de gen o la RT-PCR convencional (utilizando oligo dT), se puede utilizar el siguiente protocolo.

cebador u Oligo dT RT-Ig de ratón CST 10 uM	1 ul
2,5 ug de ARN total (esplenocitos)	x ul
dNTP 10mM	2 ul
Agua esterilizada y destilada	hasta 14 ul

Incubar la mezcla a 65°C durante 5 minutos y a continuación colocarla en hielo.

5x tampón de síntesis de ADNc	4 ul
DTT 0,1 M	1 ul
Invitrogen Thermoscript RT (15 U/ul)	1 ul

20

Mezclar los contenidos suavemente e incubar a 60°C durante 60 minutos

Terminar la reacción calentando a 85°C durante 5 minutos

El ADNc está listo para su uso en la preparación de una biblioteca

- 25 El ADNc se someterá a continuación a PCR utilizando cebadores de ratón de fusión específicos CST 454 para cadenas pesadas y ligeras. Los cebadores tendrán las siguientes secuencias:

Cebadores de amplicón de ratón 454

Cadenas Pesadas (cebadores directo e inverso)

HV1 CCATCTCATCCCTGCGTGTCTCCGACTCAGACGAGTGCCTGATGTGAAGCTTCAGGAGTC  
 (SEQ ID NO: 1)  
 HV2 CCATCTCATCCCTGCGTGTCTCCGACTCAGACGCTCGACACAGGTGCAGCTGAAGGAGTC  
 (SEQ ID NO: 2)  
 HV3 CCATCTCATCCCTGCGTGTCTCCGACTCAGAGACGGACTCCAGGTGCAGCTGAAGCAGTC  
 (SEQ ID NO: 3)  
 HV4 CCATCTCATCCCTGCGTGTCTCCGACTCAGAGCACTGTAGGAGTTACTCTGAAAAGAGTC  
 (SEQ ID NO: 4)  
 HV5 CCATCTCATCCCTGCGTGTCTCCGACTCAGATCAGACACGGAGGTCCAGCTGCAACAATCT  
 (SEQ ID NO: 5)  
 HV6 CCATCTCATCCCTGCGTGTCTCCGACTCAGATCAGATATCGCGAGGAGGTCAGCTGCAGCAGTC  
 (SEQ ID NO: 6)  
 HV7 CCATCTCATCCCTGCGTGTCTCCGACTCAGCGTGTCTACAGGTCCAACCTGCAGCAGCCT  
 (SEQ ID NO: 7)  
 HV8 CCATCTCATCCCTGCGTGTCTCCGACTCAGCTCGCGTGTGAGGTGAAGCTGGTGGAGTC  
 (SEQ ID NO: 8)  
 HV9 CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTCTATGCCGAGGTGAAGCTGGTGAATC  
 (SEQ ID NO: 9)  
 HV10 CCATCTCATCCCTGCGTGTCTCCGACTCAGTGTACGCTGATGTGAACCTTGAAGTGTC  
  
 (SEQ ID NO: 10)  
  
 HVFOR1 CCTATCCCCTGTGTGCCTTGGCAGTCAGTCAGAGAGACAGTGACCAGAGT  
 (SEQ ID NO: 11)  
 HVFOR2 CCTATCCCCTGTGTGCCTTGGCAGTCAGTGAGGAGACTGTGAGAGTGTT  
 (SEQ ID NO: 12)  
 HVFOR3 CCTATCCCCTGTGTGCCTTGGCAGTCAGTGAGGAGACGGTGAAGGT  
 (SEQ ID NO: 13)  
 HVFOR4 CCTATCCCCTGTGTGCCTTGGCAGTCAGTGAGGAGACGGTGACTGAGGT  
 (SEQ ID NO: 14)

## Cadenas Kappa (Cebadores Directo e Inverso)

KV1 CCATCTCATCCCTGCGTGTCTCCGACTCAGCATAGTAGTGGATGTTTGATGACCAAAC  
 (SEQ ID NO: 15)  
 KV2 CCATCTCATCCCTGCGTGTCTCCGACTCAGCGAGAGATACTGATATTGTGATGACCGACGCT  
 (SEQ ID NO: 16)  
 KV3 CCATCTCATCCCTGCGTGTCTCCGACTCAGATACTGACGTAGATATTGTGATGATAACCCAG  
 (SEQ ID NO: 17)  
 KV4 CCATCTCATCCCTGCGTGTCTCCGACTCAGTCACTGACTAGACATTGTGCTGACCAAATCT  
 (SEQ ID NO: 18)  
 KV5 CCATCTCATCCCTGCGTGTCTCCGACTCAGCGTCTAGTACGACATTGTGATGACCCAGCT  
 (SEQ ID NO: 19)  
 KV6 CCATCTCATCCCTGCGTGTCTCCGACTCAGTCTACGTAGCGATATTGTGCTAACCTCAGTCT  
 (SEQ ID NO: 20)  
 KV7 CCATCTCATCCCTGCGTGTCTCCGACTCAGTGTACTACTCGATATCCAGATGACACAGACT  
 (SEQ ID NO: 21)  
 KV8 CCATCTCATCCCTGCGTGTCTCCGACTCAGACGACTACAGGACATCCAGCTGACTCAGTCT  
 (SEQ ID NO: 22)  
 KV9 CCATCTCATCCCTGCGTGTCTCCGACTCAGCGTAGACTAGCAAATTGTCTCACCCAGCT  
 (SEQ ID NO: 23)  
 KVFOR1 CCTATCCCCTGTGTGCCTTGGCAGTCAGCCGTTAGCCTCCAGCTTG (SEQ ID NO: 24)  
 KVFOR2 CCTATCCCCTGTGTGCCTTGGCAGTCAGCCGTTTATTCCAGCTTG (SEQ ID NO: 25)  
 KVFOR3 CCTATCCCCTGTGTGCCTTGGCAGTCAGCCGTTTATTCCAACTTG (SEQ ID NO: 26)

5

## Cadenas Lambda (cebadores directo e inverso)

LV CCATCTCATCCCTGCGTGTCTCCGACTCAGTACGAGTATGCAGGCTGTTGTGACTCAGGAA  
 (SEQ ID NO: 27)  
 LVFOR CCTATCCCCTGTGTGCCTTGGCAGTCAGCTGGCTGACCTAGGACAGT  
 (SEQ ID NO: 28)

En todas las secuencias anteriores, las secuencias subrayadas son para la secuenciación 454, las secuencias en negrita son códigos de barras para multiplexación, y las secuencias de fuente regular son secuencias específicas del ratón.

## ES 2 666 301 T3

Los cebadores se utilizan para amplificar las bibliotecas descritas anteriormente de la siguiente manera:

PCR de cadena pesada:

Mezcla de cebadores de cadena pesada de ratón CST454	1 ul
ADNc	1 ul
2x Phusion Master Mix	12,5 ul
H <sub>2</sub> O	10,5 ul

PCR de cadena ligera:

Mezcla de cebadores de cadena ligera de ratón CST454	1 ul
ADNc	1 ul
2x Phusion Master Mix	12,5 ul
H <sub>2</sub> O	10,5 ul

5

Las condiciones del ciclo de PCR pueden ser las de la Tabla 4 siguiente:

Tabla 4

Etapa	Temperatura	Tiempo (en minutos)
1: Etapa de desnaturación	98°C	01:30
2: Etapa de desnaturación	98°C	00:10
3: Etapa de recocido	60°C	00:30
4: Etapa de extensión	72°C	00:30

10 Se aplican 20 ciclos de las etapas 2 a 4. Los productos de PCR se someterán entonces a purificación de DNA Agencourt Ampure (disponible comercialmente de Beckman Coulter Genomics, Danvers, MA) 2 veces, siguiendo el protocolo del fabricante (véanse, por ejemplo, los protocolos del sistema Agencourt AMPure XP de Beckman Coulter Genomics).

15 Una vez que se prepara la biblioteca de PCR/genética, todas las etapas posteriores seguirán protocolos de fabricación 454 para emPCR y reacciones de secuenciación. Véanse las publicaciones de 454 Life Sciences Corp., A Roche Company, Branford, CT 06405 titulada, "Sequencing Method Manual, GS Junior Titanium Series" (May 2010 (rev. Junio de 2010)) y "emPCR Amplification Method Manual - Lib-L, GS junior Titanium Series (mayo de 2010 (rev. junio 2010)).

20 Se pueden combinar múltiples muestras en esta fase en una única ronda de secuenciación. Se distinguirán por un código de barras único (o MID de la plataforma 454). Por ejemplo, en el cebador de PCR se incorpora un código de barras.

En algunas realizaciones, se sigue el emPCR Amplification Method Manual - Lib-L, GS junior Titanium Series (mayo de 2010 (rev. junio de 2010); 454 Life Sciences Corp.). En algunas realizaciones, se sigue después el Sequencing Method Manual, GS Junior Titanium Series (mayo 2010 (rev. Junio 2010); 454 Life Sciences Corp.).

Los datos de secuenciación pueden producirse como archivos FASTA (o cualquier formato de archivo convencional)

y almacenarse en una base de datos de material genético. Estos datos de secuencia se utilizarán para generar la base de datos de espectros de masas pronosticados para analizar los espectros de masas de péptidos observados generados a partir de las inmunoglobulinas de suero y/o plasma del mismo animal. Para realizar esto se pueden utilizar los programas convencionales. En este ejemplo, los espectros de masas pronosticados fueron generados por el paquete de soporte lógico Sequest.

#### Ejemplo 4

##### Identificación de cadenas de anticuerpos individuales de una población policlonal

Los métodos descritos en la presente memoria se utilizaron a continuación para identificar la secuencia de anticuerpos individuales de varias poblaciones policlonales diferentes. Los métodos de este ejemplo se muestran esquemáticamente en las Figuras 2 y 4.

Utilizando los métodos descritos anteriormente en el Ejemplo 2, tres poblaciones policlonales diferentes de anticuerpos que se unen específicamente a tres antígenos diferentes se convirtieron en tres bibliotecas diferentes. La secuenciación profunda utilizando los métodos de secuenciación 454 descritos anteriormente se realizó utilizando cebadores específicos para secuencias codificantes de cadenas de inmunoglobulina de conejo para obtener tres bases de datos de material genético diferentes.

De forma correspondiente, las bases de datos de material genético se utilizaron para generar tres bases de datos de proteínas diferentes utilizando los métodos descritos en el Ejemplo 3 anterior.

Los resultados para el primer antígeno se muestran en las Tablas 5 (cadena ligera) y 6 (cadena pesada); el segundo antígeno se muestra en las Tablas 7 (cadena ligera) y 8 (cadena pesada) y el tercer antígeno se muestra en las Tablas 9 (cadena ligera) y 10 (cadena pesada).

Tabla 5

CDR3	Recuento CDR3	Cobertura CDR3	Péptidos Totales	Péptidos Únicos	Péptido CDR3
QGEFSCRDFCTV (SEQ ID NO: 61)	16	100	58	30	CQGEFSCRDFCTVF (SEQ ID NO: 62)
AGGYKSSGDTVS (SEQ ID NO: 63)	15	100	48	24	YCAGGYKSSGDTVSF (SEQ ID NO: 64)
AGGYKSTTDGSA (SEQ ID NO: 65)	9	100	29	17	CAGGYKSFTDGSAF (SEQ ID NO: 66)
QQGRRSVDVDNV (SEQ ID NO: 67)	8	100	25	12	CADAATYYCQQGRRSVDVDNVI'GGGTI (SEQ ID NO: 68)
QGEFNCDGVGCTT (SEQ ID NO: 69)	2	100	17	9	YCQGEFNCDGVGCTTF (SEQ ID NO: 70)

Tabla 6

CDR3	Recuento CDR3	Cobertura CDR3	Péptidos Totales	Péptidos Únicos	Péptido CDR3
GVRDWGDALDL (SEQ ID NO: 71)	5	100	42	22	GVRDWGDALDLWGQQTLVTVSSGQPK (SEQ ID NO: 72)
LYNSVVVGDDI (SEQ ID NO: 73)	10	100	38	20	LYNSVVVGDDIWGPGLTVIVSLGQPK (SEQ ID NO: 74)
LYNSVVVGDDM (SEQ ID NO: 75)	4	100	37	21	LYNSVVVGDDMWGPGLTVVSLGQPK (SEQ ID NO: 76)
GMPGSTSGNSNI (SEQ ID NO: 77)	2	100	34	20	GMPGSTSGNSNIWGPGLTVVSLGQPK (SEQ ID NO: 78)
LYNSLVVGDDI (SEQ ID NO: 79)	2	100	30	15	LYNSLVVGDDIWGPGLTVVSLGQPK (SEQ ID NO: 80)
KGDPGHGPNGLIJFTM (SEQ ID NO: 81)	3	100	22	19	KGDPGHGPNGLFFTMWGPGLTVVSLGQPK (SEQ ID NO: 82)
GGGSISGSATYDMIDP (SEQ ID NO: 83)	2	100	20	14	GGGSISGSATYDMIDPWGPGLTVVSSGQPK (SEQ ID NO: 84)
GTSRGSDYRLDDL (SEQ ID NO: 85)	2	100	15	11	GTSRGSDYRLDLWGQQTLTVSSGQPK (SEQ ID NO: 86)
GMPASTSGNSNI (SEQ ID NO: 87)	2	100	14	14	GMPASTSGNSNIWGPGLTVVSLGQPK (SEQ ID NO: 88)
DAIANI (SEQ ID NO: 89)	2	100	10	8	DAIANIWGPGLTVVSLGQPK (SEQ ID NO: 90)
DKWMVFGDLRL (SEQ ID NO: 91)	2	100	9	4	DKWMVFGDLRLWGPGLTVVSSGQPK (SEQ ID NO: 92)

Tabla 7

CDR3	Reuento CDR3	Cobertura CDR3	Péptidos Totales	Péptidos Únicos	Péptido CDR3
QQGRTYSVDVANV (SEQ ID NO: 93)	1	66.67	42	20	TYSDVANVPGGGTEVVVK (SEQ ID NO: 94)
QQGYSSYNVDNA (SEQ ID NO: 95)	2	41.67	75	20	NVDNAPGGGTEVVVK (SEQ ID NO: 96)
QQGYSSSNVDNA (SEQ ID NO: 97)	2	41.67	41	19	NVDNAFGGGTEVVVK (SEQ ID NO: 98)
LGTYDCRSADCNA (SEQ ID NO: 99)	2	46.15	33	18	SADCNAFGGGTEVVVK (SEQ ID NO: 100)
QHGYYSNVDNA (SEQ ID NO: 101)	2	45.45	46	18	NVDNAFGGGTEVVVK (SEQ ID NO: 102)
QQGIPSSRNVDNA (SEQ ID NO: 103)	2	41.67	24	18	NVDNAFGGGTEVVVK (SEQ ID NO: 104)
QQGYSSVNVDNA (SEQ ID NO: 105)	2	41.67	26	18	NVDNAFGGGTEVVVK (SEQ ID NO: 106)
QQGYTYNNVDNA (SEQ ID NO: 107)	2	41.67	27	16	NVDNAFGGGTEVVVK (SEQ ID NO: 108)
LGTYDCRSGLCNV (SEQ ID NO: 109)	1	46.15	25	15	SGDCNVIFGGGTEVVVK (SEQ ID NO: 110)
QQGYTSNVDNA (SEQ ID NO: 111)	2	45.45	26	15	NVDNAFGGGTEVVVK (SEQ ID NO: 112)
QQGQTPEVNDNA (SEQ ID NO: 113)	2	41.67	22	14	NVDNAFGGGTEVVVK (SEQ ID NO: 114)
QQGSTYSVDVANV (SEQ ID NO: 115)	1	66.67	29	14	TYSDVANVPGGGTEVVVK (SEQ ID NO: 116)
QQGATYSVDVANV (SEQ ID NO: 117)	1	66.67	63	13	TYSDVANVPGGGTEVVVK (SEQ ID NO: 118)
QQGITYSDVANV (SEQ ID NO: 119)	1	66.67	25	13	TYSDVANVPGGGTEVVVK (SEQ ID NO: 120)
QQGYTRSNVDNA (SEQ ID NO: 121)	2	41.67	21	11	NVDNAFGGGTEVVVK (SEQ ID NO: 122)
AGYKSYGNADID (SEQ ID NO: 123)	4	66.67	24	10	SYGNADIDPGGGTEVVVK (SEQ ID NO: 124)
QQGYTSSNVDNA (SEQ ID NO: 125)	2	41.67	17	9	NVDNAFGGGTEVVVK (SEQ ID NO: 126)
LVSYDCSSADCNA (SEQ ID NO: 127)	2	46.15	51	8	SADCNAFGGGTEVVVK (SEQ ID NO: 128)
QQAYTSSNVDNA (SEQ ID NO: 129)	2	41.67	4	3	NVDNAFGGGTEVVVK (SEQ ID NO: 130)

Tabla 8

CDR3	Recuento CDR3	Cobertura CDR3	Péptidos Totales	Péptidos Únicos	Péptido CDR3
DGGI. (SEQ ID NO: 131)	2	100	11	6	DGGLWGPGLVTVSSGQPK (SEQ ID NO: 132)
DPYDTNTSI.DAI. (SEQ ID NO: 133)	2	100	10	4	DPYDTNTSI.DAI.WGPGLVTVSSGQPK (SEQ ID NO: 134)
EGSDDDSFDL. (SEQ ID NO: 135)	4	100	10	5	EGSDDDSFDL.WGPGLVTVSSGQPK (SEQ ID NO: 136)
GGIDL. (SEQ ID NO: 137)	2	100	9	5	GGIDL.WGQGTLVTVSSGQPK (SEQ ID NO: 138)
GHWSAGATLYGYFSL. (SEQ ID NO: 139)	2	100	11	5	GHWSAGATLYGYFSL.WGPGLVTVSSGQPK (SEQ ID NO: 140)

Tabla 9

CDR3	Recuento CDR3	Cobertura CDR3	Péptidos Totales	Péptidos Únicos	Péptido CDR3
IANYDCSSGDCSV (SEQ ID NO: 141)	1	100	28	18	CLANYDCSSGDCSVF (SEQ ID NO: 142)
QGNIPDCSSADC'SA (SEQ ID NO: 143)	2	100	37	21	CQGNFDCTSADC'SAF (SEQ ID NO: 144)
QGNFDCTSADC'SA (SEQ ID NO: 145)	2	100	37	21	CQGNFDCTSADC'SAF (SEQ ID NO: 146)

Tabla 10

CDR3	Recuento CDR3	Cobertura CDR3	Péptidos Totales	Péptidos Únicos	Péptido CDR3
DGTDHGFNIDL. (SEQ ID NO: 45)	8	100	38	22	DGTDHGFNIDLWGPGLVTVSSGQPK (SEQ ID NO: 147)
GNV	2	100	36	21	TSTTVTLQMTSLTAADTATYFCASGNV WGPGLVTVSSGQPK (SEQ ID NO: 148)
GVSTINV (SEQ ID NO: 30)	6	100	29	19	GVSTINVWGPGLVTVSSGQPK (SEQ ID NO: 149)
GVKIF (SEQ ID NO: 29)	4	100	30	18	ICTRGVKIF (SEQ ID NO: 150)
DGSDFHGFNIDL (SEQ ID NO: 52)	6	100	29	16	DGSDFHGFNIDLWGPGLVTVSSGQPK (SEQ ID NO: 151)
NAAIL					NAAILWGPGLVTVSSGQPK (SEQ ID NO: 152)
(SEQ ID NO: 152)	10	100	34	16	NAAILWGPGLVTVSSGQPK (SEQ ID NO: 153)
SRSTSYYINL					SRSTSYYINLWGPGLVTVSSGQPK (SEQ ID NO: 154)
(SEQ ID NO: 154)	12	100	33	15	SRSTSYYINLWGPGLVTVSSGQPK (SEQ ID NO: 155)
GGDAGYGSFDAGP (SEQ ID NO: 56)	6	100	30	14	GGDAGYGSFDAGPFWGPGLVTVSSGQPK (SEQ ID NO: 156)
GVSTDV (SEQ ID NO: 157)	2	100	25	14	GVSTINVWGPGLVTVSSGQPK (SEQ ID NO: 158)

<b>NVGSSYYNLNL (SEQ ID NO: 54)</b>	<b>6</b>	<b>100</b>	<b>28</b>	<b>14</b>	<b>NVGSSYYNLNLWGPGLTVVSSGQPK (SEQ ID NO: 159)</b>
<b>GVSTSV (SEQ ID NO: 160)</b>	<b>2</b>	<b>100</b>	<b>24</b>	<b>13</b>	<b>GVSTNVWGPGLTVVSSGQPK (SEQ ID NO: 161)</b>
<b>GGYAGAGYFDALNP (SEQ ID NO: 162)</b>	<b>2</b>	<b>100</b>	<b>21</b>	<b>12</b>	<b>GGYAGAGYFDALNPWGPGLTVVSSGQPK (SEQ ID NO: 163)</b>
<b>NYNL (SEQ ID NO: 164)</b>	<b>6</b>	<b>100</b>	<b>26</b>	<b>12</b>	<b>NYNLWGPGLTVVSSGQPK (SEQ ID NO: 165)</b>
<b>RDGESTDRYFNL (SEQ ID NO: 166)</b>	<b>7</b>	<b>91.67</b>	<b>25</b>	<b>12</b>	<b>DGFSSTDYIPLWGPGLTVVSSGQPK (SEQ ID NO: 167)</b>
<b>DRGIGSGDYTPFNL (SEQ ID NO: 168)</b>	<b>5</b>	<b>71.43</b>	<b>26</b>	<b>12</b>	<b>GSGDYTPFNLWGPGLTVVSSGQPK (SEQ ID NO: 169)</b>
<b>DAAIL (SEQ ID NO: 170)</b>	<b>8</b>	<b>100</b>	<b>27</b>	<b>11</b>	<b>NAAILWGPGLTVVSSGQPK (SEQ ID NO: 171)</b>
<b>GPYVDSTYYNL (SEQ ID NO: 172)</b>	<b>6</b>	<b>100</b>	<b>23</b>	<b>11</b>	<b>GPYVDSTYYNLWGPGLTVVSSGQPK (SEQ ID NO: 173)</b>
<b>GSGDYTPFNL (SEQ ID NO: 174)</b>	<b>6</b>	<b>100</b>	<b>23</b>	<b>11</b>	<b>GSGDYTPFNLWGPGLTVVSSGQPK (SEQ ID NO: 175)</b>
<b>YYDGADYHTYNL (SEQ ID NO: 176)</b>	<b>6</b>	<b>100</b>	<b>21</b>	<b>11</b>	<b>YYDGADYHTYNLWGPGLTVVSSGQPK (SEQ ID NO: 177)</b>
<b>EFGNNGWNIDL (SEQ ID NO: 178)</b>	<b>6</b>	<b>100</b>	<b>21</b>	<b>10</b>	<b>EFGNNGWNIDLWGPGLTVVSSGQPK (SEQ ID NO: 179)</b>
<b>VEYGNDWGNL (SEQ ID NO: 180)</b>	<b>6</b>	<b>100</b>	<b>20</b>	<b>10</b>	<b>VEYGNDWGNLWGPGLTVVSSGQPK (SEQ ID NO: 181)</b>
<b>YFDGADYHTYNL (SEQ ID NO: 182)</b>	<b>6</b>	<b>100</b>	<b>20</b>	<b>10</b>	<b>YFDGADYHTYNLWGPGLTVVSSGQPK (SEQ ID NO: 183)</b>
<b>RFSGGGYGYDL (SEQ ID NO: 184)</b>	<b>5</b>	<b>90.91</b>	<b>25</b>	<b>10</b>	<b>FSGGGYGYDLWGPGLTVVSSGQPK (SEQ ID NO: 185)</b>
<b>IDRDL (SEQ ID NO: 186)</b>	<b>6</b>	<b>100</b>	<b>19</b>	<b>9</b>	<b>IDRDLWGPGLTVVSSGQPK (SEQ ID NO: 187)</b>
<b>GLDL (SEQ ID NO: 188)</b>	<b>5</b>	<b>100</b>	<b>19</b>	<b>9</b>	<b>YGLDLWGPGLTVVSSGQPK (SEQ ID NO: 189)</b>
<b>YDVDSVSAYDL (SEQ ID NO: 190)</b>	<b>6</b>	<b>100</b>	<b>24</b>	<b>9</b>	<b>YDVDSVSAYDLWGPGLTVVSSGQPK (SEQ ID NO: 191)</b>
<b>EVVGYDYSGLDL (SEQ ID NO: 192)</b>	<b>6</b>	<b>100</b>	<b>18</b>	<b>8</b>	<b>EVVGYDYSGLDLWGPGLTVVSSGQPK (SEQ ID NO: 193)</b>
<b>DPYDDPTY (SEQ ID NO: 194)</b>	<b>2</b>	<b>100</b>	<b>10</b>	<b>6</b>	<b>DPYDDPTYR (SEQ ID NO: 195)</b>
<b>GCL</b>	<b>1</b>	<b>100</b>	<b>3</b>	<b>3</b>	<b>GGLVKPGASLT (SEQ ID NO: 196)</b>

Las tablas 5-10 muestran los péptidos identificados con alta confianza (>99% de certeza) mediante espectrometría de masas (péptido CDR3) que corresponden a secuencias (específicamente la región CDR3) generadas por secuenciación profunda del repertorio de anticuerpos del animal. El recuento 5 de CDR3 muestra el número de veces que se identificó un péptido a partir de la mezcla de anticuerpos policlonales que coincidía con la región CDR3. La cobertura de CDR3 indicó el porcentaje de esos aminoácidos en la región CDR3 (mostrada en la columna CDR3) que aparecen en los péptidos identificados por espectrometría de masas con respecto a los aminoácidos totales de la región CDR3. Los péptidos totales representan el número total de péptidos por secuencia identificada por espectrometría de masas correspondiente a la secuencia de la región variable completa determinada por secuenciación profunda.

Los péptidos únicos representan el número de péptidos únicos por secuencia identificada mediante espectrometría de masas correspondiente a la secuencia de la región variable completa determinada por secuenciación profunda.

#### Ejemplo 5

En otro ejemplo, se pueden utilizar los siguientes protocolos para generar las secuencias de ácido nucleico y los anticuerpos policlonales. Los resultados muestran el éxito en la generación de un anticuerpo específico de antígeno utilizando estos métodos.

En estos protocolos, los ratones se inmunizaron con un antígeno P-ERK inmunogénico. La base de datos de material genético y la base de datos de péptidos se pueden generar utilizando los siguientes métodos.

#### I. Base de datos de material genético:

**Aislamiento celular**

Los bazos de ratones inmunizados se enjuagaron 5 veces con 5 ml de RPMI/FCS al 10% utilizando una jeringa y una aguja 21G. Las células se congelaron en FCS al 90%/DMSO al 10%. Se aisló un total de 50 - 100 x 10A6 células de cada bazo.

- 5 Aislamiento de ARN y síntesis de ADNc.

Se aisló el ARN total de esplenocitos de acuerdo con el protocolo del fabricante utilizando QIAshredder (Qiagen Núm. Cat. 79654) y RNeasy mini kit (Qiagen, Hilden, Alemania, Núm. Cat. 74104). El ARN se trató con ADNasa en columna según un protocolo de secuenciación de nueva generación convencional. La concentración de ARN total se midió utilizando un espectrofotómetro ND-1000 (NanoDrop, disponible comercialmente de Thermo Scientific, Wilmington, DE).

10 El ARN aislado se utilizó para la síntesis de ADNc de primera hebra mediante transcripción inversa utilizando el sistema Thermoscript RT-PCR (Invitrogen (parte de Life Technologies), Carlsbad, CA Núm. Cat. I 1146-024). El ADNc se sintetizó utilizando 1,5 µg de ARN y cebador oligo dT de acuerdo con el protocolo del fabricante.

**Amplificación de VH y VL.**

- 15 Se utilizó una reacción de PCR en dos etapas para amplificar los genes de VH y VL. Se utilizó una mezcla de cebadores efectores y antisentido degenerados para la primera ronda de PCR y se utilizó un conjunto de cebadores universales para la segunda ronda de PCR. Debido a la gran cantidad de 10 iniciadores efectores degenerados, la PCR de cadena pesada se divide en 8 reacciones separadas. Las secuencias de los cebadores utilizados se muestran a continuación.

- 20 **Cebadores de primera ronda, la cola universal está subrayada**

**Cebadores efectores de cadena pesada:**

V<sub>H</sub>1.1:

ACGAGCTACGCACGAACTGCAGGTRTCCACTCC (SEQ ID NO: 197)

ACGAGCTACGCACGAATAGCAGGTGTCCACTCC (SEQ ID NO: 198)

- 25 ACGAGCTACGCACGARGTACAGGTGTCCACTCC (SEQ ID NO: 199)

ACGAGCTACGCACGAGCYACAGMTGTCCACTCC (SEQ ID NO: 200)

ACGAGCTACGCACGAACTGCAGGTGTCCWMTCC (SEQ ID NO: 201)

V<sub>H</sub>1.2:

ACGAGCTACGCACGARCTRCAGGTGKCACTCC (SEQ ID NO: 202)

- 30 ACGAGCTACGCACGAGCTAWMGGTGTCCACTCC (SEQ ID NO: 203)

ACGAGCTACGCACGACCTCAGGTGTCCACTCC (SEQ ID NO: 204)

ACGAGCTACGCACGAGCTACAGGTGCTCACTCC (SEQ ID NO: 205)

ACGAGCTACGCACGAACTGCAGGTGTCCCTCT (SEQ ID NO: 206)

V<sub>H</sub>1.3:

- 35 ACGAGCTACGCACGAAYTGCAGGTGTCCAYTGC (SEQ ID NO: 207)

ACGAGCTACGCACGAGCTAMMGGTGTCCACTTC (SEQ ID NO: 208)

ACGAGCTACGCACGACTCCTGTCAKTAACKCAGGT (SEQ ID NO: 209)

ACGAGCTACGCACGAAACTGCAGGTGTCTCT (SEQ ID NO: 210)

ACGAGCTACGCACGARCTRCAGGYGTCCACTCT (SEQ ID NO: 211)

- 40 V<sub>H</sub>2:

ACGAGCTACGCACGACCAAGCTGTATCCTTCC (SEQ ID NO: 212)

ACGAGCTACGCACGACCAAGCTGTGTCCRTCC (SEQ ID NO: 213)

V<sub>H</sub>3:

ACGAGCTACGCACGATGTTGACAGYCVTTCCKGGT (SEQ ID NO: 214)

ACGAGCTACGCACGATGTCACAGCCTTCCTGGT (SEQ ID NO: 215)

V<sub>H</sub>4:

5       ACGAGCTACGCACGATTAAAAGGGGTCAGTGT (SEQ ID NO: 216)

V<sub>H</sub>5:

ACGAGCTACGCACGATYTTAAAARGTGTCMAGTGT (SEQ ID NO: 217)

ACGAGCTACGCACGAGTTAAAAGGTGTCCTGTG (SEQ ID NO: 218)

V<sub>H</sub>6-8:

10      ACGAGCTACGCACGACTYTTAAAAGGKGTCCAGWG (SEQ ID NO: 219)

ACGAGCTACGCACGACYTTAMATGGTATCCAGTGT (SEQ ID NO: 220)

ACGAGCTACGCACGACTTTACATGGTTCAAGTGT (SEQ ID NO: 221)

ACGAGCTACGCACGAYTGTCCCTGCATATGTCYT (SEQ ID NO: 222)

V<sub>H</sub>9-15:

15      ACGAGCTACGCACGAATGGCAGCWGCYCCAAG (SEQ ID NO: 223)

ACGAGCTACGCACGATTATCAAGGTGTGCATTGT (SEQ ID NO: 224)

ACGAGCTACGCACGACTTTAAAAGWTGTCCAGKGT (SEQ ID NO: 225)

ACGAGCTACGCACGAGTGACAGTCCTCCTGGTAG (SEQ ID NO: 226)

ACGAGCTACGCACGACTCCTGATGGCAGTGGTT (SEQ ID NO: 227)

20      ACGAGCTACGCACGAAGCTACAGGTATCCAATCC (SEQ ID NO: 228)

**Cebadores antisentido de cadena pesada:**

IgG1:

CACTGGTGTGAGTCAATGCAGAGATGGGGGTGTCG (SEQ ID NO: 229)

IgG2a:

25      CACTGGTGTGAGTCAAGACCGATGGGCTGTTGTT (SEQ ID NO: 230)

IgG2b:

CACTGGTGTGAGTCAACAGACTGATGGGGGTGTTGTT (SEQ ID NO: 231)

IgG3:

CACTGGTGTGAGTCAAGACAGATGGGCTGTTGTT (SEQ ID NO: 232)

30      **Cebador efector de cadena Kappa:**

ACGAGCTACGCACGAGACATYWWGATGACCCAGTCTCC (SEQ ID NO: 233)

**Cebador antisentido de cadena Kappa:**

CACTGGTGTGAGTCACAGTTGGTGCAGCATCAGCCCC (SEQ ID NO: 234)

**Cebadores de segunda ronda, la cola universal está subrayada**

35      **Cebador efector de cadena ligera o pesada:**

CCTATCCCCTGTGTGCCCTGGCAGTCACGAGCTACGCACGA (SEQ ID NO: 235)

**Cebadores antisentido de cadena pesada:**

MID97:

CCATCTCATCCCTGCGTGTCTCCGACTCAGctagtcactCACTGGTGTGAGTCA (SEQ ID NO: 236)

MID81:

CCATCTCATCCCTGCGTGTCTCCGACTCAGAGAGCGTCACCACTGGTGTGAGTCA (SEQ ID NO: 237)

5 MID24:

CCATCTCATCCCTGCGTGTCTCCGACTCAGTAGAGACGAGCACTGGTGTGAGTCA (SEQ ID NO: 238)

**Cebadores antisentido de cadena ligera:**

MID34:

CCATCTCATCCCTGCGTGTCTCCGACTCAGcacgctacgtCACTGGTGTGAGTCA (SEQ ID NO: 239)

10 MID66:

CCATCTCATCCCTGCGTGTCTCCGACTCAGTCACGCGAGACACTGGTGTGAGTCA (SEQ ID NO: 240)

MID57:

CCATCTCATCCCTGCGTGTCTCCGACTCAGCGCGTATACACACTGGTGTGAGTCA (SEQ ID NO: 241)

15 En las secuencias anteriores, las secuencias subrayadas y en cursiva son para amplificación PCR 20 de 2 etapas, las secuencias subrayadas son para la secuenciación 454, las secuencias en negrita son la clave 454, las minúsculas son código de barras para multiplexación y las secuencias en mayúsculas de fuente regular son secuencias específicas de ratón.

Las reacciones de PCR se establecieron utilizando los cebadores anteriores como se describe en la Tabla 11.

**Tabla 11 (Configuración de PCR de cadena pesada de primera ronda)**

Muestra	Cebadores efectores	Cebadores antisentido
1	V <sub>H</sub> 1.1	IgG1, IgG2a, IgG2b, IgG3
2	V <sub>H</sub> 1.2	IgG1, IgG2a, IgG2b, IgG3
3	V <sub>H</sub> 1.3	IgG1, IgG2a, IgG2b, IgG3
4	V <sub>H</sub> 2	IgG1, IgG2a, IgG2b, IgG3
5	V <sub>H</sub> 3	IgG1, IgG2a, IgG2b, IgG3
6	V <sub>H</sub> 5	IgG1, IgG2a, IgG2b, IgG3
7	V <sub>H</sub> 4 Y V <sub>H</sub> 6-8	IgG1, IgG2a, IgG2b, IgG3
8	V <sub>H</sub> 9-15	IgG1, IgG2a, IgG2b, IgG3

20 Para la primera ronda, una reacción de PCR de 50 jIL para la cadena pesada contenía 0,2 (jM de cada cebador efector (5 cebadores efectores por reacción) y 0,2 jIM de cada cebador antisentido (4 cebadores antisentido por reacción), 10 jIL de 5x Tampón de reacción Phusion HF (Finnzymes (parte de Thermos Scientific), Núm. Cat. F-518), 1 jL de ADNc, dNTP 0,2 jM (NEB, Núm. Cat. N0447), 1 (jL de ADN polimerasa Phusion Hot Start II (Finnzymes, Núm. Cat. F-549L) y 28 jL de agua de grado RT-PCR (Ambion (una compañía de Life Technologies), Austin, TX, Núm. Cat. AM9935). Para la primera ronda, una reacción de 50 jL de PPL para la cadena ligera contenía 0,2 jM del cebador efector y 0,2 jM del cebador antisentido, 10 jL de tampón de reacción Phusion HF 5x (Finnzymes, Núm. Cat. F-518), 1 (jL de ADNc, dNTP 0,2 jM (NEB, Núm. Cat. N0447), 1 jL de ADN polimerasa Phusion Hot Start II (Finnzymes, Núm. Cat. F-549L) y 35 jL de agua de grado RT-PCR (Ambion, Núm. Cat. AM9935). El programa de termociclación de PCR fue el siguiente: 98°C durante 2 min; 15 ciclos (98°C durante 0,5 min, 55°C durante 0,5 min,

72°C durante 1 min); 72°C durante 5 min; almacenamiento 4°C. Los productos de PCR se purificaron de acuerdo con el protocolo del fabricante utilizando DNA clean y el kit Concentrator-5 (Zymo Research Co., Irvine, CA, Núm. Cat. DR014).

5 Para la segunda ronda, una reacción de PCR de 50 µL para la cadena pesada contenía 0,2 µM de cebador efector universal y antisentido 10 (µL de tampón de reacción 5x Phusion HF (Finnzymes, Núm. Cat. F-518), 10 µL del producto de PCR de primera ronda purificado, dNTP 0,2 µM (NEB, Núm. de cat. N0447), 1 µL de ADN polimerasa Phusion Hot Start II (Finnzymes, cat. Núm. F-549L) y 19 µL de agua de grado RT-PCR (Ambion, cat. Núm. AM9935). El programa de termociclación fue: 98°C durante 2 minutos, 10 ciclos (98°C durante 0,5 min, 55°C durante 0,5 min, 72°C durante 1 min), 72°C durante 5 min, almacenamiento a 4°C. Segunda ronda, una reacción de PCR de 50 µL para la cadena ligera contenía 0,2 µM de cebador universal efector y universal antisentido, 10 µL de tampón de reacción 5x Phusion HF (Finnzymes, Núm. Cat. F-518), 10 µL del primer producto de PCR purificado, dNTP 0,2 µM (NEB, Núm. Cat. N0447), 1 µL de ADN polimerasa Phusion Hot Start II (Finnzymes, Núm. Cat. F-549L) y 19 µL de agua de grado RT-PCR (Ambion, Núm. Cat. AM9935). El programa de termociclación de PCR fue: 98°C durante 2 minutos, 8 ciclos (98°C) durante 0,5 min, 55°C durante 0,5 min, 72°C durante 1 min); 72°C durante 5 min; 4°C de almacenamiento. Los productos de PCR se purificaron de acuerdo con el protocolo del fabricante utilizando AMPure XP (Agencourt; Beckman Coulter Genomics, Brea, CA, Núm. cat. A63881) y se analizaron utilizando Agilent 2100 BioAnalyzer.

20 Las secuencias de los productos de PCR se traducen después a secuencias de aminoácidos pronosticadas que a continuación se digieren teóricamente (p.ej., con una proteasa y/o un reactivo de escisión química de proteína) para producir fragmentos peptídicos virtuales. Estos fragmentos peptídicos virtuales se utilizan a continuación para generar espectros de masas pronosticados.

## **II. Generación de espectros de masas reales a partir de fragmentos de péptidos de anticuerpos policlonales:**

25 Los anticuerpos policlonales se purifican a partir de los sueros y/o plasma de un animal (p.ej., a partir de los sueros y/o plasma del animal del que se obtienen las secuencias de ácido nucleico). Para purificar los anticuerpos, se utilizan los siguientes métodos:

### Purificación con proteína-G:

30 Se añadió 1 ml de esferas magnéticas de proteína G (Millipore (Billerica, MA), Núm. Cat. LSKMAGG10) a cada uno de cuatro tubos cónicos de 15 ml (Falcon (BD Biosciences, Franklin Lake, NJ), Núm. Cat. 352097). Las esferas en cada tubo se lavaron dos veces con 10 ml de solución salina tamponada con fosfato pH 7,4, Tween-20 al 0,05% (PBST) y tres veces con 10 ml de PBS. Sueros de tres ratones (ID 1262-2, 1262-4, 1263-4) se agruparon y se diluyeron diez veces hasta un volumen final de 6 ml en PBS. A continuación, se añadieron 1,5 ml del suero diluido, combinado a cada tubo de esferas y se incubaron durante la noche a 40°C. El flujo se recogió y se sometió al procedimiento de purificación otras dos veces. Después de recoger el flujo, cada tubo se lavó dos veces con 10 ml de PBST y tres veces con 10 ml de PBS.

35 A continuación, cada tubo se incubó a 4°C durante 30 minutos con 0,5 ml de glicina 0,1 M de pH 2,7 para eluir la IgG. La elución se repitió cinco veces. Todos los eluatos se neutralizaron con Tris 1M de pH 8,5, se sometieron a diáisis durante la noche frente a PBS y se midió la concentración de proteína con un espectrofotómetro ND-1000 (Nanodrop). En total, se purificaron 2,5 mg de IgG.

### Preparación de la columna de antígeno

40 Se lavaron tres veces 5,0 mL de esferas magnéticas de estreptavidina (SA) (Pierce, Núm. Cat. 88817) de nueva aportación con 10 ml de PBS, y se incubaron durante la noche a 40°C con 105 ul de una provisión de partida de 20 mg/ml de péptido p-ERK-biotina (una forma biotinilada de Núm. de catálogo 1150 disponible comercialmente de Cell Signaling Technology, Inc., Danvers, MA) diluido en 5,0 ml de PBS. El flujo continuo se descartó y las esferas se lavaron tres veces con 10 ml de PBS y se dividieron en aliquotas en 10 tubos de 1,7 ml de baja unión (Axygen (Union City, CA), Núm. Cat. MCT-175-L-C). Las esferas en aliquotas se colocaron en una gradilla magnética (Invitrogen, DynaMag) y se eliminó el PBS antes de agregar los sueros diluidos.

### Purificación específica del antígeno:

50 La IgG purificada con proteína G anterior se añadió a las esferas magnéticas de SA acopladas con el péptido P-Erk con biotina. Después de una noche de incubación a 40°C, se recogió el flujo y las esferas se lavaron con tampones que contenían PBS.

55 A continuación la IgG se eluyó con 5 fracciones de 1,5 ml de glicina 0,1 M pH 3,5, a continuación 5 fracciones de 1,5 ml de glicina 0,1 M de pH 2,7, a continuación 5 fracciones de 1,5 ml de glicina 0,1 M de pH 1,8 y se neutralizó con Tris 1 M de pH 8,5. Los eluatos se sometieron a ensayo para determinar P-ERK (es decir, ERK quinasa fosforilada, el antígeno utilizado para inmunizar a los ratones) con placas de 96 pocillos recubiertas con p-ERK-péptido BSA. Las fracciones con actividad se cuantificaron mediante ELISA (Thermo, Núm. Cat. 23300) y se analizaron para determinar la reactividad de p-ERK mediante transferencia Western utilizando productos lisados de células T Jurkat

(p.ej., disponibles comercialmente de American Type Culture Collection o ATCC, Manassas, VA) tratadas con U0126 20 uM durante 1 hora o 200 nM de ácido tetramecanoil-forbol-mirístico (TPA) durante 15 minutos. Las fracciones con la reactividad p-ERK más limpia se analizaron mediante espectrometría de masas.

#### Análisis de espectrometría de masas

- 5 Las fracciones que contenían anticuerpos se digirieron con al menos una proteasa (p.ej., tripsina) y/o al menos una reacción de escisión química de proteína, y los péptidos resultantes se sometieron a análisis utilizando espectrometría de masas. Los métodos de análisis de espectrometría de masas utilizados para analizar los péptidos son convencional y se han descrito anteriormente en detalle (véase, por ejemplo, Patente de Estados Unidos Núm. 7.300.753; Geiger et al., *Nature Methods* 7: 383-385, 2010; Elias y Gygi, *Nature Methods* 4: 207-214, 2007; 10 Keshishian et al., *Molecular and Cellular Proteomics* 6: 2212-2229, 2007).

Como se describió anteriormente (véase, por ejemplo, el Ejemplo 3), los espectros de masas se analizaron utilizando como referencia la información en la base de datos de material genético. Para hacer esto, se recogen los espectros MS2 y a continuación se correlacionan uno por uno con los espectros pronosticados MS2 de las secuencias de referencia (es decir, de la base de datos de material genético) utilizando un programa computacional convencional que encuentra una coincidencia para cada espectro MS2, incluso cuando no es un espectro de buena calidad o una buena coincidencia. Tales programas están disponibles comercialmente. Por ejemplo, el soporte lógico Sequest se puede obtener como parte del paquete de soporte lógico Sorcerer de Sage-N Research, Inc. (Milpitas, CA). Los espectros que se identifican como espectros de buena calidad o buenas coincidencias con la base de datos de material genético se mapean para las secuencias de referencia de la base de datos de material genético. Si un péptido MS2 se puede mapear para más de un componente distinto de la base de datos genética, no está claro qué componente estaba presente en la fracción de anticuerpo polyclonal de unión a antígeno ya que podría ser uno o más de los componentes identificados. Por lo tanto, el procedimiento se repite y, con la repetición, se puede recopilar evidencia para mostrar que algunos componentes se correlacionan mejor con los espectros de EM recolectados que con otros. En otras palabras, gran parte de sus secuencias de región variable se observan como espectros MS2 después del enriquecimiento mediante la unión al antígeno. Se supone que estos elementos codifican anticuerpos de unión a antígeno verdaderos, y así sus secuencias se construyen (p.ej., en un generador de oligonucleótidos sintéticos), se clonian en un plásmido de expresión (p.ej., pADNcS.1 de Invitrogen), se expresan en células y se someten a ensayo para determinar la unión al antígeno.

#### Resultados

30 Como se muestra en la Fig. 5, la correlación de la espectrometría de masa real resulta de los fragmentos de péptido con la información de espectrometría de masas teórica de las secuencias de ácido nucleico que permitió la identificación de las secuencias de fragmentos de cadena pesada y ligera. Se muestran aquellos péptidos que tenían el mayor grado de confianza en lo que se refiere a la cobertura de espectrometría de masas y correlaciones con las secuencias de ácidos nucleicos. La secuencia de ácido nucleico que codifica la cadena completa que comprende los 35 fragmentos peptídicos reales se sintetizó y se clonó en un vector de expresión recombinante. Mediante emparejamiento aleatorio, las cadenas pesada y ligera se combinaron y se expresaron juntas en una célula para producir (es decir, crear) anticuerpos recombinantes (véase, por ejemplo, el método de las Patentes de Estados Unidos Núm. 4.816.397; 4.816.567; y Solicitud de Patente de Estados Unidos Núm. 20110045534). La Figura 6 es una tabla que muestra los resultados de un experimento de ELISA utilizando placas recubiertas con pERK. Como se 40 puede observar, varios emparejamientos de las cadenas identificadas en la Fig. 5 dieron como resultado anticuerpos que fueron capaces de unirse específicamente a las placas recubiertas con p-ERK (los anticuerpos positivos se muestran en color amarillo en la Fig. 6, y los péptidos positivos se muestran en color rojo en la Fig. 5).

Sorprendentemente, estos resultados mostraron que ni la frecuencia de aparición de péptido solo ni la frecuencia del recuento de CDR3 solo pronosticaron el uso de una cadena de anticuerpo concreta que se unía específicamente al antígeno. Por ejemplo, la secuencia de ácido nucleico de cadena ligera con Núm. Ref. G623FKB01A3GC7 se comparó con 235 péptidos de LC-MS/MS (es decir, cromatografía líquida, espectrometría de masas en tandem) y la secuencia de ácido nucleico de cadena ligera con Núm. Ref. G623FKB01AXJ1C tenía una secuencia que aparecía 1068 veces en una única ronda de NGS (véase la Fig. 5, tabla inferior). Sin embargo, ninguna de estas, cuando se combinó con una cadena pesada, en realidad fue capaz de formar un anticuerpo que pudiera unirse específicamente 50 al antígeno pERK. Este resultado es muy sorprendente y mostró que el método de Reddy et al., *Nature Biotechnology* 28(9):965-969, 2010, que dependía únicamente de la frecuencia de la secuencia de ácido nucleico del análisis de NGS, habría pasado por alto la verdadera secuencia de unión al antígeno. Por lo tanto, los métodos descritos en la presente memoria se pueden utilizar de manera fiable para identificar y aislar un anticuerpo que se une específicamente a un antígeno elegido.

#### Ejemplo 6

Se generó un anticuerpo de conejo específico de antígeno de acuerdo con los métodos descritos en la presente memoria. Para hacer esto, se siguieron los siguientes protocolos.

#### Purificación de ARN de esplenocitos de conejo

El antígeno p-MET (Cell Signaling Technology, Inc., Danvers, MA, Núm. de Catálogo 1645) se utilizó para inmunizar conejos utilizando métodos convencionales. Se sacrificaron conejos inmunizados que tenían sueros específicos de antígeno (es decir, sueros que contenían anticuerpos policlonales que se unían específicamente al antígeno inmunizante) después de una inyección de antígeno final (refuerzo). Se recogieron 50 ml de sangre y se recogieron

- 5 el bazo u otros órganos linfoides. Se utilizaron 10 millones de esplenocitos para la purificación de ARN. El suero y/o plasma de la sangre recolectada de 50 ml también se reservó para la purificación por afinidad de anticuerpos específicos de antígenos.

El ARN se purificó a partir de esplenocitos utilizando el kit RNeasy de Qiagen (Qiagen Núm. Cat. 74104) siguiendo el protocolo del fabricante. En la columna, se llevó a cabo el tratamiento con ADNasa I para eliminar el ADN genómico contaminante mediante la incorporación de una etapa de digestión de ADNasa I. Despues del lavado con tampón RW1, se aplicó ADNasa I (Qiagen Núm. Cat. 79254) diluida en tampón RDD a la columna de purificación de ARN y se incubó durante 20 minutos a temperatura ambiente. La columna se lavó una vez más con tampón RW 1, seguido de dos lavados con tampón RPE, y el ARN se eluyó con 30 o 50 μl de agua. La concentración del ARN se determinó mediante la absorbancia medida en un espectrofotómetro Nanodrop (Thermo Scientific) a una longitud de onda de 15 450 nm.

#### **Síntesis de ADNc y generación de amplicones mediante PCR**

El ARN aislado de esplenocitos de conejo se sometió primero a transcripción inversa utilizando la transcriptasa inversa Thermo Script de Invitrogen (Invitrogen Núm. Cat. 12236-022) como se muestra a continuación:

ARN tratado con ADNasa:	5 μL
Cebador Oligo dT (50 μM):	1 μL
dNTP (10 mM):	2 μL
dl H2O:	4 μL

- 20 Incubar a 65°C durante 5 minutos, colocar en hielo durante 2 minutos y a continuación agregar lo siguiente:

5X Tampón de ADNc:	4 μL
DTT 0,1 mM:	1 μL
RNase OUT:	1 μL
dl H2O:	1 μL
ThermoScript:	1 μL

La mezcla se incubó a 50°C durante 1 hora, seguido de una etapa de inactivación por calor a 85°C durante 5 minutos. Finalmente, la cadena de ARN complementaria se eliminó del ADNc mediante la adición de 1 μl de ARNasa H (Invitrogen (Carlsbad, CA), Núm. cat. 18021-071) y se incubó a 37°C durante 20 minutos.

- 25 Los amplicones de las regiones variables de la cadena pesada, kappa y lambda para la secuenciación se generaron mediante PCR de la siguiente manera.

#### **Cebadores de fusión de cadena pesada:**

##### **Inverso**

MID11	CCATCTCATCCCTGCGTGTCCGACTCAGtgatacgctGGGCCAGTGGAAAGACTGATGG (SEQ ID NO: 242)
-------	--

- 30 **Directo**

CCTATCCCCTGTGTGCCTGGCAGTCTCAGatcagacacgATGGAGACTGGGCTGCGCT (SEQ ID NO: 243)

**Cebadores de fusión de cadena kappa**

**Inverso**

MID16	<u>CCATCTCATCCCCTGCGTGTCTCCGACTCAGtacagtactaGAAGAGGAGGACAGWAGGCGC</u> (SEQ ID NO: 244)
-------	--

5    **Directo**

CCTATCCCCTGTGTGCCTGGCAGTCTCAGATGGACATGAGGGCCCC (SEQ ID NO: 245)

Primarios de la fusión de la cadena de Lambda

**Inverso**

MID39	<u>CCATCTCATCCCCTGCGTGTCTCCGACTCAGtacagatgtCTTGTTGCCTTGAGTTCCCTCAGAGGA</u> (SEQ ID NO: 246)
-------	---

10    **Directo**

CCTATCCCCTGTGTGCCTGGCAGTCTCAGATGGCCTGCACCCCG (SEQ ID NO: 247)

En las secuencias anteriores, las secuencias subrayadas son para la secuenciación 454, las secuencias en negrita son la clave 454, las secuencias en minúscula son el código de barras para multiplexación y las secuencias en mayúscula de fuente regular son secuencias específicas de conejo.

15    La amplificación mediante PCR se realizó utilizando la polimerasa Phusion Hot Start II de Finnzyme (Thermo Scientific Núm. Cat. F-540S) donde la mezcla y las condiciones de reacción se configuraron como se muestra a continuación:

**Mezcla de reacción:**

ADNc:	2,5 uL
5X Tampón GC:	5 uL
mezcla de dNTP 10 mM:	0,25 uL
Phusion HotStart II:	0,25 uL
Cebadores (Directo + inverso) 30 uM:	0,25 uL
Agua:	16,75 uL

20    **Programa de PCR:**

Etapa 1 98°C - 1,5 minutos

Etapa 2 98°C - 10 segundos

Etapa 3 60°C - 30 segundos

Etapa 4 72°C - 30 segundos

25    Etapa 5 Repite las etapas 2 a 4, 20 veces

Etapa 6 72°C - 2 minutos

**Etapa 7: mantenimiento**

Para asegurar la ausencia de cualquier amplificación falsa del molde contaminante en cualquiera de los reactivos, se establecieron reacciones duplicadas para cada mezcla (4 reacciones separadas para la cadena pesada, y una para cada cadena ligera) donde el molde de ADNc se sustituyó por agua. Estas reacciones de control negativo sin molde se ejecutaron al mismo tiempo que las muestras que contenían el molde. Tras completar el programa de PCR, se analizaron 3 µl de cada reacción (incluyendo los controles negativos) mediante electroforesis en un gel de agarosa-TAE al 1,5% para determinar la presencia de los amplicones cuando se añadió el molde a la reacción pero no en ausencia de ADNc. La Figura 7 muestra los resultados de estos geles de electroforesis.

**Purificación, análisis, cuantificación y preparación de amplicones para la secuenciación 454**

Para eliminar el exceso de cebadores y/o reguladores de cebador en las muestras de PCR, los amplicones se purificaron utilizando esferas magnéticas de Agentcourt Ampure (Beckman Coulter Núm. Cat. A63881) siguiendo el protocolo del fabricante (000387v001). Los amplicones eluidos después de la purificación Ampure se analizaron a continuación para determinar la pureza y ausencia de cualquier especie de ADN contaminante en el Bioanalizador Agilent 2100 utilizando el chip de ADN de alta sensibilidad (Agilent Technologies Núm. Cat. 5067-4626) siguiendo el protocolo del fabricante.

Una vez que se verificó la pureza de los amplicones, la concentración del ADN se cuantificó en un fluorómetro utilizando el kit de ensayo Quant-iT PicoGreen dsDNA (Invitrogen Núm. Cat. P7589) como se describe en el protocolo del fabricante. El ADN de Lambda proporcionado en el kit se utilizó como un patrón de concentración con el que se generó una curva patrón de 100 ng/pocillo a 1,56 ng/pocillo. La fluorescencia de cada amplicón diluido 100 veces en tampón TE se midió por duplicado, y la concentración de ADN se determinó de acuerdo con la porción lineal de la curva patrón. Todas las medidas de fluorescencia se realizaron en placas de color negro de 96-30 pocillos. Si el valor de la fluorescencia estaba fuera del intervalo lineal de la curva patrón, las muestras se volvieron a medir con diluciones más grandes o más pequeñas con el fin de capturar los valores de fluorescencia que caen dentro del intervalo lineal. utilizando el tamaño aproximado en pares de bases de cada tipo de cadena (pesada-540 pb, kappa-485 pb y lambda-510 pb), se utilizó la siguiente fórmula para determinar la concentración:

$$\text{Concentración de cada amplicón (moléculas/µl)} =$$

$$[\text{conc. de la muestra (ng/µl)} * 6,022 \times 10^{23}] / [656,6 \times 10^9 * \text{longitud del amplicón (pb)}]$$

Cada amplicón se normalizó a  $1 \times 10$  moléculas/µl, a continuación se mezcló a una razón de cadena pesada: cadena kappa: cadena lambda a 3:3:1 en volumen, se sometió a agitación vortical y finalmente se diluyó 1:10 para obtener una concentración final de la mezcla a  $1 \times 10^6$  moléculas/µl.

**Amplificación mediante PCR en emulsión, enriquecimiento de esferas, recuento de esferas y secuenciación**

La PCR en emulsión se realizó siguiendo el protocolo publicado 454: "emPCR Amplification Method Manual - Lib-L" (Edición: mayo de 2010 (Rev. abril de 2011)) con las siguientes modificaciones:

**Sección 3.1.3 Etapa 2)**

Reactivos	Volumen (µl)
Agua de grado Mol. Bio.	458
Aditivo	515
Amp Mix	270
Cebador Amp	32
Mezcla enzimática	70
PPiasa	2
Total	1347

Una vez que se enriquecieron las esferas de secuenciación, a partir del Etapa 3.7 del "emPCR Amplification Method Manual - Lib-L", las esferas se contaron en el Contador de partículas Z2 de Beckman Coulter con las siguientes configuraciones:

Abertura:	<100 µm>
Kd Abertura:	<60,04>
Ajuste corte superior:	<30,00µm>
Ajuste corte inferior:	<10,00µm>
Modo de recuento:	<entre>
Volumen medido:	<0,5 ml>
Resolución:	<256>

- 5 La concentración de esferas se calculó como:

$$\text{Concentración de esferas} = [\text{Lectura media del contador de partículas} * 4] \text{ esferas}/\mu\text{l}$$

Las esferas enriquecidas de la PCR en emulsión se secuenciaron en 454 Sequencer (Roche) siguiendo el protocolo de secuenciación 454 para GS FLX + o GS Junior.

- 10 Los fragmentos peptídicos de los anticuerpos policlonales recogidos de los sueros de 5 conejos inmunizados se generaron como se describió anteriormente para los ratones (véase, por ejemplo, el Ejemplo 6). Brevemente, se utilizó el siguiente protocolo.

#### Purificación por afinidad de péptidos de IgG de conejo

1. Volver a suspender la resina de afinidad de péptidos y llevar 0,4 ml de la suspensión a una nueva columna 10 (Bio-rad, 731-1550, 0,8 x 4 cm), y esto debería constituir 0,2 ml de resina de purificación sedimentada.
- 15 2. Si fuera necesario, preparar una columna de control de resina en blanco o una resina de afinidad peptídica no relacionada de igual volumen. La resina en blanco se preparó sin péptido en el procedimiento de conjugación.
3. Lavar la columna con 10 ml de PBS y dejar que se drene completamente.
  4. Cargar la IgG total purificada con Proteína A. Tapar la parte inferior primero y envolver con parafina. Añadir 3-5 ml de IgG total. Tapar la parte superior y envolver con parafina.
- 20 5. Girar sobre un rodillo durante 15 minutos a temperatura ambiente.
6. Recoger el flujo. Quitar la tapa primero, a continuación la parte inferior, dejar que la columna drene completamente.
  7. Lavar con 10 ml de PBS, 3 veces (lavar la pared de la columna para asegurarse de que toda la resina se empaqueta en la parte inferior).
  8. Lavar con 10 ml de acetonitrilo al 20% en PBS de pH 7,4.
  9. Lavar con 10 ml de etilenglicol al 60% en PBS de pH 7,4.
  10. Lavar con 10 ml de NaCl 2,0 M en PBS, pH 7,4.
  11. Eluir con 5 ml de glicina 0,1 M, pH 3,5, neutralizar inmediatamente con 70 µl de Tris 1M de pH 8,5.
- 25 12. Eluir con 5 ml de glicina 0,1 M pH 2,7, neutralizar inmediatamente con 300 µl de Tris 1M, de pH 8,5.
13. Eluir con 5 ml de glicina 0,1 M pH 1,8, neutralizar inmediatamente con 800 µl de Tris 1M de pH 8,5.
  14. Todas o las fracciones de interés se miden para determinar la concentración de IgG utilizando placas de

ELISA de IgG de conejo (proporcionadas por el grupo de Análisis molecular/ELISA).

15. La actividad específica de antígeno puede evaluarse utilizando ELISA y/o transferencia Western. La actividad específica también se puede evaluar después de normalizar todas las fracciones a la misma concentración.

5 16. Los materiales de anticuerpos purificados están listos para ser procesados para LC-MS/MS

Se realizó un análisis de cromatografía líquida-espectrometría de masas en tandem (LC-MS/MS) sobre los péptidos de los anticuerpos purificados (es decir, se dirigieron los anticuerpos purificados y se analizaron los péptidos) como se describió anteriormente. Los espectros de masas resultantes se correlacionaron con los datos teóricos de espectrometría de masas basados en la información de la base de datos de material genético.

10 Como se muestra en la tabla expuesta en la Figura 8, se identificaron varios péptidos de cadena pesada y ligera correlacionando la espectrometría de masas real (es decir, observada) de los péptidos con los datos teóricos de espectrometría de masas de las secuencias de ácido nucleico. La frecuencia de aparición de estos péptidos se muestra en la calle más a la derecha de la tabla. Estas cadenas se eligieron en función de su cobertura de CDR3 (en la mayoría de los casos, 100%) y las secuencias de nucleótidos subyacentes se recuperaron de la base de datos de material genético y se sintetizaron. Seis cadenas pesadas se combinaron aleatoriamente con cinco cadenas ligeras (mostradas en color rojo en la Fig. 8), y los anticuerpos resultantes se analizaron utilizando ELISA (con placas recubiertas con antígeno) y análisis de Transferencia Western (contra células Hela sin tratar (calles -) o se trataron con Factor de crecimiento humano (calles +), cuando se sabe que las células tratadas con HGF expresan el antígeno p-MET. Los resultados del análisis de transferencia Western se muestran en la Fig. 9. Un anticuerpo específico de p-MET (comercializado por Cell Signaling Technology, Inc., Danvers, MA, Núm. de catálogo 3126) se utilizó como control. Los anticuerpos generados de acuerdo con los métodos descritos en la presente memoria que mostraron alta unión específica al antígeno en los productos lisados celulares se muestran en rojo en negrita en la Fig. 8 (es decir, los Núm. de referencia de cadena pesada GXRYQP201BIQD2 y GXRYQP201A97DZ y los Núm. de referencia de cadena ligera GXRYQP201A291T y GXRYQP201BRIWK y GXRYQP201ALDF5). Obsérvese que la Fig. 9 muestra solo dos de los 6 anticuerpos diferentes que se unen específicamente al antígeno generado en este ejemplo (es decir, la Fig. 9 muestra solo los dos anticuerpos que utilizan la cadena pesada GXRYQP201BIQD2 acoplada con la cadena ligera GXRYQP201A291T y la cadena ligera GXRYQP201BRIWK).

15

20

25

30

De nuevo, como se observó con el anticuerpo de ratón, las cadenas con la frecuencia más alta no dieron como resultado la formación de un anticuerpo específico de antígeno (comparar cadena pesada 5 GXRYQP201A1C3B, que tenía una frecuencia de 9,12% pero no se unía específicamente al antígeno con cadena pesada GXRYQP201 BIQD2 que tenía una frecuencia de solo 0,19%, pero se unía específicamente al antígeno).

#### Ejemplo 7

Este ejemplo describe la generación de anticuerpos monoclonales de conejos inmunizados con cuatro antígenos diferentes y de ratones inmunizados con un antígeno diferente adicional (Tabla 12) utilizando el enfoque descrito anteriormente, demostrando adicionalmente que el presente enfoque es robusto y reproducible en al menos dos especies de animales de laboratorio.

Antígeno	Especie inmunizada	Alta confianza Cadena Pesada+Ligera	Clones ELISA+ Únicos	Clones TW+ únicos
PR A/B	Conejo	8 + 10	12	6
pMET	Conejo	11 + 10	8	4
Lip294	Conejo	7 + 4	8	3
Sox1	Conejo	9 + 5	12	1
p-EL44-47	Ratón	12 + 13	15	3

**Tabla 12:** Anticuerpos monoclonales funcionalmente relevantes contra múltiples dianas identificadas por la plataforma NGS/LC-MS/MS sometida a ensayo mediante ELISA y transferencia Western (TW)

40 Se inmunizaron conejos blancos New Zealand con péptidos específicos del Receptor de Progesterona A/B (PR A/B) humano conjugados con hemocianina de lapa californiana (KLH). La actividad del anticuerpo específico de antígeno en el suero bruto de cada animal se escrutó para seleccionar el conejo con las señales ELISA y de transferencia Western más altas para PR A/B. El suero de este animal se recogió de 20 ml de sangre y el ARN se obtuvo de células B esplénicas. Se aislaron la inmunoglobulina total e (IgG) del suero utilizando una columna de proteína A-

sefarosa, y los anticuerpos policlonales específicos de antígeno se purificaron mediante cromatografía de afinidad utilizando una columna modificada que consistía en péptido específico de antígeno conjugado con esferas de sefarosa. Las IgG unidas se lavaron exhaustivamente con PBS y a continuación se sometieron a eluciones secuenciales con tampones progresivamente ácidos (pH 3,5, pH 2,7 y pH 1,8) (Fig. 10a). Las fracciones de cada elución se recogieron, se neutralizaron, y escrutaron mediante ELISA específico de antígeno y transferencia Western de producto lisado de la línea celular T47D que expresaba PR A/B y la línea celular HT1080 negativa para PR A/B (Fig. 10a). Se descubrió que la actividad específica de la transferencia Western para PR A/B estaba muy enriquecida en la fracción de pH 1,8, en menor medida en la fracción de pH 2,7, y era indetectable en la fracción de pH 3,5 cuando la fracción polyclonal coincidía con la concentración. Por lo tanto, la fracción pH 1,8 se utilizó para el análisis LC-MS/MS.

Para generar una base de datos personalizada de secuencias de la región V de Ig por NGS, se aisló el ARN de los esplenocitos totales recogidos del mismo animal que mostró actividad específica fuerte para PR A/B. Se generaron amplicones de región variable de cadena pesada y ligera de Ig utilizando cebadores de cadena y y de cadena específica de Ig de conejo para amplificar toda la región V. Los cebadores contenían códigos de barras y seguían los requisitos específicos para el diseño del cebador de fusión 454 titanium para la plataforma Roche 454 NGS. Para aumentar el número de secuencias de la región V recolectadas, los autores de la presente invención combinaron tres rondas de secuenciación 454 GS Junior que consistían en cadenas y y K que dieron como resultado un total de 80.000 lecturas filtradas pasadas, de las cuales 44.363 contenían toda la región V y proporcionaron la base para el enfoque proteómico descrito a continuación. Las secuencias recogidas incluyeron 5.279 secuencias únicas de la región 3 determinante de la complementariedad de la cadena y (CDR3) y 11.681 secuencias CDR3 únicas de la cadena K de longitud variable que siguieron una distribución gaussiana. De acuerdo con los datos previos, este conejo utilizó preferentemente VH1 (V1S69-I-V1S40> 64%) seguido de VH4 (V1S44-I-V1S45 -30%) en el reordenamiento de VDJ de cadena pesada (Becker et al., Eur J Immunol 20: 397-402, 1990, Knight, Annu Rev Immunol 10: 593-616, 1992, Mage et al., Dev Comp Immunol 30: 137-153, 2006).

A continuación, se examinó la fracción de pH 1,8 mediante LC-MS/MS basándose en su actividad previa (Fig. 10a). Para maximizar la cobertura de la secuencia, se dividieron 5 µg de anticuerpo polyclonal de manera uniforme y se digirieron por separado mediante quimotripsina, elastasa, pepsina y tripsina. Se recogieron un total de cuatro rondas de LC-MS/MS utilizando un gradiente de 45 minutos utilizando un Orbitrap Velos (Thermo Fisher), produciendo un promedio de 10.000 espectros por ronda (Fig. 10b). Para estimar la tasa de falso descubrimiento (FDR), se utilizó el enfoque diana/señuelo generando una base de datos compuesta de secuencias orientadas directas e inversas (Elias et al., Nat Methods 4: 207-214, 2007), y cada ronda de LC-MS/MS se buscó utilizando el programa SEQUEST (Yates et al., Anal Chem 67: 1426-1436, 1995). Las coincidencias espectrales de péptidos (PSM) se filtraron a una FDR final de <2% utilizando un análisis discriminante lineal (Huttlín et al., Cell 143: 1174- 1189, 2010) teniendo en cuenta la especificidad de la enzima cuando sea posible (quimotripsina/tripsina). Un ejemplo de un péptido CDR3 de cadena pesada de alta confianza identificado utilizando este método se muestra en la Fig. 10c. Se combinaron rondas individuales y se identificaron un total de 2.356 PSM de región V con una FDR de 1,8%.

Una base de datos de secuencias de la región V del anticuerpo es análoga a una base de datos de isoformas de proteínas. Como resultado, los enfoques tradicionales que utilizan secuenciación por perdigonada mediante LC-MS/MS en la que solo se utilizan a menudo unos pocos péptidos para identificar con seguridad una proteína son insuficientes para identificar una secuencia de la región V del anticuerpo en una mezcla de anticuerpos policlónicos. Además, dado que las secuencias de la región V del anticuerpo pueden variar en tan poco como un aminoácido, la precisión de alta masa ayudó a proporcionar confianza adicional en los PSM. Cada PSM de la región V con un error de masa >-5 y <5 ppm determinado por SEQUEST se mapeó para toda la base de datos de la región V para direccionar la redundancia de PSM y la cobertura a través del conjunto de datos (Fig. 10d). Después de volver a mapear, se determinaron el número total de péptidos, el número único de péptidos, la compartición de espectro (péptidos totales que mapean la secuencia/PSM de región V totales), la cobertura total de la secuencia de la región V y la cobertura de CDR3 para cada secuencia de la región V. Para identificar las secuencias de la región V con alta confianza que probablemente se enriquecerán con la mezcla polyclonal, se aplicaron criterios empíricamente rigurosos en el análisis proteómico, que incluyen: a) cobertura general alta (> 65%), b) al menos 12 únicos péptidos debido al alto grado de homología de las secuencias de la región V, y c) alta cobertura de región hipervariable, específicamente, > 95% de cobertura de CDR3. Aunque las secuencias de la región V podían identificarse utilizando una proteasa sola, se descubrió que debido al alto grado de variabilidad en las secuencias de la región V junto con la complejidad impredecible de una mezcla polyclonal, era ventajoso utilizar proteasas múltiples para aumentar la cobertura de la región V. Por ejemplo, como se muestra en la Fig. 10e, múltiples fragmentos peptídicos solapantes de diferentes proteasas contribuyeron a la identificación de la CDR3 completa de ambas secuencias de cadena pesada y ligera.

La identificación de PSM únicos a través de múltiples rondas de proteasas múltiples que mapean para la misma secuencia de la región V aumentaron los recuentos espetrales y la cobertura en toda la secuencia de la región V, proporcionaron mayor confianza de que las secuencias específicas de la región V estaban presentes en la mezcla polyclonal, y aumentaron adicionalmente la confianza en la calidad de la secuencia NGS (Kircher 5 et al., Bioessays 32: 524-536, 2010). Utilizando los criterios de filtración descritos anteriormente, se identificaron un total de diez secuencias de cadena y y ocho K de alta confianza a partir de la fracción de elución a pH 1,8 (Tabla 13).

**100% de Cobertura de CDR3 y  $\geq 65\%$  de Cobertura de la región V**

	Núm. Ref. NGS	Recuento péptidos totales	%Cobertura región variable	Secuencia CDR3	SEQ ID NO:	Clasificación NGS mediante frecuencia CDR3	Línea germinal V(D)J
<b>Cadena γ</b>	G2IXQ001A2083	101	93.69	KIGL	396	212	IGHV1S45, D4-2, J4
	G2IXQ001AG1SI	91	92.04	GFSL	397	76	IGHV1S69, *, J4
	G2IXQ001B1E8P	78	98.26	DUGOL	398	423	IGHV1S45, D3-1, J4
	G2IXQ001BT2NA	70	86.21	DLGNL	399	461	IGHV1S45, D4-1, J4
	G2IXQ001AFBNC	61	87.27	GAL	400	58	IGHV1S64, D4-1, J4
	G2IXQ001AL49Y	59	87.72	DFHL	400	237	IGHV1S45, *, J4
	G2IXQ001BW8Z3	56	89.17	GS1G1EP	401	103	IGHV1S45, D8-1, J2
	G2IXQ001BN9MH	50	82.14	GFAL	402	109	IGHV1S69, *, J4
	G2IXQ001BPNUG	48	81.51	GHDDGYNNVYRL	403	173	IGHV1S69, D8-1, J4
	G2IXQ001BZA42	35	95.54	GFTL	404	1412	IGHV1S69, *, J4
<b>Cadena κ</b>	G2IXQ001BISIKI	93	87.77	LAGYDCTTGDXTA	405	2769	IGKV1S15, J1-2
	G2IXQ001BQM6D	47	95.5	EGGYDCDNGDCFT	406	85	IGKV1S15, J1-2
	G2IXQ001AMV3J	48	92.70	IGLYDORRADM	407	3634	IGKV1S10, J1-2
	G2IXQ001BQJFD	28	98.15	QSTLVSSTDEIV	408	86	IGKV1S10, J1-2
	G2IXQ001BIC1S	28	96.23	QCSIVNSNIE	409	4518	IGKV1S44, J1-2
	G2IXQ001AG4TB	24	65.45	LGSYDCRSDDCNV	410	179	IGKV1S2, J1-2
	G2IXQ001AH72Z	17	86.11	LGAYDQAADNS	411	252	IGKV1S19, J1-2
	G2IXQ001BFYR5	15	72.07	LGTYDCNSADCNV	412	1128	IGKV1S15, J1-2

**Tabla 13. Identificación de cadenas pesada y ligera de alta confianza.** Se identificaron cadenas pesadas y ligeras con 100% de cobertura del espectro de CDR3 y cobertura de la región variable  $\geq 65\%$  global y se clasificaron por orden de confianza según se midió mediante recuento de péptidos totales. También se indican la identidad de secuencia de CDR3 y la determinación de la línea germinal de ratón. Las cadenas pesada y ligera se eligieron para la síntesis, clonación, y expresión génicas de anticuerpos combinatorios para su caracterización. La clasificación NGS indica la clasificación de frecuencia de la secuencia de CDR3 proporcionada identificada en la base de datos NGS para cadena. \* Indica que no se puede identificar ningún posible gen D

A pesar de proporcionar evidencia de la existencia de secuencias de alta confianza en la región V presentes en suero purificado por afinidad, la información directa sobre el emparejamiento análogo de cadena pesada y ligera está ausente de los datos de LC-MS/MS debido a la proteólisis y la reducción de enlaces disulfuro durante la preparación de la muestra. Como resultado, se expresaron todas las combinaciones posibles de pares de cadenas pesadas y ligeras (matriz 8x10 para un total de 80 anticuerpos, en una transfección de placa de 96 pocillos) además de las secuencias pesadas y ligeras de más alto rango observadas mediante la frecuencia NGS y escrutadas para determinar la actividad de unión específica de antígeno al péptido PR A/B mediante ELISA. Un total de 12 pares de cadenas pesadas y ligeras fueron positivos mediante ELISA específico de antígeno (Fig. 1 la). A continuación, se sometió a ensayo cada clon positivo para ELISA específico de antígeno mediante transferencia Western para determinar su especificidad frente a PR A/B expresado endógenamente en productos lisados celulares (Fig. 1 lb). Se encontró que seis clones se unieron específicamente a PR A/B (Fig. 1 lb); dos clones mostraron una señal mucho más fuerte en comparación con la mezcla polyclonal original cuando se analizaron con la misma concentración de anticuerpo. Los clones específicos de antígeno positivos mediante transferencia Western se caracterizaron adicionalmente en ensayos adicionales. Un anticuerpo monoclonal, el clon F9 y el clon CI, exhibieron señal y especificidad superiores en la inmunotransferencia e inmunohistoquímica (IHQ) (Fig. 1 lb-c) y también reaccionaron específicamente en análisis de citometría de flujo (FC) e inmunofluorescencia (IF) donde la mezcla polyclonal fracasó (Fig. 1 Id-e). En contraste, las cadenas y y K seleccionadas en virtud de su rango más alto de NGS no produjeron anticuerpos específicos de antígeno. No se observaron péptidos que contenían CDR3 a partir de las cadenas y y K con clasificación NGS más alta, y ninguna de las secuencias de CDR3 de las 30 cadenas y o K con la clasificación más alta se identificó con alta confianza por enfoque de proteómica de los autores de la presente invención. No se puede descartar que la ausencia de actividad se deba a una falta de emparejamiento cognado, pero el hecho de que

ninguna de estas cadenas se observó por LC-MS/MS sugiere que ninguna de las cadenas de clasificación NGS más alta era específica contra el antígeno. Por lo tanto, en estos experimentos, los anticuerpos específicos de antígeno no pudieron identificarse basándose únicamente en la clasificación NGS.

5 Para visualizar la diversidad clonal, se realizó el análisis filogenético (Dereeper et al., Nucleic Acids Res 36: W465-469, 2008) en secuencias de la región V de cadena pesada y ligera de alta confianza que se muestran en la Tabla 13. Las secuencias estrechamente relacionadas para cadenas pesadas o ligeras se agrupan en grupos discretos. Curiosamente, todos los anticuerpos monoclonales específicos de PR A/B descubiertos en este informe se agruparon estrechamente en el árbol filogenético, muy probablemente debido a la expansión clonal de células B estrechamente relacionadas durante la inmunización. El uso de línea germinal también apoyó esta observación (Tabla 13). Se hicieron observaciones similares en un experimento independiente con un antígeno diferente (Lin28A, Figura 12).

10 Los métodos utilizados en los experimentos descritos en este Ejemplo son los siguientes.

15 **Inmunización y manejo de animales.** Se inmunizaron conejos blancos New Zealand mediante inyección intradérmica con cuatro dosis separadas, cada una con un intervalo de 3 semanas, con una mezcla de péptidos conjugados con hemocianina de lapa californiana derivados de la secuencia de aminoácidos de diferentes regiones de cada antígeno de proteína humana. Los péptidos se conjugaron con KLH activada con Imjetc Maleimide (Thermo-Pierce). Las inmunizaciones con ratones se llevaron a cabo de la misma manera, excepto que la ruta de inmunización fue intraperitoneal y las inyecciones se realizaron con 2 semanas de diferencia. Se extrajo sangre a los 20 3 días después del refuerzo final. El bazo completo de cada animal se recogió en el momento de la eutanasia después de la confirmación de la actividad polyclonal deseada.

25 **Secuenciación de ADN de próxima generación de repertorios de células B de conejo y ratón.** Los esplenocitos de conejos y ratones hiperinmunitizados se recogieron y se lisaron para la purificación total de ARN utilizando el kit RNeasy de Qiagen siguiendo el protocolo del fabricante. El ARN se trató en columna con ADNase I (Qiagen Núm. Cat. 79254) para eliminar el ADN genómico utilizando el protocolo proporcionado. Para generar bibliotecas de 30 amplicones de cadenas pesadas y ligeras a partir de este material a secuenciar con la plataforma 454 Life Sciences (Roche), la RT-PCR se llevó a cabo de la siguiente manera. El ADNc se generó a partir del ARN total del esplenocito como molde utilizando la transcriptasa inversa Thermoscript (Invitrogen Núm. Cat. 12236014) con oligo dT como cebador. Para la secuenciación de IgG de conejo, las regiones variables de las cadenas γ, K1, K2 y X se amplificaron con cebadores de fusión 454 específicos de secuencia (hibridación con el líder en el extremo 5' y secuencias que contienen el extremo 3' requeridas para la identificación y codificación de barras en el formato Lib-L de la plataforma de secuenciación 454) utilizando Phusion® Hot Start II High-Fidelity DNA Polymerase (Finnzymes Oy, Finlandia) con las siguientes etapas: desnaturación-98°C durante 90 segundos; 20 ciclos de [desnaturación-98°C durante 10 segundos; recocido a 60°C durante 30 segundos; extensión-72°C durante 30 segundos]. Para la secuenciación de IgG de ratón, se generaron amplicones de cadena pesada y ligera mediante un procedimiento de PCR de dos etapas. En la primera etapa, las regiones variables de la cadena γ o K se amplificaron (15 ciclos con las mismas condiciones descritas anteriormente para el conejo) con una mezcla de oligonucleótidos degenerados específicos de la familia de genes como cebadores efectores y cebadores antisentido que hibridan con una región altamente conservada al comienzo de la región constante, conteniendo cada cebador efector y antisentido secuencias adaptadoras distintas en su extremo 5'. Cada reacción de la primera ronda se purificó en 35 columna con un kit comercial (Qiagen Núm. Cat. 28104) y a continuación se amplificó adicionalmente con 10 (cadena γ) y 8 ciclos (cadena K) en la segunda etapa utilizando cebadores específicos de la secuencia adaptadora que contienen secuencias en el extremo 3' requeridas para la identificación y codificación de barras en el formato Lib-L de la plataforma de secuenciación 454. Para cualquier especie, se agruparon todas las reacciones de 40 amplificación de cadena ligera para cada animal. Los cebadores en exceso para las muestras de cadena pesada y ligera se eliminaron utilizando el sistema de purificación de ADN Agencourt AMPure XP siguiendo el protocolo 45 proporcionado. La calidad y pureza del conjunto de amplicones después de la eliminación del cebador se verificó en 50 Agilent Bioanalyzer 2100 (Agilent Technologies), y la concentración del ADN se cuantificó con precisión en un fluorómetro utilizando Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen). Siguiendo el Lib-L LV, se llevaron a cabo el protocolo GS FLX Titanium Series de 454 Life Sciences, PCR en emulsión y enriquecimiento de esferas. El número de esferas se contó en el contador de partículas Beckman Coulter Z2, y la biblioteca se secuenció en 454 GS Junior (Roche).

55 **Purificación de afinidad de IgG específica de antígeno.** La IgG total del suero de los conejos hiperinmunitizados (blancos New Zealand) se purificó utilizando esferas de proteína A-Sefarosa (GE Healthcare), a continuación se incubó en rotación durante 15 minutos en una columna con el péptido inmunógeno acoplado covalentemente a esferas de sefarosa. Por flujo de gravedad, la fracción no unida se drenó, y la columna se lavó extensamente con solución salina tamponada con fosfato 1x (PBS) para eliminar la IgG no específica. El conjunto de IgG polyclonal específico de antígeno se eluyó secuencialmente con tampón de glicina 0,1 M/HCl a pH 3,5, seguido de pH 2,7 y finalmente pH 1,8. Cada elución se neutralizó inmediatamente con tampón Tris 1 M (pH 8,5). La IgG total del suero de los ratones hiperinmunitizados se purificó utilizando esferas magnéticas de Proteína-G (Millipore, Núm. Cat. LSKMAGG10), a continuación se incubó durante la noche a 40°C con péptido inmunógeno immobilizado en esferas magnéticas (Pierce, cat Núm. 88817). Utilizando una gradilla de tubos magnéticos (Invitrogen, Núm. cat. 12321D) se lavaron exhaustivamente las esferas con PBS, a continuación se eluyó secuencialmente el anticuerpo unido a la

columna con un pH ácido progresivo como se describe para la purificación de IgG de conejo.

**Digestión con proteasa del anticuerpo purificado por afinidad.** El anticuerpo polyclonal se desnaturizó en urea 8 M en HEPES 20 mM, pH 8, a continuación se redujo en DTT 10 mM durante 1 hora a 55°C. El anticuerpo polyclonal reducido se enfrió a temperatura ambiente (RT) y se realizó la alquilación en presencia de yodoacetamida 20 mM durante 1 hora. Se realizó la digestión con quimotripsina, elastasa y tripsina en presencia de urea 2 M en HEPES 20 mM, pH 8,0, durante una noche a 37°C a una razón de enzima a sustrato de 1:50. La digestión con pepsina se realizó en presencia de ácido acético 3 M a temperatura ambiente durante la noche a una razón de enzima a sustrato de 1:50. Los péptidos digeridos fueron desalados mediante STAGE-TIPS como se publicó anteriormente (Rappaport et al., Anal Chem 75: 663-670, 2003), y analizados mediante LC-MS/MS.

**Espectrometría de masas.** La LC-MS/MS se realizó utilizando el espectrómetro de masas LTQ Orbitrap Velos (Thermo-Fisher). Las muestras se cargaron durante 7 min utilizando un automuestreador Famos (LC Packings) en una columna capilar de sílice fundida vertida a mano (125 µm de diámetro interno X 20 cm) cargada con resina Magic C18aQ (5 µm, 200 Å) utilizando una bomba binaria de la serie Agilent 1100 con un divisor de flujo en línea. La cromatografía se desarrolló utilizando 15 un gradiente binario a 400 nL/min de 5-30% de disolvente B durante 45 min (Disolvente A, ácido fórmico (AF) al 0,25%, Disolvente B, AF al 0,1%, acetonaítrilo al 97%). Se adquirieron veinte espectros de MS/MS de manera dependiente de datos de un espectro maestro precedente en el Orbitrap (300-1.500 m/z a un ajuste de resolución de 6x104) con un objetivo de control de ganancia automática (AGC) de 106. Se utilizó el escrutinio de estado de la carga para rechazar especies cargadas individualmente, y se requirió un umbral de 500 esferas para desencadenar un espectro MS/MS. Cuando fue posible, el LTQ y Orbitrap se hicieron funcionar en modo de procesamiento paralelo.

**Búsqueda de base de datos y procesamiento de datos.** Los espectros de MS/MS se buscaron utilizando el algoritmo SEQUEST (versión 28 rev 12) (Yates et al., Anal Chem 67: 1426-1436, 1995) contra una base de datos híbrida personalizada compuesta de 21.932 secuencias de región V gamma completa y 22.431 kappa completa y secuencias de región constante gamma y kappa concatenadas con 6.358 proteínas de levadura (*S. cerevisiae*, NCBI) y 42 contaminantes comunes, incluyendo varias queratinas humanas, tripsina y quimotripsina. Dado que las secuencias de la región V están muy relacionadas, el proteoma de levadura contribuyó artificialmente con secuencias más diversas a la base de datos de referencia (Beausoleil et al., Nat Biotechnol 24: 1285-1292, 2006) y proporcionó otra fuente de confianza después de filtrar el conjunto de datos final ya que los datos filtrados no deberían incluir péptidos identificados de levadura. Los parámetros de búsqueda incluyeron especificidad parcial para quimotripsina y tripsina y ninguna especificidad para elastasa y pepsina, una tolerancia de masa de ± 50 ppm, una modificación estática de 57,0214 en cisteína y modificación dinámica de 15,9949 en metionina. La tasa de falsos descubrimientos en el conjunto de datos se estimó utilizando el enfoque diana/señuelo (Elias et al., Nat Methods 4: 207-214, 2007). Los conjuntos de datos se filtraron a una FDR de <2% utilizando un análisis discriminante lineal (Hutlin et al., Cell 143: 1174-1189, 2010). Aunque la precisión de masa del Orbitrap excede en gran medida de 50 ppm, cuando se busca con una mayor tolerancia de iones precursores, las identificaciones de péptidos correctos dan como resultado pequeños errores de masa del precursor (± 1 ppm), mientras que las identificaciones peptídicas incorrectas se distribuyen en toda la ventana de 50 ppm. Como resultado, los filtros de masa del precursor restrictivos eliminan selectivamente muchos PSM incorrectos del conjunto de datos.

El análisis posterior a la adquisición se realizó como se describe en el texto. Brevemente, los péptidos que pasaban derivados de las secuencias de la región V se reasignaron a la base de datos de Ig NGS. Para los péptidos que surgieron de productos digeridos con quimiotripsina y tripsina, las coincidencias se limitaron a las derivadas de las escisiones esperadas (KR para tripsina, YWFLMA para quimotripsina). La cobertura de CDR se determinó mediante la identificación de CDR utilizando las reglas definidas por Kabat (Wu et al., J Exp Med 132: 211-250, 1970). En todos los casos, la cobertura se definió como el número total de aminoácidos identificados a partir de péptidos de alta confianza dividido por el número de aminoácidos en la secuencia de la región V madura.

**Clonación, expresión y caracterización de cadenas de inmunoglobulinas identificadas.** Las cadenas γ y κ identificadas mediante el análisis de espectrometría de masas del conjunto de IgG polyclonal purificado por afinidad se clonaron y se expresaron de la siguiente manera. Para cada cadena identificada, se sintetizó la secuencia de ácido nucleico que codificaba el dominio variable completo de FWR1 a FWR4 (Integrated DNA Technologies, Coralville, Iowa). Utilizando PCR de solapamiento, cada permutación de combinación de cadena pesada-ligera se expresó con una secuencia viral 2A que utiliza un mecanismo de salto ribosómico para generar dos polipéptidos a partir de un único marco de lectura abierto (Doronina et al., Mol Cell Biol 28: 4227-4239, 2008, Donnelly et al., J Gen Virol 82: 1027-1041, 2001). Se clonó un solo casete de marco de lectura abierto, en orden de 5' a 3', regiones variables y constantes de cadena ligera, secuencia peptídica 2A del virus Thosea asigna, y dominio variable de cadena pesada en un plásmido de expresión de mamífero dirigido por un promotor CMV que contenía la secuencia líder de cadena y de conejo y regiones constantes de cadena de conejo, 5' y 3' del sitio de clonación, respectivamente. Se transfirieron HEK293 con preparaciones de plásmido que codificaban cada combinación de cadena ligera-pesada ensamblada de esta manera utilizando polietilenimina (Boussif et al., Proc Natl Acad Sci USA 92: 7297-7301, 1995). El sobrenadante se rastreó de 2 a 5 días después de la transfección para la secreción de anticuerpo específico de antígeno mediante ELISA utilizando el péptido inmunógeno como antígeno de recubrimiento, y se caracterizaron adicionalmente las permutaciones de cadena ligera-pesada que mostraron reactividad. Para la expresión del anticuerpo de ratón, las regiones constantes fueron de IgG2a de ratón.

**Caracterización de anticuerpos policlónicos y monoclonales mediante ELISA, Transferencia Western, citometría de flujo, inmunofluorescencia e inmunohistoquímica.** Los protocolos detallados de ELISA, transferencia Western, citometría de flujo, inmunofluorescencia e inmunohistoquímica se pueden encontrar en línea en el sitio web de Cell Signaling Technology Inc. Se utilizaron placas de 96 pocillos de poliestireno de alta unión certificadas Costar Núm. Cat. 3369 para ELISA. Los antígenos utilizados para el análisis ELISA para cada diana fueron los mismos péptidos utilizados para las inmunizaciones. Para los anticuerpos del Receptor de Progesterona, la Transferencia Western se realizó en células T47D (PR+), MDA-MB-231 (PR-) y producto lisado de células HT-1080 (PR-), análisis de citometría de flujo en células T47D (PR+) y MDA-MB-231 (PR-), análisis de inmunofluorescencia confocal en células MCF-7 (PR+) en comparación con células MDA-MB-231 (PR-), y análisis inmunohistoquímico en secciones de carcinoma de mama humano primario embebidas en parafina, células T47D y MCF-7 embebidas en parafina (PR+) en comparación con las células MDA-MB-231 (PR-). Para los anticuerpos de ratón fosfo-p44/42 MAPK, se realizó la transferencia Western sobre el producto lisado de células Jurkat tratadas con U1026 (Cell Signaling Technology, Inc. Núm. Cat. 9903) o 12-O-Tetradecanoilforbol-13-acetato (TPA) (Cell Signaling) Technology, Inc. Núm. Cat. 4174). Para los anticuerpos Lin28A, se realizó la Transferencia Western sobre el producto lisado total de líneas celulares NCCIT, NTERA, MES e IGROV1, inmunofluorescencia confocal y análisis de citometría de flujo en células NTERA (Lin28A-i-) y HeLa (Lin28A-). Para los anticuerpos fosfo-Met (pMet), se utilizaron productos lisados de células MKN45 no tratadas (pMet-i-) y tratadas (pMet-) con inhibidor de quinasa Met SU11274. Para los anticuerpos Soxl, se utilizaron extracto de cerebro de ratón (Soxl-i-) y producto lisado de células NIH-3T3 (Soxl).

## 20 Ejemplo 8

En este Ejemplo, se generaron anticuerpos monoclonales humanos específicos para el antígeno de superficie pequeño del virus de la Hepatitis B (HBsAg) de acuerdo con los métodos descritos en la presente memoria. Para hacer esto, se siguió el siguiente protocolo para generar la base de datos de material 5 genético. Los anticuerpos policlónicos se purificaron como se describe a continuación y se analizaron después del análisis de espectrometría de masas como se describió anteriormente para ratón y conejo.

### I. Generación de las secuencias de ácidos nucleicos.

#### Aislamiento de células B específicas de antígeno, de memoria y totales y purificación de ARN

Se aislaron células mononucleares de sangre periférica (PBMC) a partir de sangre humana entera, de nueva aportación, recogida en tubos de vacío heparinizados, utilizando un gradiente de Ficoll. En un tubo cónico de 50 ml Greiner Leucocep (Sigma Aldrich Núm. Cat. Z642843) que contenía 20 ml de Histopaque 1077 15 (Sigma Aldrich Núm. Cat. 10771) en la parte inferior, se aplicaron hasta 25 ml de sangre en la parte superior, a continuación los tubos se centrifugaron durante 20 minutos a 1500xg a temperatura ambiente. Los leucocitos (capa leucocitaria) se recogieron utilizando una pipeta estéril, se lavaron en medio RPMI dos veces volviendo a suspender las células en 50 ml de RPMI, a continuación se centrifugaron a 300xg durante 10 minutos a 40°C. Después de los lavados, las PBMC se criopreservaron en DMSO al 20% en suero bovino fetal o se procesaron inmediatamente para el aislamiento de células B.

Para el aislamiento de células B, se utilizó un método de selección negativa para eliminar todas las células no B de las PBMC utilizando el kit de aislamiento de células B intactas Dynabeads de Invitrogen (Invitrogen Núm. Cat. 1 13-5 ID) siguiendo el protocolo del fabricante. La población de células B no marcada resultante se procesó adicionalmente para aislar células B específicas de antígeno o de memoria.

Para el aislamiento de células B específicas de antígeno, se incubaron células B no marcadas totales con antígeno biotinilado que está inmovilizado en esferas magnéticas de estreptavidina (Pierce-Thermo Scientific Núm. cat. 88816) en un rotador a temperatura ambiente durante 20 minutos. Las esferas que contenían cualquier célula B que se unía a antígeno se lavaron a continuación dos veces con 1xPBS. Las esferas lavadas se resuspendieron a continuación en tampón de lisis RLT del kit RNeasy de Qiagen (complementado con 1% de β-mercaptoetanol) para el aislamiento de ARN.

Para el aislamiento de células B de memoria, se aislaron células CD27+ e IgG+ de superficie de las células B no marcadas totales utilizando kits MACS de Miltenyi para el aislamiento de células CD27+ e IgG+ de superficie (Miltenyi Biotec (Auburn, CA) Núm. Cat. 130-051-601 y 130-047-501). Con el fin de aislar simultáneamente las células B CD27+ y sIgG +, se añadieron anticuerpos conjugados con esferas magnéticas a ambos marcadores de la superficie celular al mismo tiempo durante la etapa de incubación. Tras la purificación, las células B de memoria se centrifugaron a 300xg durante 10 minutos, y a continuación se lisaron en tampón RLT para ARN como se describió anteriormente para el aislamiento del ARN.

El ARN se purificó a partir de células seleccionadas utilizando el kit RNeasy de Qiagen (Qiagen Núm. Cat. 74104) siguiendo el protocolo del fabricante. En la columna, se llevó a cabo el tratamiento con ADNasa I para eliminar el ADN genómico contaminante mediante la incorporación de una etapa de digestión de ADNasa I. Despues del lavado con tampón RW1, se aplicó ADNasa I (Qiagen Núm. Cat. 79254) diluido en tampón RDD a la columna de purificación de ARN y se incubó durante 20 minutos a temperatura ambiente. La columna se lavó una vez más con

tampón RW 1, seguido de dos lavados con tampón RPE, y el ARN se eluyó con 30 o 50 μl de agua. La concentración del ARN se determinó mediante la absorbancia medida en un espectrofotómetro Nanodrop (Thermo Scientific) a una longitud de onda de 450 nm.

#### Síntesis de ADNc y generación de amplicones por PCR

- 5 El ARN aislado de la memoria o de las células B específicas de antígeno se sometió primero a transcripción inversa utilizando la transcriptasa inversa Thermoscript de Invitrogen (Invitrogen Núm. Cat. 12236-022) como se muestra a continuación:

ARN tratado con ADNasa:	5 uL
Cebador Oligo dT (50 uM):	1 uL
dNTP (10 uM):	2 uL
dl H2O:	4 uL

Se incubó a 65°C durante 5 minutos, se colocó en hielo durante 2 minutos y a continuación se agregó lo siguiente:

5 X Tampón de ADNc:	4 uL
DTT 0,1 uM:	1 uL
RNase OUT:	1 uL
dl H2O	1 uL
Thermo Script:	1 uL

10

La mezcla se incubó a 50°C durante 1 hora, seguido de una etapa de inactivación por calor a 85°C durante 5 minutos. Finalmente, la cadena de ARN complementario se eliminó del ADNc añadiendo 1 μl de ARNasa H (Invitrogen Núm. Cat. 8021-071) e incubando a 37°C durante 20 minutos.

- 15 Los amplicones de las regiones variables de la cadena pesada, kappa y lambda para la secuenciación se generaron mediante PCR de la siguiente manera. Para la amplificación de la cadena pesada, se realizaron 4 reacciones independientes (cada una específica para familias de genes de VH1 y 7; VH2, 5 y 6, VH3 y VH4) para cada muestra de ADNc utilizando los cebadores enumerados a continuación con el fin de preservar la distribución natural de la frecuencia de transcripción del gen VH en el conjunto de células B. Para la amplificación de la cadena kappa y lambda, se realizó una reacción única para cada cadena para cada muestra de ADNc. Para cada reacción, se utilizó 20 una mezcla equimolar de cebadores directos con la misma concentración de cebadores inversos como se indica a continuación. La amplificación se realizó con cebadores de fusión compatibles para 454 Sequencing (Roche) mediante la plataforma Lib-L. Los cebadores inversos se diseñaron para hibridarse con el extremo 5' de la región constante de cada cadena. Estos cebadores contienen las secuencias del cebador Lib-L B y MID de modo que las lecturas de secuenciación comenzarían desde el extremo 5' de cada región constante (en sentido inverso) y en el extremo 3' de la región variable. Para cadenas pesadas y kappa, se utilizó un único cebador inverso para cada MID, mientras que para la cadena lambda, se requirieron dos cebadores inversos distintos para cada MID.
- 25

#### Cebadores de fusión de cadena pesada:

# ES 2 666 301 T3

Inverso

oli551	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag <b>CTGTGCGTGGCA</b> gaa gac Sga TGG GCC CTT GGT GGA (SEQ ID NO: 248)	MID136
oli555	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag <b>ACGGCGAGTAT</b> gaa gac Sga TGG GCC CTT GGT GGA (SEQ ID NO: 249)	MID27
oli602	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag <b>CACGCTACGT</b> gaa gac Sga tgg gcc ctt ggt gga (SEQ ID NO: 250)	MID34
oli606	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag <b>TGAGTCAGTAT</b> gaa gac sga tgg gcc ctt ggt gga (SEQ ID NO: 251)	MID70
oli670	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag <b>ATCTACTGACatgtat</b> gaa gac Sga tgg gcc ctt ggt gga (SEQ ID NO: 252)	MID88
oli671	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag <b>AGTACTGATCtcaca</b> gaa gac Sga tgg gcc ctt ggt gga (SEQ ID NO: 253)	MID83

Directo

VIII/7

oli621	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC TGG ACC TGG AGV ATC (SEQ ID NO: 254)
oli622	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC TGG ATT TGG AGG RTC (SEQ ID NO: 255)
oli623	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC TGC ACC TGG AGG ATC (SEQ ID NO: 256)

oli624	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC TGG ACC TGG AGG KTC (SEQ ID NO: 257)
--------	---

VII2/S/6

oli618	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC ATA CTT TGT TCC ACG C (SEQ ID NO: 258)
oli619	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC ACA CTT TGC TAC ACA C (SEQ ID NO: 259)
oli620	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG TCT GTC TCC TIC CTC ATC T (SEQ ID NO: 260)
oli629	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GGG TCA ACC GCC ATC CTC (SEQ ID NO: 261)

VII3

oli625	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAG TTK GGR CTG AGC TGG (SEQ ID NO: 262)
oli626	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAG TTT KGG CTK AGC TGG (SEQ ID NO: 263)
oli627	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAA CTG GGG CTC CGC TGG (SEQ ID NO: 264)
oli628	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAR TIG GGG CTG WGC TGG (SEQ ID NO: 265)

VII4

oli617	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG AAR CAY CTG TGG TIC TIC CT (SEQ ID NO: 266)
--------	--

**Cebadores de fusión de la cadena kappa**

Inverso

oli552	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag ACGACAGCTC gaa gal gaa gac aga tgg lgc agc eac (SEQ ID NO: 267)	MID77
oli556	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag TCGATCACGT gaa gal gaa gac aga tgg lgc agc eac (SEQ ID NO: 268)	MID42
oli603	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag TACACACACT GAA GAT GAA GAC AGA TGG TGC AGC eac (SEQ ID NO: 269)	MID37
oli607	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag TGTAGTGTGAT GAA GAT GAA GAC AGA TGG TGC AGC eac (SEQ ID NO: 270)	MID71

Directo

oli630	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC ATG AGG GTS CCY GCT CAG CTC (SEQ ID NO: 271)
oli631	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAC ATG AGR GTC CTC GCT CAG CTC (SEQ ID NO: 272)
oli632	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAA GCC CCA GCD CAG CTT CTC (SEQ ID NO: 273)
oli633	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAA ACC CCA GCG CAG CTT CTC (SEQ ID NO: 274)
oli634	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GTG TTG CAG ACC CAG GTC TTC (SEQ ID NO: 275)
oli635	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GGG TCC CAG GTI CAC CTC CTC (SEQ ID NO: 276)
oli636	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG AGG CTC CYT GCT CAG CTC CTG (SEQ ID NO: 277)

**Cebadores de fusión de la cadena Lambda**

## Inverso

oli604	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag CGTAGACTAG AGG GCG GGA ACA GAG TGA CGC TGG (SEQ ID NO: 278)	MID21
oli605	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag CGTAGACTAG AGG GYG GGA ACA GAG TGA CGG AKG (SEQ ID NO: 279)	MID21
oli608	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag TCTAGCGACTAT AGG GCG GGA ACA GAG TGA CGC TGG (SEQ ID NO: 280)	MID45
oli609	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag TCTAGCGACTAT AGG GYG GGA ACA GAG TGA CGG AKG (SEQ ID NO: 281)	MID45
oli553	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag TAGCGCGCGCT agg gcg gca aca gag tga cMg tgg (SEQ ID NO: 282)	MID101
oli554	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag TAGCGCGCGCT agg gYg gca aca gag tga cgg aKg (SEQ ID NO: 283)	MID101
oli557	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag ATAGAGTACT agg gcg gga aca gag tga cMg tgg (SEQ ID NO: 284)	MID33
oli558	CCA TCT CAT CCC TGC GTG TCT CCG AC tcag ATAGAGTACT agg gYg gga aca gag tga cgg aKg (SEQ ID NO: 285)	MID33

## Directo

oli637	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG ACC TGC TCC CCT CTC CTC CTC A (SEQ ID NO: 286)
oli638	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC GGC TTC CCT CTC CTC CTC A (SEQ ID NO: 287)
oli639	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG TCT CCT CTC CTC CTC A (SEQ ID NO: 288)
oli640	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG ACY CCT CTC CTC CTC M (SEQ ID NO: 289)
oli641	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG CCC TGG GCT CTG CTS CTC CTS A (SEQ ID NO: 290)
oli642	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG CCC TGG GTC ATG CTC CTC CTG A (SEQ ID NO: 291)
oli643	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG ACT CCT CTC TTT CTG T (SEQ ID NO: 292)
oli644	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GAG AAG AAG AGG AGA CCT GGG G (SEQ ID NO: 293)
oli645	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG ACC GCT CTC CTC CTC A (SEQ ID NO: 294)
oli646	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG ACC GTC CTC CTC CTC G (SEQ ID NO: 295)
oli647	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCA TGG ATC CCT CTC TTC CTC G (SEQ ID NO: 296)
oli648	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG ATC CCT CTC TTC CTC C (SEQ ID NO: 297)
oli649	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG AYC CCT CTC CTC CTC C (SEQ ID NO: 298)
oli650	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCA TGG GCC ACA CTC CTG CTCC (SEQ ID NO: 299)
oli651	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG ACC CCT CTC TGG CTC A (SEQ ID NO: 300)
oli652	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG GTC TCC TTC TAC CTA C (SEQ ID NO: 301)
oli653	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG ACC CCA CTC CTC CTC C (SEQ ID NO: 302)

oli654	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG GCT CCT CTG CTC CTC A (SEQ ID NO: 303)
oli655	CCT ATC CCC TGT GTG CCT TGG CAG TC tcag ATG GCC TGG GCT CCA CTA CTT CTC A (SEQ ID NO: 304)

La amplificación mediante PCR se realizó utilizando la polimerasa Phusion Hot Start n de Finnzyme (Thermo Scientific Núm. Cat. F-540S) donde la mezcla y las condiciones de reacción se configuraron como se muestra a continuación:

**Mezcla de reacción:**

ADNc:	2,5 uL
5X Tampón GC:	5 uL
Mezcla de dNTP 10 mM:	0,25 uL
Phusion HotStart II:	0,25 uL
Cebadores (Directo + inverso) 30 uM:	0,25 uL
Agua:	16,75 uL

**Programa de PCR:**

- Etapa 1 98°C - 2 minutos
- Etapa 2 98°C - 10 segundos
- 5 Etapa 3 60°C - 30 segundos
- Etapa 4 72°C -30 segundos 20
- Etapa 5 Repite las etapas 2 a 4
- Etapa 6 72°C - 2 minutos
- Etapa 7: mantenimiento
- 10 Para la amplificación de la cadena pesada, se ejecutaron 25 o 30 ciclos (etapa 5 repetida 24 o 29 veces), y para las cadenas kappa y lambda, 20 o 30 ciclos al amplificar el molde de ADNc generado a partir de células B de memoria o de células B específicas de antígeno, respectivamente, puesto que se requieren 5 ciclos suplementarios para una amplificación suficiente a partir de ADNc de células B específicas de antígeno para cada cadena. Para asegurar la ausencia de cualquier amplificación falsa del molde contaminante en cualquiera de los reactivos, se establecieron 15 reacciones duplicadas para cada mezcla (4 reacciones separadas para cadena pesada, y una para cada cadena ligera) donde el molde de ADNc se sustituyó por agua. Estas reacciones de control negativo sin molde se ejecutaron al mismo tiempo que las muestras que contenían el molde. Al completar el programa de PCR, se analizaron 3 µl de 20 cada reacción (incluyendo los controles negativos) mediante electroforesis en un gel de agarosa- TAE al 1,5% para determinar la presencia de los amplicones (aproximadamente 540 pb para la cadena pesada, aproximadamente 485 pb para la cadena kappa y aproximadamente 510 pb para la cadena lambda) cuando se añadió el molde a la reacción, pero no en ausencia de ADNc.

**Purificación, análisis, cuantificación y preparación de amplicones para la secuenciación 454**

- Para eliminar el exceso de cebadores y/o dímeros de cebador en las muestras de PCR, los amplicones se purificaron utilizando esferas magnéticas Agentcourt Ampure (Beckman Coulter Núm. Cat. A63881) siguiendo el 25 protocolo del fabricante (000387v001). Para la cadena pesada, las cuatro reacciones (VH1/7, VH2/5/6, VH3, VH4) se agruparon y purificaron como una muestra, por lo que se purificaron un total de 3 muestras de amplicones (cadenas pesadas, kappa y lambda) para cada ADNc amplificación. El protocolo para la purificación de Ampure se modificó debido a que las purificaciones se realizaron en microtubos únicos de 1,5 ml utilizando una gradilla magnética genérico que es adecuada para tubos de 1,5 ml en lugar de en un formato de placa de 96 pocillos. Todos los volúmenes y otros procedimientos fueron los descritos en el protocolo. Los amplicones eluidos después de la purificación Ampure se analizaron a continuación para determinar la pureza y ausencia de cualquier especie de ADN 30 contaminante en el Bioanalizador Agilent 2100 utilizando el chip de ADN de alta sensibilidad (Agilent Technologies Núm. Cat. 5067-4626) siguiendo el protocolo del fabricante.

- Una vez que se verificó la pureza de los amplicones, la concentración del ADN se cuantificó en un fluorómetro 35 utilizando el kit de ensayo Quant-iT PicoGreen dsDNA (Invitrogen Núm. Cat. P7589) como se describe en el protocolo del fabricante. El ADN de Lambda proporcionado en el kit se utilizó como un patrón de concentración con el que se generó una curva patrón de 100 ng/pocillo a 1,56 ng/pocillo. La fluorescencia de cada amplicón diluido 100 veces en tampón TE se midió por duplicado, y la concentración de ADN se determinó de acuerdo con la porción lineal de la curva patrón. Todas las medidas de fluorescencia se realizaron en placas de color negro de 96 pocillos. 40 Si el valor de la fluorescencia estaba fuera del rango lineal de la curva convencional, las muestras se volvieron a medir con diluciones más grandes o más pequeñas con el fin de capturar los valores de fluorescencia que caen dentro del intervalo lineal. Utilizando el tamaño aproximado en pares de bases de cada tipo de cadena (pesada-540

pb, kappa-485 pb y lambda-510 pb), se utilizó la siguiente fórmula para determinar la concentración:

Concentración de cada amplicón (moléculas/ $\mu$ l) =

$$[\text{conc. de la muestra (ng}/\mu\text{l}) * 6,022 \times 10^{23}] / [656,6 \times 10^9 * \text{longitud del amplicón (pb)}]$$

Cada amplicón se normalizó a  $1 \times 10^7$  moléculas/ $\mu$ l, a continuación se mezcló en una proporción de Hc:Kc:Lc a 3:3:1 en volumen, se sometió a agitación vortical y finalmente se diluyó 1:10 para obtener una concentración final de la mezcla a  $1 \times 10^6$  moléculas/ $\mu$ l.

#### **Amplificación mediante PCR en emulsión, enriquecimiento de esferas, recuento de esferas y secuenciación**

La PCR en emulsión se realizó siguiendo el protocolo publicado 454: "emPCR Amplification Method Manual - Lib-L" (Edición: mayo de 2010 (abril de 2011)) con las siguientes modificaciones:

#### **10 Sección 3.1.3 Etapa 2)**

Reactivos	Volumen ( $\mu$ l)
Agua de grado Mol. Bio.	458
Aditivo	515
Amp Mix	270
Cebador Amp	32
Mezcla enzimática	70
PPiasa	2
Total	1347

Una vez que se enriquecieron las esferas de secuenciación, a partir del Etapa 3.7 del "emPCR Amplification Method Manual - Lib-L", las esferas se contaron en el Contador de partículas Z2 de Beckman Coulter con las siguientes configuraciones:

Abertura:	<100 $\mu$ m>
Kd Abertura:	<60,04>
Ajuste corte superior:	<30,00 $\mu$ m>
Ajuste corte inferior:	<10,00 $\mu$ m>
Modo de recuento:	<entre>
Volumen medido:	<0,5 ml>
Resolución:	<256>

15

La concentración de esferas se calculó como:

$$\text{Concentración de esferas} = [\text{Lectura media del contador de partículas} * 4] \text{ esferas}/\mu\text{l}$$

Las esferas enriquecidas de la PCR en emulsión se secuenciaron en 454 Sequencer (Roche) siguiendo el protocolo de secuenciación 454: "Sequencing Method Manual - GS Junior Titanium Series" - mayo de 2010 (Rev. junio 2010).

## II. Generación de fragmentos de péptidos

### Purificación de IgG específica de antígeno de plasma de donante humano

#### 5 Aislamiento y escrutinio del plasma del donante para determinar la reactividad con antígenos específicos.

Se recogió la sangre completa de voluntarios humanos siguiendo las pautas de IRB en tubos heparinizados. Durante la separación en gradiente de Ficoll de PBMC (como se describió anteriormente), las muestras de plasma se recogieron simultáneamente y se almacenaron a -80°C. La reactividad de IgG plasmática a diversos antígenos se sometió a ensayo mediante ELISA. Brevemente, placas de 96 pocillos de alta unión (Costar Núm. Cat.) se recubrieron con 100 µl/pocillo de antígeno a 2 µg/ml disuelto en tampón de carbonato a 37°C durante dos horas o 4°C durante la noche. Las placas se enjuagaron tres veces con PBS-Tween (0,1%), a continuación se bloquearon con 300 µl/pocillo de leche en polvo desnatada al 5% en PBS-Tween a 37°C durante 1 hora. Las muestras de plasma se diluyeron a 1/100, 1/500 y 1/1000 y 1/2000 en leche al 5% en PBS-Tween, y se añadieron 100 µl de cada dilución en duplicados de pocillos bloqueados de la placa de 96 pocillos y se incubaron durante 2 horas a 37°C. Las placas se lavaron 3 veces con 1xPBS-Tween y se añadieron anticuerpos anti-IgG humana conjugados con peroxidasa de rábano picante (Southern Biotech 2040-05) diluidos 1/4000 en PBS-Tween a cada pocillo (100 µl) y se incubaron a 37°C durante una hora. Las placas se lavaron 6 veces con PBS-Tween y se desarrollaron mediante la adición de 50 µl de solución de sustrato TMB (BioFX Núm. Cat. TMBW-1000-01), seguido de 50 µl de solución de parada (BioFX Núm. Cat. STPR1000-01). Las señales se midieron a una densidad óptica de 450 nm. Los donantes cuyo plasma mostró señal significativa a 1/500 o mayor dilución se seleccionaron para el escrutinio mediante el procedimiento NG-XMT.

El subtipo adw del antígeno de superficie de la hepatitis B (HBsAg) se adquirió de Prospec (Rehovot, Israel, Núm. Cat. HBS-872).

### Purificación de IgG específica de antígeno a partir de IgG plasmática total

#### 25 Purificación con proteína G

1. Se aplicaron 5 ml de suspensión de esferas (2,5 ml de volumen de lecho de esferas) de Protein G Sepharose 4 Fast Flow (GE Healthcare Núm. Cat. 17-0618-05) a una columna de flujo por gravedad y se lavó con 1xPBS dos veces.
2. Se aplicaron a la columna con esferas 5 ml de plasma humano diluido con 1xPBS a 15 ml, y la columna se incubó en un rotador durante la noche a 4°C, o a temperatura ambiente durante 2 horas.
3. La columna se lavó 4 veces con 20 ml de 1xPBS.
4. Se eluyó IgG con 20 ml de tampón de glicina/HCl 0,1 M a pH 2,7, y se recogió en un tubo que contenía 1,2 ml de Tris 1M de pH 8,5 para la neutralización.
5. Se añadieron 10 ml de 1xPBS (pH 7,4) al producto eluido neutralizado para minimizar la precipitación debida a la alta concentración de IgG.
6. La IgG purificada se sometió a diáisis dos veces frente a 4 litros de 1xPBS en un casete de diáisis de corte de 10 kDa (Pierce Núm. Cat. 66456).
7. La concentración de IgG se determinó midiendo la absorbancia a 280 nm en un espectrómetro fotográfico Nanodrop (Thermo Scientific).

#### 40 Purificación de afinidad

1. El HBsAg se conjugó con biotina (Pierce Núm. Cat 20217) siguiendo los protocolos del fabricante. El antígeno conjugado se sometió a diáisis extensamente en 1xPBS.
2. Se incubaron 2 mg de antígeno conjugado con biotina con 5 ml de esferas magnética de estreptavidina (Thermo Scientific Núm. Cat. 8816) durante la noche a 40°C o durante dos horas a temperatura ambiente en un rotador. Las esferas se enjuagaron con 1xPBS dos veces, a continuación se dividieron en nueve tubos de 1,5 ml.
3. La eficacia de la inmovilización del antígeno sobre las esferas se evaluó mediante Elisa HBsAg y mostró consistentemente una unión superior a 80%.
4. A cada tubo que contenía antígeno inmovilizado, se le añadió 1 mg de IgG purificada con proteína G de un solo donante, las esferas se resuspendieron completamente mediante agitación vortical y se incubaron

mediante rotación a temperatura ambiente durante 15 minutos.

- 5 5. Los tubos se colocaron en una gradilla magnética, se separó el sobrenadante y las esferas se lavaron 5 veces con 1 ml de 1xPBS.
6. Después de la última etapa de lavado, se aplicaron 0,9 ml de tampón de glicina-HCl 0,1 M a pH 1,8 a un tubo, se sometieron a agitación vortical y se incubaron a temperatura ambiente durante 5 minutos. Después de 5 minutos, el primer tubo se colocó en la gradilla magnética, a continuación se retiró el tampón ácido del tubo y se colocó en un segundo tubo. Este procedimiento se repitió hasta que se incubaron los nueve tubos con el tampón ácido. La IgG eluida finalmente se recogió en un tubo que contenía 0,14 ml de Tris 1M de pH 8,5 para la neutralización.
- 10 7. Después de que cada tubo se sometió a elución, las esferas se lavaron con 1xPBS dos veces antes de reiniciar la purificación de la etapa en la que se añadió 1 mg de IgG purificada con proteína G a las esferas. El procedimiento se repitió varias veces para generar suficiente material para el tratamiento con proteasa antes del análisis de MS.

### III. Espectrometría de masas

- 15 15. El análisis de espectrometría de masas se realizó como se describió anteriormente. Brevemente, después de la digestión con una proteasa (p.ej., tripsina) y/o un reactivo de escisión química de proteína (p.ej., bromuro de cianógeno), se realizó un análisis de espectrometría de masas sobre los péptidos. Los espectros MS2 resultantes se correlacionaron con los espectros teóricos MS2 derivados de la información en la base de datos de material genético, con el fin de identificar las secuencias genéticas que codifican anticuerpos que se unen específicamente al 20 antígeno de superficie pequeño del virus de la hepatitis B.

### IV. Expresión e identificación de anticuerpos monoclonales

- 25 Se expresaron en un formato combinatorio 24 distintos clones de región variable de cadena pesada (gamma), 20 clones de región variable de cadena kappa distintos y 10 clones de región variable de cadena lambda distintos y se escrutaron para determinar la actividad de unión específica de antígeno (Véanse las Tablas 14-15, donde se incluyen clones de cadena gamma se indican en la columna vertical más a la izquierda y los clones de cadena ligera se indican en la fila horizontal superior). Cada cadena gamma se emparejó con cada cadena ligera (kappa y lambda) para expresar anticuerpos por medio de transfección transitoria de células HEK293E en placas de cultivo tisular convencional de 96 pocillos.

- 30 30. El anticuerpo que se secretó a partir de las células transfectadas en cada pocillo se escrutó para determinar su unión al antígeno de superficie de hepatitis B recombinante purificado (subtipo HBsAg-adw adquirido de Prospec, Ness-Ziona, ISRAEL) mediante un ensayo inmunoabsorbente ligado a enzimas (ELISA). Se recubrieron placas de ELISA de 96 pocillos de alta unión (Costar-3369) con 50 µl/pocillo de HBsAg diluido en tampón de carbonato a 2 µg/ml incubando a 37°C durante dos horas y a continuación se bloquearon con 300 µl/pocillo de leche en polvo al 5% en solución salina tamponada con fosfato (PBS) incubando a 37°C durante una hora. El sobrenadante de las células HEK293E transfectadas transitoriamente se diluyó cinco veces en leche al 5% en PBS con Tween 20 al 0,05% (PBS-T), a continuación se aplicaron 50 µl del sobrenadante diluido a cada pocillo de placas de ELISA recubiertas con HBsAg. Para evaluar la unión no específica, también se aplicó el mismo sobrenadante a placas recubiertas solo con leche al 5% en PBS. Después de la adición del sobrenadante, las placas de ELISA se incubaron a 37°C durante 2 horas seguido de 3 lavados con 250 µl/pocillo de PBS-T. Para detectar cualquier unión del anticuerpo, se añadieron a cada pocillo 50 µl/pocillo de anti-IgG humana conjugada con peroxidasa de rábano (HRP) diluida 4000 veces en PBS-T y se incubaron a 37°C durante una hora. Las placas se lavaron 6 veces como se describió anteriormente, y a continuación se añadieron 50 µl/pocillo de un sustrato cromogénico para HRP, 3,3',5,5'-Tetrametilbenzidina, que se neutralizó con 50 µl/pocillo de ácido aproximadamente 10 minutos después. La señal del sustrato cromogénico neutralizado con ácido se midió mediante absorbancia (densidad óptica) a 450 nm.

- 45 45. Las **Tablas 14-15** muestran los valores obtenidos de la absorbancia de las placas de HBsAg de las cuales se restó la absorbancia de las placas que solo tenían leche en cada pocillo. Las siguientes muestras de sobrenadante se utilizaron como controles (los valores son promedios de dos pocillos independientes en cada caso): positivo = sobrenadante de la transfección de la cadena pesada y ligera del anticuerpo humano anti-HBsAg; negativo = sobrenadante de células transfectadas solo con PEI. Los pocillos con señal mayor que la señal de control negativo 50 10-20 veces, 20-40 veces y más de 40 veces se indican en tonos crecientes de color gris. Treinta permutaciones pesada-ligera mostraron una fuerte reactividad con HBsAg, más de 40 veces el fondo en dos o más de cuatro pocillos, 26 estaban entre 20 y 40 veces el fondo, y 18 estaban entre 10 y 20 veces el fondo (uno de los 18, expresado como una combinación de la cadena gamma EVUGG, que se muestra con \* en la tabla, y a continuación se descubrió que la cadena AKUOL lambda no era reactiva). Por tanto, de los 24 clones de cadena gamma de 55 región variable distintos sometidos a ensayo, 17 expresaron anticuerpos específicos de HBsAg cuando se emparejaron con al menos uno de los 30 clones de cadena ligera sometidos a ensayo.

Tabla 14

## Cadena Kappa

	<b>Cadena Gamma</b>	<b>GRAWI</b>	<b>CHUZK</b>	<b>EDLIN</b>	<b>JPSHA</b>	<b>DO1SE</b>	<b>IKQES</b>	<b>JQIBX</b>	
<b>HSP82</b>		2.86 2.59	2.94 2.71	2.72 0.68	-0.04 0.00	-0.04 0.04	0.01 0.08	2.58 2.62	2.80 2.63
<b>B1RWB</b>		0.92 0.07	0.41 0.05	0.06 0.03	0.03 0.01	0.06 0.08	0.16 0.09	0.71 0.04	0.22 0.02
<b>BDGOL</b>		0.20 0.18	0.09 0.15	1.84 1.89	1.97 2.11	0.77 0.85	0.82 0.83	1.85 1.91	2.05 2.16
<b>H42OK</b>		0.01 0.02	-0.01 -0.01	0.03 0.04	0.05 0.05	0.05 0.10	0.03 0.06	-0.02 -0.03	0.00 0.00
<b>CP4Y8</b>		0.02 0.04	0.02 0.06	0.02 0.06	0.05 0.06	0.04 0.05	0.04 0.07	0.03 0.04	0.02 0.02
<b>GBOGO</b>		0.06 0.13	0.04 0.06	0.06 0.07	0.05 0.05	0.03 0.08	0.06 0.08	0.05 0.05	0.04 0.06
<b>JJESP</b>		0.15 0.05	0.05 0.05	0.10 0.06	0.12 0.08	0.04 0.05	0.07 0.08	0.15 0.06	0.06 0.05
<b>APABY</b>		-0.06 0.05	0.04 0.08	0.03 0.09	0.05 0.04	0.03 0.05	0.04 0.04	0.04 0.05	0.07 0.13
<b>H9L3G</b>		0.75 0.50	0.71 0.93	2.23 2.46	2.59 2.48	0.97 1.50	1.27 1.53	2.55 2.64	2.63 2.66
<b>EZKNT</b>		0.06 0.28	0.05 0.07	0.18 0.13	0.13 0.10	0.84 0.80	0.72 1.05	0.17 0.14	0.31 0.18
<b>A9F9Q</b>		2.60 2.57	2.47 2.62	0.23 0.15	0.10 0.05	0.05 0.06	0.07 0.04	0.11 0.06	0.25 0.29
<b>JXGB9</b>		0.11 -0.04	0.24 0.03	0.05 0.01	0.05 0.02	0.03 0.03	0.06 0.03	0.24 0.06	0.43 0.05
<b>JD1J9</b>		-0.17 0.01	0.02 0.03	0.02 0.04	0.02 0.05	0.02 0.03	0.02 0.03	0.01 0.02	-0.02 0.00
<b>AGTR9</b>		0.01 0.03	0.03 0.02	0.03 0.03	0.04 0.05	0.02 0.03	0.03 0.04	0.04 0.05	0.05 0.05
<b>EOZZF</b>		0.05 0.06	0.04 0.04	0.03 0.05	0.04 0.06	0.04 0.05	0.05 0.04	0.05 0.05	0.01 0.01
<b>ASNDH</b>		0.05 0.15	0.06 0.04	0.04 0.03	0.04 0.06	0.05 0.14	0.05 0.03	0.05 0.08	0.02 0.04
<b>G7MQM</b>		0.16 0.22	0.11 0.24	0.01 0.04	0.02 0.06	0.04 0.08	0.02 0.07	0.18 0.05	0.24 0.63
<b>IJDRC</b>		0.02 0.05	0.04 0.04	0.04 0.03	0.05 0.05	0.05 0.04	0.03 0.04	0.05 0.04	0.05 0.04
<b>D4LT2</b>		0.06 0.04	0.06 0.03	0.03 0.04	0.04 0.04	0.04 0.05	0.05 0.07	0.03 0.05	-0.01 0.06
<b>EVUGG</b>		0.06 0.03	0.05 0.04	0.06 0.03	0.03 0.02	0.04 0.05	0.05 0.02	0.05 0.04	0.03 0.04
<b>DUIFY</b>		-0.01 0.01	0.04 0.03	0.01 0.03	0.05 0.05	0.03 -0.03	0.04 0.02	-0.01 0.01	-0.04 0.00
<b>DTFSS</b>		0.03 0.03	0.05 0.03	0.04 0.03	0.26 0.25	0.20 0.20	0.03 0.03	0.04 0.04	0.05 0.04
<b>DZ3YC</b>		0.04 0.10	0.05 0.03	0.02 0.03	0.04 0.05	0.04 0.05	0.03 0.04	0.04 0.03	-0.02 0.06
<b>CIZNS</b>		0.04 0.05	0.04 0.03	0.03 0.04	0.02 0.05	0.04 0.05	0.05 0.04	0.05 0.07	0.02 0.01
	<b>GRAWI</b>	<b>CHUZK</b>	<b>EDLIN</b>	<b>JPSHA</b>	<b>DO1SE</b>		<b>IKQES</b>	<b>JQIBX</b>	

## Cadena Kappa

Tabla 14 continuación

## Cadena Kappa

	HRONB	BAHZA	HSHZE	H6RCI	FPPWBB	HHFLL	CRY2A
<b>HSPB2</b>	0.04	0.04	0.03	0.03	0.04	0.03	0.02
	0.04	0.03	0.03	0.03	0.04	0.03	0.02
<b>B1RWB</b>	0.04	0.03	0.04	0.02	0.02	0.03	0.04
	0.03	0.02	0.03	0.03	0.03	0.09	0.07
<b>BD00L</b>	1.18	1.06	0.81	1.04	1.52	1.46	2.58
	2.06	1.51	0.59	0.95	2.01	2.05	2.65
<b>H420K</b>	2.68	2.64	0.04	0.04	0.04	0.04	0.04
	2.48	2.39	0.04	0.03	0.03	0.02	0.01
<b>CP4Y8</b>	0.09	0.04	0.02	0.04	0.03	0.02	0.01
	0.19	0.14	0.03	0.03	0.02	0.03	0.01
<b>GBOGO</b>	2.23	2.36	0.05	0.03	0.03	0.04	0.04
	0.21	0.16	0.03	0.03	0.05	0.02	0.03
<b>JJESP</b>	0.07	0.08	0.03	0.02	0.03	0.05	0.06
	0.04	0.03	0.03	0.04	0.05	0.04	0.05
<b>APABY</b>	0.03	0.03	0.02	0.04	0.04	0.03	0.09
	0.04	0.03	0.03	0.04	0.03	0.01	0.01
<b>H9L3G</b>	-0.33	0.38	-0.11	0.20	0.37	0.35	2.34
	0.75	0.75	0.38	0.73	0.84	0.95	2.41
<b>EZKNT</b>	0.33	0.33	-0.06	0.05	0.05	0.03	0.35
	0.44	0.73	0.02	0.04	0.03	0.20	0.26
<b>A9F9Q</b>	-0.30	-0.18	0.03	0.02	0.03	0.07	0.07
	-0.38	-0.01	0.02	0.03	0.04	0.03	0.02
<b>JXGB9</b>	-0.94	0.01	-0.05	0.02	0.03	0.03	0.04
	-0.02	0.03	0.02	0.01	0.03	0.02	0.01
<b>JD1J9</b>	0.01	0.00	0.01	0.02	0.01	0.02	0.02
	-0.03	0.02	0.02	0.02	0.02	0.01	0.01
<b>A6TR9</b>	0.03	0.03	0.02	0.02	0.03	0.04	0.04
	0.02	0.02	0.02	0.01	0.01	0.01	0.03
<b>EOZ2F</b>	0.01	0.01	0.01	0.01	0.01	0.01	0.06
	0.03	0.01	0.01	0.02	0.02	0.01	0.02
<b>ASNDH</b>	0.01	0.00	0.01	0.01	0.00	0.02	0.04
	-0.02	-0.04	0.00	-0.16	-0.18	-0.26	-0.08
<b>G7MQM</b>	0.03	0.05	0.11	0.09	0.61	0.45	0.18
	0.05	0.07	0.13	0.25	0.89	0.52	0.59
<b>IJOR6</b>	0.06	0.06	0.03	0.04	0.01	0.02	0.05
	-0.05	0.00	-0.01	-0.01	0.01	0.04	0.00
<b>D4LT2</b>	0.00	0.04	0.03	0.04	-0.02	0.00	0.04
	0.00	0.01	0.05	0.00	0.04	0.06	0.06
<b>EVUGG</b>	0.04	0.06	0.04	0.07	0.04	0.06	0.06
	0.00	-0.01	0.03	-0.04	0.03	0.06	0.01
<b>DUIFV</b>	0.03	0.03	0.03	0.03	0.04	0.05	0.05
	0.04	0.05	0.06	0.08	0.05	0.06	0.04
<b>DTF5S</b>	0.05	0.04	0.04	0.05	0.05	0.07	0.06
	0.04	0.05	0.05	0.03	0.07	0.06	0.04
<b>DZ3YC</b>	0.03	0.04	0.00	0.04	0.05	0.06	0.05
	0.03	0.03	0.03	0.04	0.05	0.06	0.04
<b>CIZNS</b>	0.02	0.04	0.05	0.05	0.04	0.05	0.05
	0.03	0.02	0.02	0.05	0.03	0.03	0.02
	HRONB	BAHZA	HSHZE	H6RCI	FPPWBB	HHFLL	CRY2A

## Cadena Kappa

Tabla 14 continuación

## Cadena Kappa

	<b>ARW3R</b>	<b>GRBL0</b>	<b>JAVSR</b>	<b>BLP3K</b>	<b>AEJTS</b>	<b>G00Z5</b>						
<b>HSP82</b>	0.05 0.06	0.04 0.06	0.06 0.07	0.05 0.06	0.09 0.07	0.05 0.05	0.02 0.08	0.01 0.02	-0.02 0.02	0.01 0.00	0.00 0.02	0.00 0.04
<b>B1RW8</b>	0.03 0.03	0.06 0.08	0.06 0.17	0.07 0.06	0.07 0.07	0.08 0.06	0.06 0.08	0.10 0.08	-0.02 0.01	0.02 0.12	0.21 0.08	0.18 0.18
<b>BD00L</b>	2.11 1.95	2.38 2.12	2.59 2.53	2.69 2.61	2.50 2.53	2.58 2.38	0.66 0.49	0.38 0.42	0.21 0.14	0.86 0.63	0.95 0.91	0.60 0.35
<b>H420K</b>	0.07 0.03	0.07 0.05	0.06 0.03	0.05 0.03	0.04 0.04	0.04 0.04	0.05 0.05	0.05 0.05	-0.01 -0.11	0.00 0.00	0.13 0.01	0.03 0.01
<b>CP4Y8</b>	0.05 0.04	0.01 0.05	0.06 0.08	0.05 0.05	0.06 0.06	0.03 0.00	0.04 0.00	0.04 0.00	-0.02 -0.01	-0.02 0.00	0.00 0.01	0.01 0.01
<b>GBOGO</b>	0.06 0.03	0.00 0.05	0.06 0.07	0.07 0.08	0.08 0.06	0.08 0.07	0.06 0.06	0.06 0.06	-0.03 0.00	0.00 0.03	0.00 0.01	0.01 0.01
<b>JJESP</b>	0.04 0.08	0.06 0.07	0.08 0.06	0.08 0.06	0.07 0.07	0.07 0.06	0.08 0.06	0.05 0.03	-0.01 -0.01	0.04 0.02	0.01 0.12	0.04 0.03
<b>APABY</b>	0.05 0.04	0.07 0.02	0.06 0.00	0.05 0.02	0.01 0.00	0.06 -0.01	0.07 0.01	0.07 0.04	0.03 0.03	0.00 -0.04	-0.01 -0.03	0.00 0.01
<b>H9L3G</b>	0.30 0.80	0.58 1.11	2.22 2.41	2.39 2.65	0.36 1.01	0.51 1.14	0.04 0.08	-0.05 0.05	0.09 0.16	0.17 0.14	0.20 0.49	0.28 0.44
<b>EZKNT</b>	0.04 0.03	0.05 0.30	0.25 0.42	0.61 0.05	0.07 0.06	0.09 0.04	0.07 0.04	0.31 0.23	0.07 0.20	0.08 0.16	0.07 0.17	0.07 0.17
<b>A9F9Q</b>	0.04 0.06	0.05 0.05	0.12 0.06	0.14 0.04	0.05 0.05	0.09 0.05	0.06 0.05	0.05 0.03	0.13 0.16	0.09 0.12	0.09 0.12	0.10 0.11
<b>JXGB9</b>	0.04 0.00	0.03 0.02	0.05 0.03	0.02 0.02	0.03 0.03	0.04 0.04	0.04 0.07	0.04 0.13	0.19 0.06	0.10 0.06	0.09 0.07	0.22 0.08
<b>JD119</b>	0.03 0.03	0.05 0.03	0.32 0.40	0.32 0.02	0.07 0.06	0.06 0.03	0.04 0.01	0.02 0.03	0.01 0.04	0.05 0.04	0.02 0.04	-0.02 0.04
<b>ASTR9</b>	0.03 0.02	0.03 0.01	0.04 0.05	0.05 0.06	0.07 0.05	0.12 0.07	0.05 0.05	0.03 0.04	0.17 0.04	0.04 0.06	0.05 0.05	0.01 0.07
<b>EOZ2F</b>	0.05 0.04	0.05 0.06	0.06 0.05	0.04 0.05	0.05 0.05	0.06 0.05	0.06 0.05	0.06 0.01	0.06 -0.02	0.05 0.03	0.04 0.04	0.06 0.05
<b>ASNDH</b>	0.04 0.02	0.04 0.02	0.07 0.03	0.29 0.05	0.04 0.01	0.06 0.01	0.07 0.04	0.07 0.05	0.04 0.00	0.03 -0.03	0.08 0.22	0.06 0.54
<b>G7MQM</b>	0.06 0.05	0.12 0.09	0.06 0.04	0.07 0.06	0.07 0.07	0.07 0.03	0.04 0.00	0.02 0.09	0.04 0.09	0.10 0.62	2.56 2.49	2.50 2.88
<b>I3OR6</b>	0.04 0.03	0.05 0.07	0.04 0.05	0.07 0.05	0.08 0.05	0.03 0.05	0.02 0.02	0.03 0.03	0.03 0.04	0.03 0.04	0.09 0.05	0.15 0.08
<b>D4LT2</b>	0.05 0.04	0.06 0.04	0.07 0.05	0.05 0.05	0.05 0.05	0.05 0.04	0.05 0.04	0.04 0.02	0.04 0.04	0.04 0.03	0.04 0.04	0.04 0.04
<b>EVUGG</b>	0.03 0.01	0.05 -0.01	0.04 -0.01	0.04 0.01	0.04 0.03	0.04 0.02	0.04 0.03	0.03 0.04	0.04 0.04	0.05 0.05	0.04 0.05	0.04 0.05
<b>DUIFV</b>	0.03 0.03	0.06 0.04	0.04 0.05	0.05 0.06	0.04 0.05	0.04 0.05	0.02 0.01	0.02 0.02	0.00 0.04	0.03 0.04	0.03 0.02	0.03 0.02
<b>DTFS5</b>	0.04 0.05	0.04 0.06	0.05 0.06	0.04 0.05	0.06 0.06	0.03 0.05	0.04 0.04	0.02 0.03	0.01 0.03	0.05 0.03	0.03 0.04	0.04 0.03
<b>DZ3YC</b>	0.03 0.03	0.04 0.04	0.06 0.03	0.05 0.05	0.05 0.04	0.06 0.05	0.04 0.04	0.04 0.02	0.03 0.02	0.05 0.04	0.06 0.04	0.05 0.04
<b>CIZNS</b>	0.01 0.02	0.03 0.02	0.05 0.05	0.04 0.02	0.03 0.03	0.04 0.04	0.02 0.03	0.03 0.03	0.04 0.04	0.04 0.04	0.03 0.04	0.03 0.03

## Cadena Kappa

**Tabla 15**

Cadena Lambda													Cadena Gamma																						
HERFZ	ECGVK	FQB4M	B5BPK	G09XX	C85FM	GRBRE	FSK3G	AKUOL	DPJJC	HSPB2	SIRWB	BD001	H42DK	CP4YB	GB060	J3ESP	APABY	HBL3G	EZKNT	A9F9Q	JXGB8	JB119	AGTR9	ED23F	ASNDH	G7MQM	EDN9B	D4L1Z	EVUGM	DUIFV	OTF55	D23YC	CIZNS		
-0.01	0.01	0.03	0.06	0.01	0.01	-0.04	0.02	-0.02	0.03	-0.01	0.02	-0.02	0.03	0.03	1.16	0.19	0.00	-0.02																	
0.04	0.02	0.04	0.05	0.02	0.04	-0.01	0.01	0.04	0.03	0.03	0.03	0.03	0.04	0.06	0.57	0.07	0.01	0.00																	
0.03	0.17	0.08	0.07	0.13	0.04	0.06	0.01	0.01	0.02	0.03	0.03	0.04	0.03	0.03	0.03	0.08	0.17	0.04	0.03																
0.02	0.14	0.04	0.07	0.01	0.04	0.02	0.03	0.03	0.02	0.03	0.03	0.04	0.03	0.03	0.03	0.28	0.10	0.03	-0.03																
0.15	0.27	0.91	0.65	0.09	0.31	0.33	0.11	0.04	0.09	0.18	0.17	0.36	0.03	0.39	0.36	0.38	0.38	0.05	-0.01																
0.16	0.21	1.28	0.45	0.28	0.25	0.25	0.13	0.04	0.14	0.29	0.09	0.64	0.01	0.36	0.72	0.28	0.12	0.05	-0.02																
0.26	0.04	0.08	0.05	0.03	0.03	0.02	0.02	0.03	0.06	0.03	0.05	0.01	0.04	0.04	0.03	0.23	0.08	0.02	0.00																
0.02	0.03	0.02	0.04	0.02	0.03	-0.01	0.00	0.02	0.02	0.03	0.05	0.02	0.02	0.04	0.04	0.39	0.35	0.03	0.03																
0.00	0.06	0.02	0.02	0.02	0.11	0.01	0.00	0.04	0.02	0.01	0.03	0.03	0.03	0.03	0.01	0.45	0.00	0.01																	
0.01	0.06	0.02	0.01	0.02	0.06	0.04	0.07	0.03	0.02	0.02	0.03	0.02	0.05	0.04	0.18	0.01	-0.02																		
0.02	0.01	0.01	0.04	0.04	0.02	0.00	0.00	0.03	0.04	0.03	0.04	0.03	0.04	0.05	0.05	0.13	0.02	-0.01																	
-0.01	0.01	0.23	0.00	0.04	0.02	0.01	0.00	0.02	0.02	0.03	0.03	0.08	0.04	0.04	0.04	0.14	0.02	0.02																	
0.05	0.05	0.04	0.08	0.02	0.04	0.02	0.01	0.04	0.04	0.03	0.03	0.04	0.04	0.06	0.09	0.08	0.04	-0.03																	
0.06	0.03	0.85	0.26	0.02	0.04	0.00	0.03	0.04	0.04	0.03	0.03	0.02	0.04	0.03	0.11	0.03	-0.02																		
0.02	0.02	2.51	2.77	0.02	0.03	0.02	0.05	0.02	0.03	0.03	0.03	0.27	0.02	0.01	0.08	0.35	-0.01	0.01																	
0.00	0.05	2.82	2.64	-0.01	-0.01	-0.03	0.03	-0.02	0.02	0.01	0.02	0.03	0.02	0.02	0.06	0.17	0.04	0.08																	
0.19	0.34	0.25	0.20	0.13	0.15	0.09	0.07	0.02	0.04	0.08	0.14	0.08	0.11	0.08	0.06	0.13	0.16	0.01	-0.04																
0.37	1.02	0.56	0.37	0.15	0.14	0.18	0.16	0.05	0.12	0.48	0.44	0.31	0.32	0.22	0.22	0.57	0.48	0.04	0.01																
3.65	1.73	1.71	1.39	1.82	2.00	0.12	0.09	2.01	2.03	0.04	0.03	0.05	0.05	2.25	2.34	1.32	0.99	0.07	0.01																
2.64	2.85	1.88	1.98	1.94	2.33	0.08	0.11	2.09	2.71	0.03	0.03	0.04	0.04	3.61	2.14	0.72	0.74	0.04	0.02																
0.19	0.13	0.13	0.09	0.11	0.09	0.13	0.14	0.04	0.04	0.02	0.03	0.02	0.03	0.03	0.04	0.16	0.14	0.03	0.02																
0.11	0.10	0.10	0.11	0.10	0.11	0.11	0.15	0.01	0.01	0.03	0.02	0.02	0.00	0.02	0.02	0.28	0.09	0.03	0.00																
0.15	0.13	0.10	0.09	0.05	0.07	0.11	0.10	0.03	0.03	0.03	0.03	0.04	0.03	0.03	0.03	0.09	0.05	0.02	0.02																
0.08	0.11	0.09	0.15	0.08	0.06	0.06	0.13	0.07	0.05	0.01	0.05	0.01	0.05	0.04	0.05	0.04	0.02	0.10	0.07																
0.09	0.01	0.05	0.05	0.05	0.04	0.10	0.05	0.08	0.07	0.07	0.08	0.08	0.02	0.03	0.45	0.23	2.82	2.57	0.07	0.01															
0.06	0.04	0.04	0.04	0.49	0.10	0.02	0.02	0.35	0.38	0.10	0.42	0.03	0.03	0.25	0.82	2.78	2.62	0.09	0.02																
0.06	0.05	0.04	0.06	1.88	1.83	0.10	0.06	2.30	2.51	2.43	2.49	0.04	0.09	2.34	2.62	2.12	2.30	0.06	0.09																
0.06	0.03	0.01	0.06	1.69	1.79	0.04	0.02	2.32	2.44	2.46	2.43	0.05	0.02	2.48	2.85	0.57	0.45	0.25	0.02																
0.03	0.06	0.07	0.12	0.04	0.11	0.06	0.01	0.33	0.08	0.84	0.03	0.03	0.02	0.41	0.19	0.12	0.15	2.38	2.50																
0.04	0.03	0.07	0.10	0.06	0.07	0.07	0.04	0.07	0.05	0.04	0.03	0.01	0.01	0.04	0.07	2.00	2.10	2.52	2.50																
0.05	0.06	0.07	0.09	0.07	0.03	0.01	0.28	0.08	0.09	0.09	0.06	0.03	0.41	0.43	2.73	2.74	0.05	0.03																	
0.09	0.14	0.17	0.18	0.10	0.12	0.06	0.06	0.14	0.12	0.09	0.14	0.04	0.03	0.30	0.36	2.49	1.57	0.03	0.05																
0.04	0.04	0.13	0.15	0.08	0.03	0.72	0.56	0.05	0.21	0.05	0.08	0.03	0.03	1.47	1.37	0.35	0.21	0.01	-0.01																
0.06	0.91	1.73	1.87	0.81	1.18	1.77	3.21	0.29	0.19	0.11	0.12	0.04	0.03	1.90	2.02	0.72	0.23	0.01	-0.02																
0.05	0.03	0.05	0.07	2.28	2.21	0.05	0.00	2.29	2.68	2.39	2.49	0.31	0.04	2.28	2.42	0.55	0.37	0.03	0.02																
0.04	0.02	0.07	0.09	2.21	2.25	0.08	0.03	2.25	2.67	2.44	2.53	0.33	0.03	2.34	2.32	0.72	0.28	0.02	-0.02																
0.05	0.04	0.06	0.07	0.04	0.04	2.66	2.73	0.08	0.04	0.03	0.03	0.03	0.19	0.26	0.37	0.23	0.02	0.01																	
0.04	0.01	0.08	0.04	0.03	0.03	2.04	2.89	0.00	0.01	0.03	0.03	0.04	0.32	0.31	0.40	0.28	0.03	-0.02																	
0.04	0.03	0.05	0.04	0.04	0.06	0.05	0.01	0.02	0.04	0.03	0.02	0.04	0.13	0.35	0.52	0.44	0.05	-0.02																	
0.04	0.05	0.04	0.10	0.03	0.05	0.06	0.01	0.01	0.01	0.00	-0.01	0.05	0.38	0.49	0.88	0.81	0.02	0.01																	
0.05	0.04	0.11	0.07	0.05	-0.01	-0.01	-0.02	0.03	-0.01	-0.01	-0.02	0.06	0.00	0.07	0.08	0.29	0.09	0.03	0.02																
0.05	0.05	0.04	0.05	0.03	0.02	-0.01	-0.05</td																												

## I. Generación de las secuencias de ácidos nucleicos.

### Aislamiento de células B, de memoria y totales específicas de antígeno y purificación de ARN

Para el aislamiento de células B, se utiliza un método de selección negativa para eliminar todas las células no B de las PBMC utilizando el kit Dynabeads Untouched B-cell Isolation de Invitrogen (Invitrogen Núm. Cat. 1 13-5 ID) 5 siguiendo el protocolo del fabricante. La población resultante de células B no marcadas se procesa adicionalmente para aislar células B específicas de antígenos o de memoria.

Para el aislamiento de células B específicas de antígeno, se incuban células B no marcadas totales con antígeno biotinilado que se inmoviliza sobre esferas magnéticas de estreptavidina (Pierce-Thermo Scientific Núm. cat. 88816) en un rotador a temperatura ambiente durante 20 minutos. Las esferas que contienen cualquier célula B que se une 10 a un antígeno se lavan a continuación dos veces con 1xPBS. Las esferas lavadas se resuspenden en tampón de lisis RLT del kit RNeasy de Qiagen (con un suplemento de β-mercaptoetanol al 1%) para el aislamiento del ARN.

Para el aislamiento de células B de memoria, las células CD27+ e IgG+ de superficie se aíslan del total de células B 15 no marcadas utilizando kits MACS de Miltenyi para el aislamiento de células CD27+ e IgG+ de superficie (Miltenyi Biotec (Auburn, CA) Núm. Cat. 130-051-601 y 130-047-501). Para aislar simultáneamente las células B CD27+ y slgG1", se añaden anticuerpos conjugados con esferas magnéticas a ambos marcadores de superficie celular al mismo tiempo durante la etapa de incubación. Tras la purificación, las células B de memoria se centrifugan a 300xg durante 10 minutos, y a continuación se lisán en tampón RLT para ARN como se describió anteriormente para el aislamiento de ARN.

El ARN se purifica a partir de células seleccionadas utilizando el kit RNeasy de Qiagen (Qiagen Núm. Cat. 74104) 20 siguiendo el protocolo del fabricante. El tratamiento con ADNase I en columna se lleva a cabo para eliminar el ADN genómico contaminante mediante la incorporación de una etapa de digestión de ADNase I. Despues del lavado con tampón RW1, la ADNase I (Qiagen Núm. Cat. 79254) diluida en tampón RDD se aplica a la columna de purificación de ARN y se incuba durante 20 minutos a temperatura ambiente. La columna se lava una vez más con tampón RW1, seguido de dos lavados con tampón RPE, y el ARN se eluye con 30 o 50 µl de agua. La concentración del ARN se 25 determina por la absorbancia medida en un espectrofotómetro Nanodrop (Thermo Scientific) a una longitud de onda de 450 nm.

### Síntesis de ADNc y generación de amplicones mediante PCR

El ARN aislado de las células B de memoria o específicas de antígeno se somete en primer lugar a transcripción inversa como se describe en el Ejemplo 8. Los amplicones de las regiones variables de la cadena pesada, kappa y lambda para la secuenciación se generan por PCR de la siguiente manera. Para la amplificación de la cadena pesada, se ejecutan para cada muestra de ADNc cuatro reacciones independientes (cada una específica para familias de genes de V<sub>H</sub>1 y 7; V<sub>H</sub>2, 5 y 6; V<sub>H</sub>3; y V<sub>H</sub>4) utilizando los cebadores descritos en el Ejemplo 8 para preservar la distribución natural de la frecuencia del transcripto génico de V<sub>H</sub> en la reserva de células B. Para la amplificación de la cadena kappa y lambda, se ejecuta una única reacción para cada cadena para cada muestra de ADNc. Para cada reacción, se utiliza una mezcla equimolar de cebadores directos con la misma concentración de 30 cebadores inversos. La amplificación se realiza con cebadores de fusión compatibles para 454 Sequencing (Roche) mediante la plataforma Lib-L. Los cebadores inversos están diseñados para hibridar con el extremo 5' de la región 35 constante de cada cadena. Estos cebadores contienen las secuencias del cebador Lib-L B y MID de modo que las lecturas de secuenciación comienzan desde el extremo 5' de cada región constante (en sentido inverso) y en el extremo 3' de la región variable. Para las cadenas pesadas y kappa, se utiliza un único cebador inverso para cada 40 MID, mientras que para la cadena lambda, se requieren dos cebadores inversos distintos para cada MID.

La amplificación mediante PCR se realiza utilizando la polimerasa Phusion Hot Start II de Finnzyme (Thermo Scientific Núm. Cat. F-540S) donde la mezcla y las condiciones de reacción se preparan como se describe en el Ejemplo 8.

45 Para garantizar la ausencia de cualquier amplificación falsa del molde contaminante en cualquiera de los reactivos, se preparan reacciones duplicadas para cada mezcla (cuatro reacciones separadas para la cadena pesada, y una para cada cadena ligera) donde el molde de ADNc se sustituye por agua. Estas reacciones de control negativo sin molde se ejecutan al mismo tiempo que las muestras que contienen el molde. Al completar el programa de PCR, se analizan 3 µl de cada reacción (incluyendo los controles negativos) mediante electroforesis en un gel de agarosa-TAE al 1,5% para determinar la presencia de los amplicones (aproximadamente 540 pb para la cadena pesada, aproximadamente 485 pb para la cadena kappa y aproximadamente 510 pb para cadena lambda) cuando se añade el molde a la reacción, pero no en ausencia de ADNc.

50 Para preservar el emparejamiento análogo de las cadenas de anticuerpos durante la secuenciación, las células B aisladas se someten a encapsulación de células individuales utilizando encapsulación de microgotas de células individuales (Raindance Technologies, Inc., Lexington, MA). Las células B encapsuladas a continuación se fusionan con un reactivo de RT-PCR de una sola célula (el reactivo comercializado por Qiagen, como Núm. Cat. 210210) con cebadores de amplificación para generar productos de PCR de cadena pesada y ligera unidos a partir de cada célula B individual. Se utiliza PCR de solapamiento (Meijer P. J. et al., J. Mol. Biol. 358(3):764-72, 2006) para unir los

productos de PCR de cadena pesada y ligera en un solo ADN para preservar los pares de cadenas de anticuerpos mediante la secuenciación aguas abajo.

#### **Purificación, análisis, cuantificación y preparación de amplicones para la secuenciación 454**

- 5 Para eliminar el exceso de cebadores y/o dímeros de cebadores en las muestras de PCR, los amplicones se purifican utilizando esferas magnéticas Agentcourt Ampure (Beckman Coulter Núm. Cat. A63881) siguiendo el protocolo del fabricante (000387v001). Para la cadena pesada, las cuatro reacciones (VH1/7, VH2/5/6, VH3, VH4) se agrupan y se purifican como una muestra, por lo que se purifican un total de tres muestras de amplicón (cadenas pesada, kappa y lambda) para cada amplificación de ADNc. El protocolo para la purificación de Ampure se modifica al realizar las purificaciones en microtubos únicos de 1,5 ml utilizando una gradilla magnética genérica que es adecuada para tubos de 1,5 ml en lugar de en un formato de placa de 96 pocillos. Todos los volúmenes y otros procedimientos son los descritos en el protocolo. Los amplicones eluidos después de la purificación Ampure se analizan después para determinar la pureza y ausencia de cualquier especie de ADN contaminante en el Bioanalizador Agilent 2100 utilizando el chip de ADN de alta sensibilidad (Núm. Cat. 5067-4626 de Agilent Technologies) siguiendo el protocolo del fabricante.
- 10 15 Una vez que se verifica la pureza de los amplicones, la concentración del ADN se cuantifica en un fluorómetro utilizando Quant-iT PicoGreen dsDNA Assay Kit (Invitrogen Núm. Cat. P7589) como se describe en el protocolo del fabricante. El ADN de Lambda proporcionado en el kit se utiliza como un patrón de concentración con el que se generó una curva convencional de 100 ng/pocillo a 1,56 ng/pocillo. La fluorescencia de cada amplicón diluido 100 veces en tampón TE se mide por duplicado, y la concentración de ADN se determina de acuerdo con la porción 20 lineal de la curva patrón. Todas las medidas de fluorescencia se realizan en placas de color negro de 96 pocillos. Utilizando el tamaño aproximado en pares de bases de cada tipo de cadena (pesada-540 pb, kappa-485 pb y lambda-510 pb), se utiliza la siguiente fórmula para determinar la concentración:

Concentración de cada amplicón (moléculas/ $\mu$ l) =

$$[\text{conc. de la muestra (ng/ $\mu$ l)} * 6,022 \times 10^{23}] / [656,6 \times 10^9 * \text{longitud del amplicón (pb)}]$$

- 25 Cada amplicón se normalizó a  $1 \times 10^7$  moléculas/ $\mu$ l, a continuación se mezcló a una razón de Hc:Kc:Lc de 3:3:1 en volumen, se sometió a agitación vortical y finalmente se diluyó 1:10 para obtener una concentración final de la mezcla a  $1 \times 10^6$  moléculas/ $\mu$ l.

#### **Amplificación mediante PCR en emulsión, enriquecimiento de esferas, recuento de esferas y secuenciación**

- 30 La PCR en emulsión se realizó siguiendo el protocolo publicado 454: "emPCR Amplification Method Manual - Lib-L" (Edición: mayo de 2010 (Rev. abril de 2011)) con las 20 modificaciones descritas en el Ejemplo 8.

Una vez que las esferas de secuenciación se enriquecen, a partir del Etapa 3.7 del "emPCR Amplification Method Manual - Lib-L", las esferas se cuentan en el Contador de partículas Z2 de Beckman Coulter, y la concentración de esferas se calcula como:

Concentración de esferas = [Lectura media del contador de partículas \* 4] cuentas/ $\mu$ l

- 35 Las esferas enriquecidas de la PCR en emulsión se secuencian en 454 Sequencer (Roche) siguiendo el protocolo de secuenciación 454: "Sequencing Method Manual - GS Junior Titanium Series" - Mayo de 2010 (Rev. Junio 2010)

#### **II. Generación de fragmentos de péptidos:**

##### **Purificación de IgG específica de antígeno a partir de plasma de donante humano**

##### **Escrutinio para determinar reactividad al antígeno.**

- 40 La reactividad de la IgG plasmática al antígeno o antígenos de interés se somete a ensayo mediante ELISA. Brevemente, las placas de 96 pocillos de alta unión (Costar Núm. Cat.) se recubren con 100  $\mu$ l/pocillo de antígeno a 2  $\mu$ g/ml disuelto en tampón de carbonato a 37°C durante dos horas o 4°C durante la noche. Las placas se enjuagan tres veces con PBS-Tween (0,1%), a continuación se bloquean con 300  $\mu$ l/pocillo de leche en polvo al 5% sin grasa en PBS-Tween a 37°C durante 1 hora. Las muestras de plasma se diluyen 1/100, 1/500 y 1/1000 y 1/2000 en leche 45 al 5% en PBS-Tween, y se añaden 100  $\mu$ l de cada dilución en duplicados de pocillos bloqueados de la placa de 96 pocillos y se incuban durante 2 horas a 37°C. Las placas se lavan tres veces con 1x PBS-TWEEN y se añade anticuerpo anti-IgG humana conjugado con peroxidasa de rábano picante (Southern Biotech 2040-05) diluido 1/4000 en PBS-Tween a cada pocillo (100  $\mu$ l) y se incuba a 37°C por una hora. Las placas se lavan 6 veces con PBS-Tween y se desarrollan mediante la adición de 50  $\mu$ l de solución de sustrato TMB (BioFX Núm. Cat. TMBW-1000-01), seguido de 50  $\mu$ l de solución de parada (BioFX Núm. Cat. STPR1000-01). Las señales se miden a una densidad óptica de 450 nm. Se observa que los títulos séricos generalmente aumentan con el tiempo después de la vacunación.

### Purificación de IgG específica de antígeno a partir de IgG plasmática total

Las IgG totales se purifican de cada muestra de suero utilizando Proteína G como se describe en el Ejemplo 8. Las IgG purificadas se someten a diálisis dos veces contra 4 litros de 1xPBS en un casete de diálisis de corte de 10 kDa (Pierce Núm. Cat. 66456), y la concentración de IgG se determina midiendo la absorbancia a 280 nm en un 5 espectrómetro fotográfico Nanodrop (Thermo Scientific). La IgG purificada con Proteína G se purifica a continuación por afinidad utilizando esferas unidas al antígeno como se describe en el Ejemplo 8. Los anticuerpos purificados por afinidad de cada muestra se recogen para el análisis de spectrometría de masas.

### III. Espectrometría de masas

El análisis de spectrometría de masas se realiza como se describió anteriormente. Brevemente, después de la 10 digestión con una proteasa (p.ej., tripsina) y/o un reactivo de escisión química de proteína (p.ej., bromuro de cianógeno), se realiza un análisis de spectrometría de masas sobre los péptidos. Los espectros MS2 resultantes se correlacionan con los espectros teóricos MS2 derivados de la información de la base de datos de material genético, con el fin de identificar las secuencias genéticas que codifican anticuerpos que se unen específicamente al antígeno 15 de interés. Determinando las secuencias de los anticuerpos en las muestras, se determina la composición de la población de anticuerpos específicos de antígeno en el sujeto en múltiples puntos temporales después de la vacunación.

### Ejemplo 10

Este Ejemplo describe la producción de anticuerpos humanos específicos de antígeno utilizando un animal transgénico que expresa genes de anticuerpos humanos.

20 Los ratones XMG1-KL de la cepa XENOMOUSE (Amgen, Thousand Oaks, CA) tienen inactivada su maquinaria de anticuerpos endógenos de ratón y contienen loci de cadena pesada y ligera de inmunoglobulina humana (Jakovovits et al., 2007, Nature Biotechnol., 25:1134-43) Estos ratones producen anticuerpos IgGIK e IgG IX completamente humanos. Los ratones se inmunizan con un antígeno humano de interés, y se generan una base de datos de material genético y una base de datos de péptidos utilizando los siguientes métodos.

#### I. Base de datos de material genético:

##### Aislamiento de células.

Los bazos de ratones inmunizados se lavan cinco veces con 5 ml de RPMI/FCS al 10% utilizando una jeringa y una aguja 21G. Las células se congelan en FCS al 90%/DMSO al 10%. Se aíslan un total de 50 - 100 x 10<sup>6</sup> células de cada bazo.

##### Aislamiento de ARN y síntesis de ADNc.

El ARN total se aisló a partir de esplenocitos de acuerdo con el protocolo del fabricante utilizando QIAshredder (Qiagen Núm. Cat. 79654) y RNeasy mini kit (Qiagen, Hilden, Alemania, Núm. Cat. 74104). El ARN se trata con ADNasa en columna según un protocolo de secuenciación de nueva generación convencional. La concentración total de ARN se mide utilizando un espectrofotómetro ND-1000 (NanoDrop, disponible comercialmente de Thermo 35 Scientific, Wilmington, DE).

El ARN aislado se utiliza para la síntesis de ADNc de primera hebra mediante transcripción inversa 30 utilizando el sistema Thermoscript RT-PCR (Invitrogen (parte de Life Technologies), Carlsbad, CA Núm. Cat. I 1146-024). El ADNc se sintetiza utilizando 1,5 µg de ARN y cebador oligo dT de acuerdo con el protocolo del fabricante.

##### Amplificación de VH y VL.

40 Los amplicones de las regiones variables de la cadena pesada, kappa y lambda para secuenciación se generan mediante PCR de la siguiente manera utilizando cebadores específicos para secuencias de anticuerpos humanos como se describe en el Ejemplo 8. Para la amplificación de cadena pesada, cuatro reacciones independientes (cada una específica para familias de genes de V<sub>H</sub>1 y 7; V<sub>H</sub>2, 5 y 6; V<sub>H</sub>3; y V<sub>H</sub>4) se ejecutan para cada muestra de ADNc para preservar la distribución natural de la frecuencia del transcripto génico de VH en la reserva de células B. Para la 45 amplificación de la cadena kappa y lambda, se ejecuta una única reacción para cada cadena para cada muestra de ADNc. Para cada reacción, se utiliza una mezcla equimolar de cebadores directos con la misma concentración de cebadores inversos. La amplificación se realiza con cebadores de fusión compatibles para 454 Sequencing (Roche) mediante la plataforma Lib-L. Los cebadores inversos están diseñados para hibridar con el extremo 5' de la región constante de cada cadena. Estos cebadores contienen las secuencias del cebador Lib-L B y MID de modo que las 50 lecturas de secuenciación comienzan desde el extremo 5' de cada región constante (en sentido inverso) y en el extremo 3' de la región variable. Para las cadenas pesada y kappa, se utiliza un único cebador inverso para cada MID, mientras que para la cadena lambda, se requieren dos cebadores inversos distintos para cada MID.

La amplificación mediante PCR se realiza utilizando la polimerasa Phusion Hot Start II de Finnzyme (Thermo Scientific Núm. Cat. F-540S) donde la mezcla y las condiciones de reacción se configuran como se describe en el

**Ejemplo 8.**

Para garantizar la ausencia de cualquier amplificación falsa del molde contaminante en cualquiera de los reactivos, se preparan reacciones duplicadas para cada mezcla (cuatro reacciones separadas para cadena pesada, y una para cada cadena ligera) donde el molde de ADNc se sustituye por agua. Estas reacciones de control negativo sin molde

- 5 se ejecutan al mismo tiempo que las muestras que contienen el molde. Al completar el programa de PCR, se analizan 3 µl de cada reacción (incluyendo los controles negativos) mediante electroforesis en un gel de agarosa TAE al 1,5% para la presencia de los amplicones (aproximadamente 540 pb para la cadena pesada, aproximadamente 485 pb para la cadena kappa y aproximadamente 510 pb para cadena lambda) cuando se añade el molde a la reacción, pero no en ausencia de ADNc. Los productos de PCR se purifican de acuerdo con el protocolo del fabricante utilizando AMPure XP (Agencourt; Beckman Coulter Genomics, Brea, CA, Núm. cat. A63881) y se analizan utilizando un Agilent 2100 Bio Analyzer.

10 Las secuencias de los productos de PCR se traducen a continuación a las secuencias de aminoácidos pronosticadas, que a continuación se digieren teóricamente (p.ej., con una proteasa y/o un reactivo de escisión química de proteína) para producir fragmentos peptídicos virtuales. Estos fragmentos peptídicos virtuales se utilizan 15 a continuación para generar los espectros de masas pronosticados.

**Generación de espectros de masas reales a partir de fragmentos de péptidos de anticuerpos policlonales:**

Los anticuerpos policlonales se purifican a partir de los sueros y/o plasma de los ratones. Para purificar los anticuerpos, se utilizan los siguientes métodos:

**Purificación con proteína-G:**

- 20 Se añade 1 ml de esferas magnéticas de proteína G (Millipore (Billerica, MA), Núm. Cat. LSKMAGG10) a cada uno de cuatro tubos cónicos de 15 ml (Falcon (BD Biosciences, Franklin Lake, NJ), Núm. Cat. 352097). Las esferas de cada tubo se lavan dos veces con 10 ml de solución salina tamponada con fosfato de pH 7,4, Tween-20 al 0,05% (PBST) y tres veces con 10 ml de PBS. Los sueros de tres ratones (ID 1262-2, 1262-4, 1263-4) se agrupan y se diluyen diez veces hasta un volumen final de 20 de 6 ml en PBS. A continuación, se añaden 1,5 ml del suero diluido 25 combinado a cada tubo de esferas y se incuba durante la noche a 4°C. El flujo continuo se recoge y pasa por el procedimiento de purificación otras dos veces. Después de recoger el flujo continuo, cada tubo se lava dos veces con 10 ml de PBST y tres veces con 10 ml de PBS. A continuación, cada tubo se incuba a 4°C durante 30 minutos con 0,5 ml de glicina 0,1 M de pH 2,7 para eluir la IgG. La elución es 25 repetida cinco veces. Todos los eluatos se neutralizan con Tris 1M de pH 8,5, se someten a diálsis durante la noche frente a PBS y se mide la concentración 30 de proteína con un espectrofotómetro ND-1000 (Nanodrop). En total, se purifican aproximadamente 2,5 mg de IgG.

**Preparación de la columna de antígeno:**

Se lavan tres veces 5,0 mL de esferas magnéticas de estreptavidina (SA) de nueva aportación (Pierce, Núm. Cat. 88817) con 10 ml de PBS, y se incuban durante la noche a 4°C con 105 µl de una provisión de partida de 20 mg/ml 35 del antígeno de interés conjugado con biotina diluida en 5,0 mL de PBS. El flujo se desecha, y las esferas se lavan tres veces con 10 ml de PBS y se dividen en alícuotas en diez tubos de 1,7 ml de baja unión (Axygen (Union City, CA), Núm. Cat. MCT-175-L-C). Las esferas en alícuotas se colocan en una gradilla magnética (Invitrogen, DynaMag), y el PBS se elimina antes de agregar los sueros diluidos.

**Purificación específica del antígeno:**

- 40 La IgG purificada con proteína G anterior se añade a las esferas magnéticas de SA acopladas con antígeno biotinilado. Después de una noche de incubación a 4°C, se recoge el flujo continuo y las esferas se lavan con un total de 10 ml de cada uno de los siguientes tampones, en serie:

**PBS**

Tampón RIPA (es decir, tampón de ensayo de radioinmunoprecipitación; Alcaraz et al.,/. Vet. Diagn. Invest. 2(3):191-196, 1990; Ngoka, L.C., Proteome Sci. 6 (1): 30, 2008)

- 45 Acetonitrilo al 20% en PBS

Etilenglicol al 60% en PBS

NaCl 0,5M en PBS

PBS (es decir, solución salina tamponada con fosfato)

- 50 A continuación la IgG se eluye con 5 fracciones de 1,5 ml de glicina 0,1 M de pH 3,5, a continuación 5 fracciones de 1,5 ml de glicina 0,1 M de pH 2,7, a continuación 5 fracciones de 1,5 ml de glicina 0,1 M de pH 1,8 y se neutraliza con Tris 1 M de pH 8,5. Los eluatos se someten a ensayo para determinar la reactividad al antígeno de interés utilizando placas de 96 pocillos recubiertas con el antígeno. Las fracciones con actividad se cuantifican mediante

ELISA (Thermo, Núm. cat. 23300) y se analizan para determinar la reactividad del antígeno mediante transferencia Western. Las fracciones con la reactividad más limpia se analizan mediante espectrometría de masas.

#### **Espectrometría de masas**

El análisis de espectrometría de masas se realiza como se describió anteriormente. Brevemente, después de la digestión con una proteasa (p.ej., tripsina) y/o un reactivo de escisión química de proteína (p.ej., bromuro de cianógeno), se realiza un análisis de espectrometría de masas sobre los péptidos. Los espectros de MS2 resultantes se correlacionan con los espectros teóricos MS2 derivados de la información en la base de datos de material genético, con el fin de identificar las secuencias genéticas que codifican anticuerpos que se unen específicamente al antígeno de interés.

#### **10 Expresión e identificación de anticuerpos monoclonales**

Los clones de la región variable de la cadena pesada (gamma) diferenciados, los clones de la región variable de la cadena kappa y los clones de la región variable de la cadena lambda se expresan en un formato combinatorio y se escrutan para determinar la actividad de unión específica del antígeno. Cada cadena gamma se empareja con cada cadena ligera (kappa y lambda) para expresar anticuerpos mediante transfección transitoria de células HEK293E en placas de cultivo de tejidos de 96 pocillos convencionales.

El anticuerpo que se secreta a partir de las células transfectadas en cada pocillo se escruta para determinar la unión al antígeno recombinante purificado mediante el ensayo inmunoabsorbente ligado a enzimas (ELISA), como se describió anteriormente. Varios emparejamientos de las cadenas pesada y ligera dan como resultado anticuerpos que se unen específicamente a las placas recubiertas de antígeno. Estos pares de cadena pesada y ligera se seleccionan, dando como resultado la producción de anticuerpos completamente humanos que se unen específicamente al antígeno humano de interés.

#### **Equivalentes**

Los expertos en la técnica reconocerán, o podrán determinar, utilizando no más que la experimentación de rutina, numerosos equivalentes a las realizaciones específicas descritas específicamente en la presente memoria. Se pretende que dichos equivalentes estén incluidos en el alcance de las siguientes reivindicaciones.

#### **ENUNCIADOS DE LA DESCRIPCIÓN**

A. Un método para obtener secuencias de nucleótidos o secuencias de aminoácidos de cadenas pesadas o ligeras de inmunoglobulinas que se unen específicamente a un antígeno, que comprende: (a) proporcionar secuencias de ácido nucleico que codifican cadenas de inmunoglobulina o regiones variables de las mismas de al menos un animal; (b) obtener una población de inmunoglobulinas policlonales que se unen específicamente a un antígeno, y obtener información de los espectros de masas de fragmentos peptídicos de cadenas pesadas o cadenas ligeras o regiones variables de las mismas de dicha población; (c) correlacionar la información de espectros de masas de los fragmentos peptídicos con información pronosticada de los espectros de masas de las secuencias de ácidos nucleicos, dicha información pronosticada de los espectros de masas derivada de secuencias de aminoácidos pronosticadas codificadas por dichas secuencias de ácidos nucleicos, para identificar secuencias de nucleótidos o secuencias de aminoácidos de cadenas de inmunoglobulina (o regiones variables de las mismas) que comprenden los fragmentos peptídicos; y (d) seleccionar entre las secuencias de nucleótidos identificadas o secuencias de aminoácidos de cadenas de inmunoglobulina (o regiones variables de las mismas) basándose en la cobertura de la secuencia de aminoácidos de las cadenas de inmunoglobulina o fragmentos de las mismas por los fragmentos peptídicos, para obtener secuencias de nucleótidos o secuencias de aminoácidos de cadenas pesadas o ligeras de inmunoglobulinas que se unen específicamente a un antígeno.

B. El método del enunciado A, en donde dicho al menos un animal de la etapa (a) es un animal expuesto al antígeno.

C. El método del enunciado A, en donde las secuencias de ácidos nucleicos son secuencias de ácidos nucleicos expresadas.

D. El método del enunciado A, en donde las secuencias de ácido nucleico que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) se obtienen de dicho al menos un animal: (1) aislando moléculas de ácido nucleico de glóbulos blancos de dicho al menos un animal; y (2) amplificando la cadena de inmunoglobulina (o región variable de la misma) que codifica las moléculas de ácido nucleico utilizando cebadores específicos para secuencias de polinucleótidos adyacentes a dicha cadena de inmunoglobulina (o región variable de la misma que codifica las moléculas de ácido nucleico, y (3) obteniendo secuencias de nucleótidos de dichas moléculas de ácido nucleico amplificadas que codifican cadenas de inmunoglobulina (o regiones variables de las mismas).

E. El método del enunciado D, en donde las moléculas de ácido nucleico que codifican los polipéptidos son moléculas de ARN y dicha etapa de amplificación incluye una etapa de transcripción inversa inicial.

- F. El método del enunciado D, en donde dichas secuencias de polinucleótidos adyacentes a dicha cadena de inmunoglobulina (o región variable de la misma que codifica las moléculas de ácido nucleico se seleccionan del grupo que consiste en genes de inmunoglobulina que flanquean ADN genómico, secuencias de polinucleótidos que codifican regiones constantes de cadena de inmunoglobulina, y secuencias de polinucleótidos que codifican para la región del marco de la cadena de inmunoglobulina.
- 5 G. El método del enunciado A, en donde la información del espectro de masas pronosticado se obtiene utilizando un método que comprende las etapas de: (i) realizar un resumen teórico de secuencias de aminoácidos pronosticadas codificadas por secuencias de nucleótidos de las secuencias de ácido nucleico con una o más proteasas y/o uno o más reactivos de escisión química de proteínas para generar fragmentos de péptidos virtuales; y (ii) crear espectros de masas pronosticados de dichos fragmentos peptídicos virtuales.
- 10 H. El método del enunciado A, en donde las secuencias de ácido nucleico de la etapa (a), las secuencias de aminoácidos pronosticadas y los espectros de masas pronosticados derivados de dichas secuencias de ácido nucleico se ubican dentro de una base de datos de material genético.
- 15 I. El método del enunciado A, en donde dicha población policlonal de inmunoglobulinas de la etapa (b) se obtiene a partir de una muestra de fluido corporal o fracción de la misma de un animal.
- J. El método del enunciado I, en donde dicho fluido corporal se selecciona del grupo que consiste en sangre, fluido cerebroespinal, fluido sinovial, fluido peritoneal, secreciones de la mucosa, lágrimas, secreciones nasales, saliva, leche y secreciones genitourinarias.
- K. El método del enunciado I, en donde el animal es un animal previamente expuesto al antígeno.
- 20 L. El método del enunciado K, en donde el animal previamente expuesto al antígeno es un animal previamente inmunizado con el antígeno.
- M. El método del enunciado I, en donde el animal es el mismo que dicho al menos un animal en la etapa (a).
- N. El método del enunciado I, en donde dicha población policlonal de inmunoglobulinas de la etapa (b) se purifica a partir de dicha muestra de fluido corporal o fracciones de la misma.
- 25 O. El método del enunciado N, en donde la purificación se consigue mediante purificación de proteína A o proteína G, o purificación por afinidad basada en el antígeno.
- P. El método del enunciado A, en donde dicha población policlonal de inmunoglobulinas de la etapa (b) se obtiene a partir del medio de células cultivadas in vitro.
- 30 Q. El método del enunciado A, en donde la información del espectro de masas de los fragmentos peptídicos de la etapa (b) se obtiene de la población policlonal de inmunoglobulinas dirigiendo la población con una o más proteasas y/o uno o más reactivos de escisión química de proteínas para generar fragmentos, y someter los fragmentos generados a análisis de espectrometría de masas.
- R. El método del enunciado A, en donde la selección en la etapa (d) se basa en la cobertura de la secuencia de aminoácidos de las secuencias de CDR3 de las inmunoglobulinas por los fragmentos peptídicos.
- 35 S. El método del enunciado A, en donde la selección en la etapa (d) se basa en la cobertura de la secuencia de aminoácidos de las regiones variables de las inmunoglobulinas mediante los fragmentos peptídicos.
- T. El método del enunciado A, en donde la selección en la etapa (d) se basa en la cobertura de la secuencia de aminoácidos de las secuencias de CDR3 y la cobertura de la secuencia de aminoácidos de las regiones variables de las inmunoglobulinas por los fragmentos peptídicos.
- 40 U. El método del enunciado A, en donde la selección en la etapa (d) se basa adicionalmente en al menos un parámetro seleccionado del grupo consistente en el número de péptidos únicos mapeados, la compartición de espectro, el recuento de péptidos totales, el recuento de péptidos únicos, la frecuencia de las secuencias codificantes de ácidos nucleicos, y la relación clonal.
- V. Un método para crear una inmunoglobulina que se une específicamente a un antígeno, que comprende: obtener secuencias de nucleótidos o secuencias de aminoácidos de una cadena pesada y una cadena ligera utilizando el método del enunciado A, haciendo que dicha cadena pesada y dicha cadena ligera se basen en las secuencias de nucleótidos o secuencias de aminoácidos obtenidas mediante técnicas recombinantes de biología molecular o técnicas de síntesis de genes, y el ensamblaje de dicha cadena pesada con dicha cadena ligera para crear una inmunoglobulina que se une específicamente al antígeno.
- 45 W. El método del enunciado V, que comprende adicionalmente evaluar la inmunoglobulina en un inmunoensayo para confirmar que dicha inmunoglobulina se une específicamente al antígeno.
- 50

- 5 X. El método del enunciado W, en donde el inmunoensayo se selecciona del grupo que consiste en un ensayo de citometría de flujo, un ensayo de inmunoabsorción ligado a enzimas (ELISA), un ensayo de transferencia Western, un ensayo de inmunohistoquímica, un ensayo de inmunofluorescencia, un radioinmunoensayo, un ensayo de neutralización, un ensayo de unión, un ensayo de afinidad, un ensayo de inmunoprecipitación de proteínas y un ensayo de inmunoprecipitación de péptidos.
- 10 Y. Un método para obtener secuencias de nucleótidos o secuencias de aminoácidos de una región variable de cadena de inmunoglobulina de una inmunoglobulina que se une específicamente a un antígeno que comprende: (a) proporcionar secuencias de ácido nucleico que codifican regiones variables de inmunoglobulina de múltiples inmunoglobulinas de al menos un animal; (b) obtener información de los espectros de masas de los fragmentos peptídicos de las regiones variables de la cadena de inmunoglobulina de una población de inmunoglobulinas policlonales que se unen específicamente a un antígeno; (c) correlacionar información de espectros de masas de los fragmentos peptídicos con información de espectros de masas pronosticados de las secuencias de ácidos nucleicos, dicha información pronosticada de los espectros de masas derivada de secuencias de aminoácidos pronosticadas codificadas por dichas secuencias de ácidos nucleicos para identificar secuencias de nucleótidos o secuencias de aminoácidos de regiones variables de cadenas de inmunoglobulinas que comprenden los fragmentos peptídicos; y (d) seleccionar entre las secuencias de nucleótidos o secuencias de aminoácidos identificadas de las regiones variables de cadena de inmunoglobulina basadas en la cobertura de secuencias de aminoácidos de las regiones variables por los fragmentos peptídicos, para obtener secuencias de nucleótidos o secuencias de aminoácidos de regiones variables de inmunoglobulinas que se unen específicamente a un antígeno.
- 15 20 Z. El método del enunciado Y, que comprende adicionalmente (e) preparar las regiones variables de la cadena de inmunoglobulina basadas en las secuencias de nucleótidos o secuencias de aminoácidos obtenidas en la etapa (d) mediante técnicas recombinantes de biología molecular o técnicas de síntesis de genes.
- 25 AA. El método del enunciado Z, que comprende adicionalmente (f) escrutar las regiones variables de la cadena de inmunoglobulina con un inmunoensayo para aislar una región variable de la cadena de inmunoglobulina de una inmunoglobulina que se une específicamente al antígeno.
- AB. El método del enunciado Y, en donde la cadena de inmunoglobulina es una cadena pesada.
- AC. El método del enunciado Y, en donde la cadena de inmunoglobulina es una cadena ligera.
- 30 AD. Un método para crear un dominio de unión a antígeno de una inmunoglobulina que se une específicamente a un antígeno que comprende: (a) obtener secuencias de nucleótidos o secuencias de aminoácidos de las regiones variables de la cadena de inmunoglobulina de inmunoglobulinas que se unen específicamente a un antígeno de acuerdo con un método del enunciado Y; y (b) ensamblar una región variable de cadena pesada de inmunoglobulina con una región variable de cadena ligera de inmunoglobulina para crear un dominio de unión a antígeno de una inmunoglobulina que se une específicamente al antígeno.
- AE. El método del enunciado A, V, Y o AD, en donde el animal es un ser humano.
- 35 AF. El método del enunciado A, V, Y o AD, en donde el animal es un conejo o un ratón.
- AG. Una inmunoglobulina que se une específicamente a un antígeno creado de acuerdo con el método del enunciado V.
- AH. Una región variable de cadena de inmunoglobulina de una inmunoglobulina que se une específicamente a un antígeno aislado de acuerdo con el método del enunciado Y.
- 40 AI. Un dominio de unión a antígeno de una inmunoglobulina que se une específicamente a un antígeno creado de acuerdo con el método del enunciado AD.
- AJ. Una composición que comprende la inmunoglobulina del enunciado AG y un vehículo farmacéuticamente aceptable.
- AK. Una composición que comprende la región variable de la cadena de inmunoglobulina del enunciado AH y un vehículo farmacéuticamente aceptable.
- 45 AL. Una composición que comprende el dominio de unión a antígeno del enunciado AI y un vehículo farmacéuticamente aceptable.
- AM. Un método para tratar un animal que tiene o se sospecha que tiene una enfermedad caracterizada por un antígeno de enfermedad, en donde el método comprende administrar una cantidad eficaz de una composición del enunciado 36, 37 o 38, en donde el antígeno y el antígeno de enfermedad son el mismo.
- AN. El método del enunciado AM, en donde el animal es un ser humano.
- AO. Un método para reducir la probabilidad de aparición en un animal de una enfermedad caracterizada por la

presencia en el animal de un antígeno de enfermedad, en donde el método comprende administrar una cantidad eficaz de una composición del enunciado AJ, AK o AL, en donde el antígeno y el antígeno de enfermedad son el mismo.

AP. El método del enunciado AO, en donde la composición comprende además un coadyuvante.

5 AQ. El método del enunciado AO, en donde el animal es un ser humano.

AR. Un método para controlar una población de anticuerpo específico de antígeno de un sujeto expuesto a un antígeno que comprende: (a) obtener secuencias de ácido nucleico que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) de inmunoglobulinas múltiples de un sujeto expuesto a un antígeno; (b) obtener información de los espectros de masas de fragmentos de péptidos de cadenas pesadas de inmunoglobulina y cadenas ligeras de inmunoglobulina de una población de inmunoglobulinas policlonales del sujeto que se unen específicamente al antígeno; (c) correlacionar información de espectros de masas de los fragmentos peptídicos con información de espectros de masas pronosticados de las secuencias de ácidos nucleicos, dicha información pronosticada de los espectros de masas derivada de secuencias de aminoácidos pronosticadas codificadas por dichas secuencias de ácidos nucleicos, para identificar secuencias de la población de inmunoglobulinas policlonales del sujeto que se unen específicamente al antígeno; y (d) seleccionar entre las secuencias de nucleótidos identificadas o secuencias de aminoácidos de cadenas de inmunoglobulina (o regiones variables de las mismas) basándose en la cobertura de la secuencia de aminoácidos de las cadenas de inmunoglobulina o fragmentos de las mismas por los fragmentos peptídicos, para obtener secuencias de nucleótidos o secuencias de aminoácidos de cadenas pesadas o ligeras de inmunoglobulinas que se unen específicamente al antígeno.

20 AS. El método del enunciado AR, en donde las secuencias de ácidos nucleicos y la población de inmunoglobulinas policlonales corresponden cada una a un punto temporal después de la exposición del sujeto al antígeno.

AT. El método del enunciado AR, en donde las secuencias de ácidos nucleicos y la población de inmunoglobulinas policlonales corresponden cada una a múltiples puntos temporales después de la exposición del sujeto al antígeno.

25 AU. El uso de la inmunoglobulina del enunciado AG para la fabricación de un medicamento útil para tratar o reducir la probabilidad de enfermedad en un animal caracterizado por un antígeno de enfermedad, en donde el antígeno y el antígeno de enfermedad son el mismo.

AV. El uso de la región variable de cadena de inmunoglobulina del enunciado AH para la fabricación de un medicamento útil para tratar o reducir la probabilidad de enfermedad en un animal caracterizado por un antígeno de enfermedad, en donde el antígeno y el antígeno de enfermedad son el mismo.

30 AW. El uso del dominio de unión a antígeno del enunciado AI para la fabricación de un medicamento útil para tratar o reducir la probabilidad de enfermedad en un animal caracterizado por un antígeno de enfermedad, en donde el antígeno y el antígeno de enfermedad son el mismo.

AX. El uso de acuerdo con cualquiera de las declaraciones AU-AW, en donde el animal es un ser humano.

## LISTA DE SECUENCIAS

5      <110> Polakiewicz, Roberto  
       Cheung, Wan Cheung  
 Rush, II, John Edward  
 Beausoleil, Sean Andre  
 10     <120> Métodos y reactivos para crear anticuerpos monoclonales  
 <130> 28446  
 <140> Para ser asignada  
 <141> 09-03-2012  
 15     <150> 61/450,922  
 <151> 09-03-2011  
 <150> 61/560,006  
 <151> 15-11-2011  
 20     <150> 61/566,876  
 <151> 05-12-2011  
 <150> 61/594,729  
 25     <151> 03-02-2012  
 <160> 412  
 <170> PatentIn versión 3.5  
 30     <210> 1  
 <211> 60  
 <212> DNA  
 <213> Secuencia Artificial  
 35     <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 1  
 40     ccatctcatc cctgcgtgtc tccgactctag acgagtgcgt gatgtgaagc ttcaggagtc        60  
 <210> 2  
 <211> 60  
 <212> DNA  
 45     <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 2  
 50     ccatctcatc cctgcgtgtc tccgactctag acgctcgaca caggtgcagc tgaaggagtc        60  
 <210> 3  
 <211> 60  
 55     <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 60     <400> 3  
 ccatctcatc cctgcgtgtc tccgactctag agacgcactc caggtgcagc tgaaggagtc        60  
 <210> 4  
 65     <211> 60

ES 2 666 301 T3

<212> DNA  
 <213> Secuencia Artificial  
 5      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 4  
 ccatctcatc cctgcgtgtc tccgactcaag agcactgttag cagttactct gaaaagagtcc 60  
 10     <210> 5  
 <211> 61  
 <212> DNA  
 <213> Secuencia Artificial  
 15     <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 5  
 ccatctcatc cctgcgtgtc tccgactcaag atcagacacg gaggtccagc tgcaacaatc 60  
 20     t 61  
 <210> 6  
 <211> 60  
 <212> DNA  
 <213> Secuencia Artificial  
 25     <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 6  
 ccatctcatc cctgcgtgtc tccgactcaag atatcgcgag gaggtccagc tgcaaggc 60  
 30     <210> 7  
 <211> 61  
 <212> DNA  
 <213> Secuencia Artificial  
 35     <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 7  
 ccatctcatc cctgcgtgtc tccgactcaag cgtgtctcta caggtccaaac tgcaaggcc 60  
 40     t 61  
 <210> 8  
 <211> 60  
 <212> DNA  
 <213> Secuencia Artificial  
 45     <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 8  
 ccatctcatc cctgcgtgtc tccgactcaag ctgcgtgtc gaggtgaagc tggtgagtc 60  
 50     <210> 9  
 <211> 60  
 <212> DNA  
 <213> Secuencia Artificial  
 55     <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 9  
 ccatctcatc cctgcgtgtc tccgactcaag tctctatggc gaggtgaagc tggtgaaatc 60

5       <210> 10  
 <211> 60  
 <212> DNA  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
 10      <400> 10  
 ccatctcatc cctgcgtgtc tccgacttcag tgatacgatct gatgtgaact tgaaagtgtc       60  
  
 <210> 11  
 <211> 51  
 15      <212> DNA  
 <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
 20      <400> 11  
 cctatccccct gtgtgccttg gcagtctcag tgccagagaca gtgaccagag t       51  
  
 <210> 12  
 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial  
  
 <220>  
 30      <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
 <400> 12  
 cctatccccct gtgtgccttg gcagtctcag tgaggagact gtgagagtgg t       51  
  
 35      <210> 13  
 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial  
  
 40      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
 <400> 13  
 cctatccccct gtgtgccttg gcagtctcag tgaggagacg gtgactgagg t       51  
  
 45      <210> 14  
 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial  
  
 50      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
 <400> 14  
 55      cctatccccct gtgtgccttg gcagtctcag tgaggagacg gtgaccgtgg t       51  
  
 <210> 15  
 <211> 61  
 <212> DNA  
 60      <213> Secuencia Artificial  
  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
 65      <400> 15

# ES 2 666 301 T3

	ccatctcatc cctgcgtgta tccgactca gatacgatgt gatgtttga tgacccaaac	60
	t	61
	<210> 16	
	<211> 61	
5	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
10	<400> 16	
	ccatctcatc cctgcgtgta tccgactca gcaagatatac gatattgtga tgacgcaggc	60
	t	61
	<210> 17	
15	<211> 58	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
20	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
	<400> 17	
	ccatctcatc cctgcgtgta tccgactca atacgacgtt gatattgtga taacccag	58
25	<210> 18	
	<211> 61	
	<212> DNA	
	<213> Secuencia Artificial	
30	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
	<400> 18	
	ccatctcatc cctgcgtgta tccgactca tcacgtacta gacattgtgc tgacccaaatc	60
	t	61
35	<210> 19	
	<211> 61	
	<212> DNA	
	<213> Secuencia Artificial	
40	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
	<400> 19	
	ccatctcatc cctgcgtgta tccgactca cgtcttagtac gacattgtga tgacccagtc	60
45	t	61
	<210> 20	
	<211> 61	
	<212> DNA	
50	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
55	<400> 20	
	ccatctcatc cctgcgtgta tccgactca tctacgttagc gatattgtgc taactcagtc	60
	t	61
	<210> 21	

<211> 61		
<212> DNA		
<213> Secuencia Artificial		
 5 <220>		
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético		
 10 <400> 21	cctatctcatc cctgcgtgtc tccgactca g tgtactactc gatatccaga tgacacagac	60
	t	61
 15 <210> 22		
<211> 61		
<212> DNA		
<213> Secuencia Artificial		
 20 <220>		
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético		
 25 <400> 22	cctatctcatc cctgcgtgtc tccgactca g acgactacag gacatccago tgactcagtc	60
	t	61
 30 <210> 23		
<211> 61		
<212> DNA		
<213> Secuencia Artificial		
 35 <220>		
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético		
 40 <400> 23	cctatctcatc cctgcgtgtc tccgactca g ctagactca g caaattgttc tcacccagtc	60
	t	61
 45 <210> 24		
<211> 49		
<212> DNA		
<213> Secuencia Artificial		
 50 <220>		
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético		
 55 <400> 24	cctatcccct gtgtgccttg gcagtctca g ccgtttcagc tccagcttg	49
 55 <210> 25		
<211> 50		
<212> DNA		
<213> Secuencia Artificial		
 60 <220>		
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético		

<400> 26  
 cctatcccct gtgtgccttg gcagtcctcg ccgttttatt tccaaacttg 50

5 <210> 27  
 <211> 61  
 <212> DNA  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 27  
 ccatctcatac cctgcgtgtc tccgactcgatc tacgagtatgcaggctgttg tgactcagga 60

a 61

15 <210> 28  
 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 28  
 cctatcccct gtgtgccttg gcagtcctcg ctggggctga cctaggacatg 51

25 <210> 29  
 <211> 4  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 29  
 Gly Val Lys Phe  
 35 1

<210> 30  
 <211> 6  
 <212> PRT

40 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45 <400> 30  
 Gly Val Ser Thr Asn Val  
 1 5

<210> 31  
 <211> 15  
 <212> PRT

50 <213> Secuencia Artificial

<220>  
 55 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 31  
 Asp Pro Tyr Asp Asp Pro Thr Tyr Arg Gly Tyr Gly Met Asp Leu  
 1 5 10 15

60 <210> 32  
 <211> 9  
 <212> PRT

<213> Secuencia Artificial

<220>

5 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 32  
**Asn Pro Ala Val Asn Thr Tyr Ala Ser**  
 1               5

10 <210> 33  
 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial

<220>

15 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 33  
**His Leu Phe Leu His Phe**  
 1               5

20 <210> 34  
 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial

25 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 34  
**His Leu Phe Leu Asn Leu**  
 1               5

30 <210> 35  
 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 35  
**His Leu Phe Leu Asn Phe**  
 1               5

40 <210> 36  
 <211> 17  
 <212> PRT  
 <213> Secuencia Artificial

<220>

45 <223> Descripción de la Secuencia Artificial: Péptido sintético

50 <400> 36  
**Gly Leu Gly Tyr Val Gly Ser Ser Val Tyr Ile Val Lys Tyr Ile Asn**  
 1               5               10               15

Leu

<210> 37  
 <211> 17  
 55 <212> PRT  
 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 37  
 Asp Leu Ile Arg Val Ala Gly Asp Thr Phe Tyr Asp Gly Ala Phe Asn  
 1               5                           10                           15

Leu

5   <210> 38  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

10   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 38  
 Gly Arg Tyr Asn Gly Trp Gly Tyr Ser Asn Asp Leu  
 1               5                           10

15   <210> 39  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

20   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 39  
 Gly Gly Gly Thr Thr Leu Tyr Thr Tyr Phe Asp Leu  
 1               5                           10

25   <210> 40  
 <211> 17  
 <212> PRT  
 30   <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

35   <400> 40  
 Gly Leu Gly Tyr Val Gly Ser Asp Val Tyr Ile Val Lys Tyr Ile Asn  
 1               5                           10                           15

Leu

40   <210> 41  
 <211> 14  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45   <400> 41  
 Gly Gly Tyr Gly Tyr Gly Asn Thr Asp Phe Asn Leu  
 1               5                           10

50   <210> 42  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 42  
 Asp Asp Gly Gly Val Arg Val Asp Phe Asp Leu  
 1               5                   10

5   <210> 43  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial

10   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 43  
 Val Asp Asp Ser Gly Trp Met Pro Phe Lys Leu  
 1               5                   10

15   <210> 44  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

20   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 44  
 Asn Val Gly Ser Ser Ser His Tyr Asn Leu Asn Leu  
 1               5                   10

25   <210> 45  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial

30   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 45  
 Asp Gly Thr Asp His Gly Phe Asn Ile Asp Leu  
 1               5                   10

35   <210> 46  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

40   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 46  
 Ser Thr Phe Arg Asn Ser Tyr Ala Arg Leu Ala Leu  
 1               5                   10

45   <210> 47  
 <211> 17  
 50   <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

55   <400> 47  
 Ile Pro Tyr Gly Trp Tyr Ser Gly Gly Ala Ala Pro Tyr Phe Asp  
 1               5                   10                   15

Leu

5           <210> 48  
 <211> 5  
 <212> PRT  
 <213> Secuencia Artificial  
 10          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 48  
 Asn Ala Ala Ile Leu  
 1               5  
 15          <210> 49  
 <211> 14  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 20          <400> 49  
 Ala Val Ser Asp Asn Gly Tyr Gly Met Tyr Trp Phe Asn Leu  
 1               5               10  
 <210> 50  
 <211> 11  
 25          <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 30          <400> 50  
 Glu Leu Ala Gly Tyr Asp Val Gly Val Glu Phe  
 1               5               10  
 <210> 51  
 <211> 10  
 35          <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 40          <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 51  
 Ser Arg Ser Thr Ser Tyr Tyr Ile Asn Leu  
 1               5               10  
 45          <210> 52  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial  
 50          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 52  
 Asp Gly Ser Asp His Gly Phe Asn Ile Asp Leu  
 1               5               10  
 55          <210> 53  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 60          <220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

5           <400> 53  
       Gly Ala Asp Ser Ile Tyr Arg Ile Tyr Phe Asp Leu  
       1                       5                       10

10          <210> 54  
       <211> 12  
       <212> PRT  
       <213> Secuencia Artificial

15          <220>  
       <223> Descripción de la Secuencia Artificial: Péptido sintético

20          <400> 54  
       Asn Val Gly Ser Ser Ser Tyr Tyr Asn Leu Asn Leu  
       1                       5                       10

25          <210> 55  
       <211> 14  
       <212> PRT  
       <213> Secuencia Artificial

30          <220>  
       <223> Descripción de la Secuencia Artificial: Péptido sintético

35          <400> 55  
       Gly Gly Asp Ala Gly Tyr Gly Tyr Phe Asp Ala Phe Gly Pro  
       1                       5                       10

40          <210> 56  
       <211> 14  
       <212> PRT  
       <213> Secuencia Artificial

45          <220>  
       <223> Descripción de la Secuencia Artificial: Péptido sintético

50          <400> 56  
       Gly Gly Asp Ala Gly Tyr Gly Ser Phe Asp Ala Phe Gly Pro  
       1                       5                       10

55          <210> 57  
       <211> 17  
       <212> PRT  
       <213> Secuencia Artificial

60          <220>  
       <223> Descripción de la Secuencia Artificial: Péptido sintético

65          <400> 57  
       Gly Leu Gly Tyr Val Gly Ser Ser Val Tyr Ile Ser Lys Tyr Ile Asn  
       1                       5                       10                       15

70           Leu

75          <210> 58  
       <211> 17  
       <212> PRT  
       <213> Secuencia Artificial

80          <220>  
       <223> Descripción de la Secuencia Artificial: Péptido sintético

85          <400> 58

ES 2 666 301 T3

Val Pro Trp Thr Gly Gly Ser Gly Asp Ala Arg Leu Thr Arg Leu Asp  
1 5 10 15

Leu

<210> 59

<211> 19

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

10

<400> 59

Asp Leu Gly Tyr Ala Ser Tyr Ile Gly Tyr Gly Tyr Pro Ser Tyr Tyr  
1 5 10 15

Phe Lys Leu

<210> 60

<211> 19

15 <212> PRT

<213> Secuencia Artificial

<220>

20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 60

Asp Leu Gly Tyr Ala Ser Tyr Arg Gly Tyr Gly Tyr Pro Ser Tyr Tyr  
1 5 10 15

Phe Lys Leu

25 <210> 61

<211> 13

<212> PRT

<213> Secuencia Artificial

30 <220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 61

Gln Gly Glu Phe Ser Cys Arg Asp Phe Asp Cys Thr Val  
1 5 10

35 <210> 62

<211> 15

<212> PRT

<213> Secuencia Artificial

40 <220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 62

Cys Gln Gly Glu Phe Ser Cys Arg Asp Phe Asp Cys Thr Val Phe  
1 5 10 15

<210> 63

<211> 12

50 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 63  
 Ala Gly Gly Tyr Lys Ser Ser Gly Asp Thr Val Ser  
 1               5                           10

5   <210> 64  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial

10   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 64  
 Tyr Cys Ala Gly Gly Tyr Lys Ser Ser Gly Asp Thr Val Ser Phe  
 1               5                           10                   15

15   <210> 65  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

20   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 65  
 Ala Gly Gly Tyr Lys Ser Thr Thr Asp Gly Ser Ala  
 1               5                           10

25   <210> 66  
 <211> 14  
 <212> PRT  
 30   <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

35   <400> 66  
 Cys Ala Gly Gly Tyr Lys Ser Thr Thr Asp Gly Ser Ala Phe  
 1               5                           10

40   <210> 67  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45   <400> 67  
 Gln Gln Gly Arg Arg Ser Val Asp Val Asp Asn Val  
 1               5                           10

50   <210> 68  
 <211> 27  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 68  
 Cys Ala Asp Ala Ala Thr Tyr Tyr Cys Gln Gln Gly Arg Arg Ser Val  
 1               5                       10                   15  
 Asp Val Asp Asn Val Phe Gly Gly Thr Glu  
 20                                       25

5           <210> 69  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial  
 10          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 69  
 Gln Gly Glu Phe Asn Cys Asp Gly Val Gly Cys Thr Thr  
 1               5                   10  
 15          <210> 70  
 <211> 16  
 <212> PRT  
 <213> Secuencia Artificial  
 20          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 70  
 Tyr Cys Gln Gly Glu Phe Asn Cys Asp Gly Val Gly Cys Thr Thr Phe  
 1               5                   10                   15  
 25          <210> 71  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 30          <400> 71  
 Gly Val Arg Asp Trp Gly Asp Ala Leu Asp Leu  
 1               5                   10  
 35          <210> 72  
 <211> 26  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 40          <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 72  
 Gly Val Arg Asp Trp Gly Asp Ala Leu Asp Leu Trp Gly Gln Gly Thr  
 1               5                   10                   15  
 <210> 73  
 <211> 10  
 <212> PRT  
 <213> Secuencia Artificial  
 50          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 73  
 Leu Tyr Asn Ser Val Val Gly Asp Asp Ile  
 1               5                   10  
 55          <210> 74  
 <211> 25  
 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

5 <400> 74  
Leu Tyr Asn Ser Val Val Gly Asp Asp Ile Trp Gly Pro Gly Thr Leu  
1 5 10 15

10 Val Thr Val Ser Leu Gly Gln Pro Lys  
20 25

15 <210> 75  
<211> 10  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

25 <400> 75  
Leu Tyr Asn Ser Val Val Gly Asp Asp Met  
1 5 10

30 <210> 76  
<211> 25  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

40 <400> 76  
Leu Tyr Asn Ser Val Val Gly Asp Asp Met Trp Gly Pro Gly Thr Leu  
1 5 10 15

45 Val Thr Val Ser Leu Gly Gln Pro Lys  
20 25

50 <210> 77  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

55 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

60 <400> 77  
Gly Met Pro Gly Ser Thr Ser Gly Asn Ser Asn Ile  
1 5 10

65 <210> 78  
<211> 27  
<212> PRT  
<213> Secuencia Artificial

70 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

75 <400> 78  
Gly Met Pro Gly Ser Thr Ser Gly Asn Ser Asn Ile Trp Gly Pro Gly  
1 5 10 15

80 Thr Leu Val Thr Val Ser Leu Gly Gln Pro Lys  
85 90

5 <210> 79  
 <211> 10  
 <212> PRT  
 <213> Secuencia Artificial  
 10 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 10 <400> 79  
 Leu Tyr Asn Ser Leu Val Gly Asp Asp Ile  
 1               5                   10  
 15 <210> 80  
 <211> 25  
 <212> PRT  
 <213> Secuencia Artificial  
 20 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 20 <400> 80  
 Leu Tyr Asn Ser Leu Val Gly Asp Asp Ile Trp Gly Pro Gly Thr Leu  
 1               5                   10                   15  
 25 Val Thr Val Ser Leu Gly Gln Pro Lys  
 20                   25  
 25 <210> 81  
 <211> 14  
 <212> PRT  
 <213> Secuencia Artificial  
 30 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 30 <400> 81  
 Lys Gly Asp Pro Gly His Pro Asn Gly Leu Phe Phe Thr Met  
 1               5                   10  
 35 <210> 82  
 <211> 29  
 <212> PRT  
 <213> Secuencia Artificial  
 40 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 40 <400> 82  
 Lys Gly Asp Pro Gly His Pro Asn Gly Leu Phe Phe Thr Met Trp Gly  
 1               5                   10                   15  
 45 Pro Gly Thr Leu Val Thr Val Ser Phe Gly Gln Pro Lys  
 20                   25  
 45 <210> 83  
 <211> 15  
 <212> PRT  
 50 <213> Secuencia Artificial  
 50 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 55 <400> 83

Gly Gly Gly Ser His Ser Gly Ser Ala Ile Tyr Asp Met Asp Pro  
 1                   5                   10                   15

<210> 84  
<211> 30  
 5 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Polipéptido sintético

<400> 84  
Gly Gly Gly Ser His Ser Gly Ser Ala Ile Tyr Asp Met Asp Pro Trp  
 1               5               10               15  
Gly Pro Gly Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20              25              30

<210> 85  
<211> 12  
 15 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 85  
Gly Thr Ser Arg Gly Ser Asp Tyr Arg Leu Asp Leu  
 1               5               10

<210> 86  
<211> 27  
 25 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 86  
Gly Thr Ser Arg Gly Ser Asp Tyr Arg Leu Asp Leu Trp Gly Gln Gly  
 1               5               10               15

Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 35              20              25

<210> 87  
<211> 12  
 40 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 87  
Gly Met Pro Ala Ser Thr Ser Gly Asn Ser Asn Ile  
 1               5               10

<210> 88  
<211> 27  
 45 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 88

Gly Met Pro Ala Ser Thr Ser Gly Asn Ser Asn Ile Trp Gly Pro Gly  
 1               5               10               15  
 Thr Leu Val Thr Val Ser Leu Gly Gln Pro Lys  
 20               25

5   <210> 89  
 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial

10   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 89  
**Asp Ala Ile Ala Asn Ile**  
 1               5

15   <210> 90  
 <211> 21  
 <212> PRT  
 <213> Secuencia Artificial

20   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 90  
**Asp Ala Ile Ala Asn Ile Trp Gly Pro Gly Thr Leu Val Thr Val Ser**  
 1               5               10               15

25   **Leu Gly Gln Pro Lys**  
 20

25   <210> 91  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial

30   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 91  
**Asp Lys Trp Met Val Phe Gly Asp Leu Arg Leu**  
 1               5               10

35   <210> 92  
 <211> 26  
 <212> PRT  
 40   <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45   <400> 92  
**Asp Lys Trp Met Val Phe Gly Asp Leu Arg Leu Trp Gly Pro Gly Thr**  
 1               5               10               15  
**Leu Val Thr Val Ser Ser Gly Gln Pro Lys**  
 20               25

50   <210> 93  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

55   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 93  
 Gln Gln Gly Arg Thr Tyr Ser Asp Val Ala Asn Val  
 1               5                   10

5    <210> 94  
 <211> 18  
 <212> PRT  
 <213> Secuencia Artificial

10   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 94  
 Thr Tyr Ser Asp Val Ala Asn Val Phe Gly Gly Gly Thr Glu Val Val  
 1               5                   10                   15

Val Lys

15   <210> 95  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

20   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 95  
 Gln Gln Gly Tyr Ser Ser Tyr Asn Val Asp Asn Ala  
 1               5                   10

25   <210> 96  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial

30   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 96  
 Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
 1               5                   10                   15

35   <210> 97  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

40   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45   <400> 97  
 Gln Gln Gly Tyr Ser Ser Ser Asn Val Asp Asn Ala  
 1               5                   10

50   <210> 98  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial

55   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 98  
 Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
 1               5                   10                   15

5                   <210> 99  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial  
 10                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 99  
 Leu Gly Thr Tyr Asp Cys Arg Ser Ala Asp Cys Asn Ala  
 1                    5                   10  
 15                  <210> 100  
 <211> 16  
 <212> PRT  
 <213> Secuencia Artificial  
 20                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 100  
 Ser Ala Asp Cys Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
 1                    5                   10                   15  
 25                  <210> 101  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial  
 30                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 101  
 Gln His Gly Tyr Tyr Ser Asn Val Asp Asn Ala  
 1                    5                   10  
 35                  <210> 102  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial  
 40                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 102  
 Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
 1                    5                   10                   15  
 45                  <210> 103  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 50                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 103  
 Gln Gln Gly Phe Ser Ser Arg Asn Val Asp Asn Ala  
 1                    5                   10  
 55                  <210> 104  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial  
 60                  <220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 104  
Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Val Lys  
1 5 10 15

5 <210> 105  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 105  
Gln Gln Gly Tyr Ser Ser Val Asn Val Asp Asn Ala  
1 5 10

15 <210> 106  
<211> 15  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 106  
Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
1 5 10 15

25 <210> 107  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

30 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 107  
Gln Gln Gly Tyr Thr Tyr Asn Asn Val Asp Asn Ala  
1 5 10

35 <210> 108  
<211> 15  
<212> PRT  
<213> Secuencia Artificial

40 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 108  
Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
1 5 10 15

45 <210> 109  
<211> 13  
<212> PRT  
<213> Secuencia Artificial

50 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 109  
Leu Gly Thr Tyr Asp Cys Arg Ser Gly Asp Cys Asn Val  
1 5 10

55 <210> 110  
<211> 14  
<212> PRT  
<213> Secuencia Artificial

60 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 110  
Leu Gly Thr Tyr Asp Cys Arg Ser Gly Asp Cys Asn Val  
1 5 10

<210> 110  
<211> 16  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 110  
Ser Gly Asp Cys Asn Val Phe Gly Gly Gly Thr Glu Val Val Val Val Lys  
1 5 10 15

<210> 111  
<211> 11  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 111  
Gln Gln Gly Tyr Thr Ser Asn Val Asp Asn Ala  
1 5 10

<210> 112  
<211> 15  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 112  
Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
1 5 10 15

<210> 113  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 113  
Gln Gln Gly Gln Thr Pro Glu Asn Val Asp Asn Ala  
1 5 10

<210> 114  
<211> 15  
<212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 114  
Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
1 5 10 15

<210> 115  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

5 <400> 115  
Gln Gln Gly Ser Thr Tyr Ser Asp Val Ala Asn Val  
1 5 10

10 <210> 116  
<211> 18  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

15 <400> 116  
Thr Tyr Ser Asp Val Ala Asn Val Phe Gly Gly Gly Thr Glu Val Val  
1 5 10 15

15 <210> 117  
<211> 12  
<212> PRT  
20 <213> Secuencia Artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

25 <400> 117  
Gln Gln Gly Ala Thr Tyr Ser Asp Val Ala Asn Val  
1 5 10

30 <210> 118  
<211> 18  
<212> PRT  
30 <213> Secuencia Artificial

35 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

35 <400> 118  
Thr Tyr Ser Asp Val Ala Asn Val Phe Gly Gly Gly Thr Glu Val Val  
1 5 10 15  
Val Lys

40 <210> 119  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

45 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

45 <400> 119  
Gln Gln Gly Thr Thr Tyr Ser Asp Val Ala Asn Val  
1 5 10

50 <210> 120  
<211> 18  
<212> PRT  
<213> Secuencia Artificial

55 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 120

Thr Tyr Ser Asp Val Ala Asn Val Phe Gly Gly Gly Thr Glu Val Val  
 1                   5                   10                   15

**Val Lys**  
 <210> 121  
 <211> 12  
 5 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

10 <400> 121  
 Gln Gln Gly Tyr Thr Arg Ser Asn Val Asp Asn Ala  
 1                   5                   10

<210> 122  
 15 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 122  
 Asn Val Asp Asn Ala Phe Gly Gly Thr Glu Val Val Lys  
 1                   5                   10                   15

25 <210> 123  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 123  
 Ala Gly Tyr Lys Ser Tyr Gly Asn Ala Asp Ile Asp  
 1                   5                   10

35 <210> 124  
 <211> 18  
 <212> PRT  
 <213> Secuencia Artificial

40 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 124  
 Ser Tyr Gly Asn Ala Asp Ile Asp Phe Gly Gly Thr Glu Val Val  
 1                   5                   10                   15

45 **Val Lys**  
 <210> 125  
 <211> 12  
 <212> PRT  
 50 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

55 <400> 125

Gln Gln Gly Tyr Thr Ser Ser Asn Val Asp Asn Ala  
1 5 10

<210> 126  
<211> 15  
5 <212> PRT  
<213> Secuencia Artificial

<220>  
10 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 126  
Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
1 5 10 15

<210> 127  
15 <211> 13  
<212> PRT  
<213> Secuencia Artificial

<220>  
20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 127  
Leu Val Ser Tyr Asp Cys Ser Ser Ala Asp Cys Asn Ala  
1 5 10

25 <210> 128  
<211> 16  
<212> PRT  
<213> Secuencia Artificial

30 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 128  
Ser Ala Asp Cys Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
1 5 10 15

35 <210> 129  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

40 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 129  
Gln Gln Ala Tyr Thr Ser Ser Asn Val Asp Asn Ala  
45 1 5 10

<210> 130  
<211> 15  
<212> PRT  
50 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

55 <400> 130  
Asn Val Asp Asn Ala Phe Gly Gly Gly Thr Glu Val Val Val Lys  
1 5 10 15

<210> 131  
<211> 4  
60 <212> PRT

<213> Secuencia Artificial

<220>

5 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 131  
Asp Gly Gly Leu  
1

10 <210> 132  
<211> 19  
<212> PRT  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 132  
Asp Gly Gly Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser Gly  
1 5 10 15

Gln Pro Lys

20 <210> 133  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 133  
Asp Pro Tyr Asp Thr Asn Thr Ser Leu Asp Ala Leu  
1 5 10

30 <210> 134  
<211> 27  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 134  
Asp Pro Tyr Asp Thr Asn Thr Ser Leu Asp Ala Leu Trp Gly Pro Gly  
1 5 10 15

Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
40 20 25

<210> 135  
<211> 10  
<212> PRT  
45 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

50 <400> 135  
Glu Gly Ser Asp Asp Asp Ser Phe Asp Leu  
1 5 10

<210> 136  
<211> 25  
55 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

5 <400> 136  
Glu Gly Ser Asp Asp Asp Ser Phe Asp Leu Trp Gly Pro Gly Thr Leu  
1 5 10 15

10 Val Thr Val Ser Ser Gly Gln Pro Lys  
20 25

10 <210> 137  
<211> 4  
<212> PRT  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 137  
Gly Gly Asp Leu  
1

20 <210> 138  
<211> 19  
<212> PRT  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 138  
Gly Gly Asp Leu Trp Gly Gln Gly Thr Leu Val Thr Val Ser Ser Gly  
1 5 10 15

30 Gln Pro Lys

30 <210> 139  
<211> 15  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 139  
Gly His Trp Ser Ala Gly Ala Thr Leu Tyr Gly Tyr Phe Ser Leu  
1 5 10 15

40 <210> 140  
<211> 30  
<212> PRT  
<213> Secuencia Artificial

45 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 140  
Gly His Trp Ser Ala Gly Ala Thr Leu Tyr Gly Tyr Phe Ser Leu Trp  
1 5 10 15

50 Gly Pro Gly Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
55 60 65 70 75 80 85 90

5       <210> 141  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial  
  
 10      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 10      <400> 141  
 Leu Ala Asn Tyr Asp Cys Ser Ser Gly Asp Cys Ser Val  
 1                   5                   10  
  
 15      <210> 142  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial  
  
 20      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 20      <400> 142  
 Cys Leu Ala Asn Tyr Asp Cys Ser Ser Gly Asp Cys Ser Val Phe  
 1                   5                   10                   15  
  
 25      <210> 143  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial  
  
 30      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 30      <400> 143  
 Gln Gly Asn Phe Asp Cys Ser Ser Ala Asp Cys Ser Ala  
 1                   5                   10  
  
 35      <210> 144  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial  
  
 40      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 40      <400> 144  
 Cys Gln Gly Asn Phe Asp Cys Ser Ser Ala Asp Cys Ser Ala Phe  
 1                   5                   10                   15  
  
 45      <210> 145  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial  
  
 50      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 50      <400> 145  
 Gln Gly Asn Phe Asp Cys Thr Ser Ala Asp Cys Ser Ala  
 1                   5                   10  
  
 55      <210> 146  
 <211> 15  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 146  
Cys Gln Gly Asn Phe Asp Cys Thr Ser Ala Asp Cys Ser Ala Phe  
1 5 10 15

<210> 147  
<211> 26  
<212> PRT  
10 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

15 <400> 147  
Asp Gly Thr Asp His Gly Phe Asn Ile Asp Leu Trp Gly Pro Gly Thr  
1 5 10 15

Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
20 25

<210> 148  
<211> 42  
20 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

25 <400> 148  
Thr Ser Thr Thr Val Thr Leu Gln Met Thr Ser Leu Thr Ala Ala Asp  
1 5 10 15

Thr Ala Thr Tyr Phe Cys Ala Ser Gly Asn Val Trp Gly Pro Gly Thr  
20 25 30

Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
35 40

<210> 149  
<211> 21  
30 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

35 <400> 149  
Gly Val Ser Thr Asn Val Trp Gly Pro Gly Thr Leu Val Thr Val Ser  
1 5 10 15

Ser Gly Gln Pro Lys  
20

40 <210> 150  
<211> 8  
<212> PRT  
<213> Secuencia Artificial

45 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 150

Phe Cys Thr Arg Gly Val Lys Phe  
 1               5

<210> 151  
<211> 26  
 5 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 151  
Asp Gly Ser Asp His Gly Phe Asn Ile Asp Leu Trp Gly Pro Gly Thr  
 1               5               10               15

Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20               25

<210> 152  
 15 <211> 5  
<212> PRT  
<213> Secuencia Artificial

<220>  
 20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 152  
Asn Ala Ala Ile Leu  
 1               5

<210> 153  
 25 <211> 20  
<212> PRT  
<213> Secuencia Artificial

<220>  
 30 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 153  
Asn Ala Ala Ile Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser  
 1               5               10               15

Gly Gln Pro Lys  
 35               20

<210> 154  
<211> 10  
<212> PRT  
 40 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 154  
Ser Arg Ser Thr Ser Tyr Tyr Ile Asn Leu  
 45               1               5               10

<210> 155  
<211> 25  
<212> PRT  
 50 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 155  
 Ser Arg Ser Thr Ser Tyr Tyr Ile Asn Leu Trp Gly Pro Gly Thr Leu  
 1               5               10               15

Val Thr Val Ser Ser Gly Gln Pro Lys  
 20               25

5    <210> 156  
 <211> 29  
 <212> PRT  
 <213> Secuencia Artificial

10   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 156  
 Gly Gly Asp Ala Gly Tyr Gly Ser Phe Asp Ala Phe Gly Pro Trp Gly  
 1               5               10               15

Pro Gly Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20               25

15   <210> 157  
 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial

20   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 157  
 Gly Val Ser Thr Asp Val  
 1               5

25   <210> 158  
 <211> 21  
 <212> PRT  
 <213> Secuencia Artificial

30   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 158  
 Gly Val Ser Thr Asn Val Trp Gly Pro Gly Thr Leu Val Thr Val Ser  
 1               5               10               15

Ser Gly Gln Pro Lys  
 35    20

35   <210> 159  
 <211> 27  
 <212> PRT  
 40   <213> Secuencia Artificial

40   <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45   <400> 159  
 Asn Val Gly Ser Ser Ser Tyr Tyr Asn Leu Asn Leu Trp Gly Pro Gly  
 1               5               10               15

Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20               25

5           <210> 160  
 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial  
 10          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 160  
 Gly Val Ser Thr Ser Val  
 1                 5  
 15          <210> 161  
 <211> 21  
 <212> PRT  
 <213> Secuencia Artificial  
 20          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 161  
 Gly Val Ser Thr Asn Val Trp Gly Pro Gly Thr Leu Val Thr Val Ser  
 1                 5                 10                 15  
 Ser Gly Gln Pro Lys  
 20  
 <210> 162  
 <211> 14  
 25          <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 30          <400> 162  
 Gly Gly Tyr Ala Gly Ala Gly Tyr Phe Asp Ala Phe Asn Pro  
 1                 5                 10  
 <210> 163  
 35          <211> 29  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 40          <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 163  
 Gly Gly Tyr Ala Gly Ala Gly Tyr Phe Asp Ala Phe Asn Pro Trp Gly  
 1                 5                 10                 15  
 Pro Gly Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20                 25  
 45          <210> 164  
 <211> 4  
 <212> PRT  
 <213> Secuencia Artificial  
 50          <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 164  
 Asn Tyr Asn Leu  
 1

5       <210> 165  
 <211> 19  
 <212> PRT  
 <213> Secuencia Artificial  
 10      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 10      <400> 165  
 Asn Tyr Asn Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser Gly  
 1               5                   10                   15  
 Gln Pro Lys  
 15      <210> 166  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 20      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 20      <400> 166  
 Arg Asp Gly Phe Ser Thr Asp Arg Tyr Phe Asn Leu  
 1               5                   10  
 <210> 167  
 25      <211> 26  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 30      <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 167  
 Asp Gly Phe Ser Thr Asp Arg Tyr Phe Asn Leu Trp Gly Pro Gly Thr  
 1               5                   10                   15  
 Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20               25  
 35      <210> 168  
 <211> 14  
 <212> PRT  
 <213> Secuencia Artificial  
 40      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 168  
 Asp Arg Gly Thr Gly Ser Gly Asp Tyr Thr Pro Phe Asn Leu  
 1               5                   10  
 45      <210> 169  
 <211> 25  
 <212> PRT  
 <213> Secuencia Artificial  
 50      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 169

Gly Ser Gly Asp Tyr Thr Pro Phe Asn Leu Trp Gly Pro Gly Thr Leu  
 1                   5                   10                   15

Val Thr Val Ser Ser Gly Gln Pro Lys  
 20                   25

<210> 170  
<211> 5  
 5 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 170  
Asp Ala Ala Ile Leu  
 1                   5

<210> 171  
<211> 20  
 15 <212> PRT  
<213> Secuencia Artificial

<220>  
 20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 171  
Asn Ala Ala Ile Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser  
 1                   5                   10                   15

Gly Gln Pro Lys  
 20

<210> 172  
<211> 11  
 25 <212> PRT  
<213> Secuencia Artificial

<220>  
 30 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 172  
Gly Pro Tyr Val Asp Ser Thr Tyr Tyr Asn Leu  
 1                   5                   10

<210> 173  
<211> 26  
 35 <212> PRT  
<213> Secuencia Artificial

<220>  
 40 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 173  
Gly Pro Tyr Val Asp Ser Thr Tyr Tyr Asn Leu Trp Gly Pro Gly Thr  
 1                   5                   10                   15

Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 45                   20                   25

<210> 174  
<211> 10  
 50 <212> PRT  
<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 174  
 Gly Ser Gly Asp Tyr Thr Pro Phe Asn Leu  
 5               5                   10

<210> 175  
<211> 25  
<212> PRT

10 <213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

15 <400> 175  
 Gly Ser Gly Asp Tyr Thr Pro Phe Asn Leu Trp Gly Pro Gly Thr Leu  
 1               5                   10                   15

Val Thr Val Ser Ser Gly Gln Pro Lys  
 20                                   25

<210> 176  
<211> 12  
20 <212> PRT  
<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

25 <400> 176  
 Tyr Tyr Asp Gly Ala Asp Tyr His Thr Tyr Asn Leu  
 1               5                   10

<210> 177  
<211> 27  
<212> PRT  
<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

35 <400> 177  
 Tyr Tyr Asp Gly Ala Asp Tyr His Thr Tyr Asn Leu Trp Gly Pro Gly  
 1               5                   10                   15

Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20                                   25

40 <210> 178  
<211> 11  
<212> PRT  
<213> Secuencia Artificial

45 <220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 178  
 Glu Phe Gly Asn Asn Gly Trp Asn Ile Asp Leu  
 50 1               5                   10

<210> 179  
<211> 26  
<212> PRT

55 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

5 <400> 179  
 Glu Phe Gly Asn Asn Gly Trp Asn Ile Asp Leu Trp Gly Pro Gly Thr  
 1               5                           10                           15

Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20   25

10 <210> 180  
 <211> 10  
 <212> PRT  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 180  
 Val Glu Tyr Gly Asn Asp Trp Gly Asn Leu  
 1               5                           10

20 <210> 181  
 <211> 25  
 <212> PRT  
 <213> Secuencia Artificial

25 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 181  
 Val Glu Tyr Gly Asn Asp Trp Gly Asn Leu Trp Gly Pro Gly Thr Leu  
 1               5                           10                           15

Val Thr Val Ser Ser Gly Gln Pro Lys  
 20   25

30 <210> 182  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 182  
 Tyr Phe Asp Gly Ala Asp Tyr His Thr Tyr Asn Leu  
 1               5                           10

40 <210> 183  
 <211> 27  
 <212> PRT  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 183  
 Tyr Phe Asp Gly Ala Asp Tyr His Thr Tyr Asn Leu Trp Gly Pro Gly  
 1               5                           10                           15

Thr Leu Val Thr Val Ser Ser Gly Gln Pro Lys  
 20   25

50

5                   <210> 184  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial  
 10                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 184  
 Arg Phe Ser Gly Gly Gly Tyr Gly Tyr Asp Leu  
 1                    5                   10  
 15                  <210> 185  
 <211> 25  
 <212> PRT  
 <213> Secuencia Artificial  
 20                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 185  
 Phe Ser Gly Gly Gly Tyr Gly Tyr Asp Leu Trp Gly Pro Gly Thr Leu  
 1                    5                   10                   15  
 Val Thr Val Ser Ser Gly Gln Pro Lys  
 20                   25  
 25                  <210> 186  
 <211> 4  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 30                  <400> 186  
 Asp Arg Asp Leu  
 1  
 <210> 187  
 35                  <211> 19  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 40                  <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 187  
 Asp Arg Asp Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser Gly  
 1                    5                   10                   15  
 Gln Pro Lys  
 45                  <210> 188  
 <211> 4  
 <212> PRT  
 <213> Secuencia Artificial  
 50                  <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 188  
 Gly Leu Asp Leu  
 1  
 55

5 <210> 189  
<211> 20  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 189  
Tyr Gly Leu Asp Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser  
1 5 10 15

15 Gly Gln Pro Lys  
20

15 <210> 190  
<211> 11  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 190  
Tyr Asp Val Asp Ser Val Ser Ala Tyr Asp Leu  
1 5 10

25 <210> 191  
<211> 26  
<212> PRT  
<213> Secuencia Artificial

30 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 191  
Tyr Asp Val Asp Ser Val Ser Ala Tyr Asp Leu Trp Gly Pro Gly Thr  
1 5 10 15

35 Ieu Val Thr Val Ser Ser Gly Gln Pro Lys  
20 25

35 <210> 192  
<211> 11  
<212> PRT  
<213> Secuencia Artificial

40 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 192  
Glu Val Val Gly Tyr Asp Tyr Ser Gly Asp Leu  
1 5 10

45 <210> 193  
<211> 26  
<212> PRT  
<213> Secuencia Artificial

50 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 193

```

Glu Val Val Gly Tyr Asp Tyr Ser Gly Asp Leu Trp Gly Pro Gly Thr
1           5          10          15

Leu Val Thr Val Ser Ser Gly Gln Pro Lys
20          25

<210> 194
<211> 8
5 <212> PRT
<213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: Péptido sintético
10 <400> 194
Asp Pro Tyr Asp Asp Pro Thr Tyr
1           5

<210> 195
15 <211> 9
<212> PRT
<213> Secuencia Artificial

<220>
20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 195
Asp Pro Tyr Asp Asp Pro Thr Tyr Arg
1           5

25 <210> 196
<211> 12
<212> PRT
<213> Secuencia Artificial

30 <220>
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 196
Gly Gly Leu Val Lys Pro Gly Ala Ser Leu Thr Leu
1           5          10

35 <210> 197
<211> 33
<212> DNA
<213> Secuencia Artificial
40 <220>
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 197
45 acgagctacg cacgaactgc aggttccac tcc      33

<210> 198
<211> 33
<212> DNA
50 <213> Secuencia Artificial

<220>
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético
55 <400> 198
acgagctacg cacgaatagc aggtgtccac tcc      33

<210> 199
<211> 33

```

<212> DNA  
<213> Secuencia Artificial

5 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 199  
acgagctacg cacgargtac aggtgtccac tcc 33

10 <210> 200  
<211> 33  
<212> DNA  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 200  
acgagctacg cacgagcyac agmtgtccac tcc 33

20 <210> 201  
<211> 33  
<212> DNA  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 201  
acgagctacg cacgaactgc aggtgtccwm tcc 33

<210> 202  
<211> 33  
<212> DNA

35 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

40 <400> 202  
acgagctacg cacgarctrac aggtgtkac tcc 33

<210> 203  
<211> 33

45 <212> DNA  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

50 <400> 203  
acgagctacg cacgagctaw mggtgtccac tcc 33

<210> 204  
55 <211> 32  
<212> DNA  
<213> Secuencia Artificial

<220>  
60 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 204  
acgagctacg cacgacacctca ggtgtccact cc 32

65 <210> 205  
<211> 33

<212> DNA  
 <213> Secuencia Artificial  
 5      <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 205  
       acgagctacg cacgagctac aggtgctcac tcc        33  
 10     <210> 206  
       <211> 33  
       <212> DNA  
       <213> Secuencia Artificial  
 15     <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 206  
       acgagctacg cacgaactgc aggtgtcctc tct        33  
 20     <210> 207  
       <211> 33  
       <212> DNA  
       <213> Secuencia Artificial  
 25     <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 207  
 30     acgagctacg cacgaaytgc aggtgtccay tgc        33  
       <210> 208  
       <211> 33  
       <212> DNA  
 35     <213> Secuencia Artificial  
       <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 40     <400> 208  
       acgagctacg cacgagctam mggtgtccac ttc        33  
       <210> 209  
       <211> 36  
 45     <212> DNA  
       <213> Secuencia Artificial  
       <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 50     <400> 209  
       acgagctacg cacgactcct gtcaktaact kcagg        36  
       <210> 210  
 55     <211> 33  
       <212> DNA  
       <213> Secuencia Artificial  
       <220>  
 60     <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 210  
       acgagctacg cacgaaaactg caggtgtctc tct        33  
 65     <210> 211  
       <211> 33

<212> DNA  
<213> Secuencia Artificial

5 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 211  
acgagctacg cacgarctr c aggygtccac tct 33

10 <210> 212  
<211> 33  
<212> DNA  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 212  
acgagctacg cacgaccaag ctgtatcctt tcc 33

20 <210> 213  
<211> 33  
<212> DNA  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 213  
acgagctacg cacgaccaag ctgtgtccctr tcc 33

30 <210> 214  
<211> 35  
<212> DNA  
<213> Secuencia Artificial

35 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 214  
acgagctacg cacgatgtt acagycvttc ckgg 35

40 <210> 215  
<211> 35  
<212> DNA  
<213> Secuencia Artificial

45 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 215  
acgagctacg cacgatgtt acagcctttc ctgg 35

50 <210> 216  
<211> 34  
<212> DNA  
<213> Secuencia Artificial

55 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 216  
acgagctacg cacgattaa aagggtcca gtgt 34

60 <210> 217  
<211> 37

<212> DNA  
 <213> Secuencia Artificial  
 5      <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 217  
       acgagctacg cacgataytt taaaargtgt cmagtgt        37  
 10     <210> 218  
       <211> 35  
       <212> DNA  
       <213> Secuencia Artificial  
 15     <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 218  
       acgagctacg cacgagttt aaaagggtgc ctgtg        35  
 20     <210> 219  
       <211> 35  
       <212> DNA  
       <213> Secuencia Artificial  
 25     <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 219  
 30     acgagctacg cacgactyt ttaaaaggkgtc cagwg        35  
       <210> 220  
       <211> 36  
       <212> DNA  
 35     <213> Secuencia Artificial  
       <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 40     <400> 220  
       acgagctacg cacgacytt amatggatc cagtgt        36  
       <210> 221  
       <211> 36  
 45     <212> DNA  
       <213> Secuencia Artificial  
       <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 50     <400> 221  
       acgagctacg cacgactttt acatggttc aagtgt        36  
       <210> 222  
 55     <211> 34  
       <212> DNA  
       <213> Secuencia Artificial  
       <220>  
 60     <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 222  
       acgagctacg cacgatgttc cctgcatatg tcyt        34  
 65     <210> 223  
       <211> 32

<212> DNA  
 <213> Secuencia Artificial  
 5      <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 223  
       acgagctacg cacgaatggc agcwgcyc ca ag            32  
 10     <210> 224  
       <211> 35  
       <212> DNA  
       <213> Secuencia Artificial  
 15     <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 224  
       acgagctacg cacgatttat caagggtgtc attgt            35  
 20     <210> 225  
       <211> 36  
       <212> DNA  
       <213> Secuencia Artificial  
 25     <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 225  
 30     acgagctacg cacgactttt aaaagwtgtc cagkgt            36  
       <210> 226  
       <211> 35  
       <212> DNA  
 35     <213> Secuencia Artificial  
       <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 40     <400> 226  
       acgagctacg cacgagtgtac agtccttcgt ggttag            35  
       <210> 227  
       <211> 34  
 45     <212> DNA  
       <213> Secuencia Artificial  
       <220>  
       <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 50     <400> 227  
       acgagctacg cacgacttcc tgatggcagt ggtt            34  
       <210> 228  
 55     <211> 34  
       <212> DNA  
       <213> Secuencia Artificial  
       <220>  
 60     <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
       <400> 228  
       acgagctacg cacgaagcta caggtatcca atcc            34  
 65     <210> 229  
       <211> 37

<212> DNA  
<213> Secuencia Artificial

5 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 229  
cactggtg agtcaatgca gacagatggg ggtgtcg 37

10 <210> 230  
<211> 35  
<212> DNA  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 230  
cactggtg agtcaagacc gatggggctg ttgtt 35

20 <210> 231  
<211> 37  
<212> DNA  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 231  
cactggtg agtcaacaga ctgatgggg tgttttt 37

30 <210> 232  
<211> 35  
<212> DNA  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

40 <400> 232  
cactggtg agtcaagaca gatggggctg ttgtt 35

<210> 233  
<211> 38

45 <212> DNA  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

50 <400> 233  
acgagctacg cacgagacat ywwgatgacc cagtctcc 38

<210> 234  
55 <211> 37  
<212> DNA  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 234  
cactggtg agtcacagtt ggtgcagcat cagcccg 37

65 <210> 235  
<211> 41

<212> DNA  
 <213> Secuencia Artificial

5 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 235  
 cctatccctt gtgtgccttg gcagtcacga gctacgcacg a 41

10 <210> 236  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 236  
 ccatctcatc cctgcgtgtc tccgactca gtagtcactc cactggtgtg agtca 55

20 <210> 237  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

25 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 237  
 ccatctcatc cctgcgtgtc tccgactca gtagcgtcac cactggtgtg agtca 55

30 <210> 238  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

35 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 238  
 ccatctcatc cctgcgtgtc tccgactca gtagacgag cactggtgtg agtca 55

40 <210> 238  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

45 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 239  
 ccatctcatc cctgcgtgtc tccgactca gacgctacgt cactggtgtg agtca 55

50 <210> 239  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

55 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 240  
 ccatctcatc cctgcgtgtc tccgactca tcacgcgaga cactggtgtg agtca 55

60 <210> 241  
 <211> 55

<212> DNA  
 <213> Secuencia Artificial

5 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 241  
 ccatctcatc cctgcgtgtc tccgactca cgcgataaca cactggtg agtca 55

10 <210> 242  
 <211> 62  
 <212> DNA  
 <213> Secuencia Artificial

15 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 242  
 ccatctcatc cctgcgtgtc tccgactca tgataacgtct gggccagtgg gaagactgat 60  
 gg 62

20 <210> 243  
 <211> 59  
 <212> DNA  
 <213> Secuencia Artificial

25 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 243  
 cctatccccct gtgtccctg gcagtctca atcagacacg atggagactg ggctgcgct 59  
 <210> 244  
 <211> 61  
 <212> DNA

35 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

40 <400> 244  
 ccatctcatc cctgcgtgtc tccgactca tcacgtacta gaagaggagg acagwaggcg 60  
 c 61

<210> 245  
 <211> 48

45 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

50 <400> 245  
 cctatccccct gtgtccctg gcagtctca atggacatga gggcccc 48

<210> 246  
 <211> 67

55 <212> DNA  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

60 <400> 246

	ccatctcatc cctgcgtgtc tccgactca	60
	g tacagatcg tttgtcc ttgagttcct	
	cagagga	67
	<210> 247	
	<211> 45	
5	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
10	<400> 247	
	cctatcccgttg cgcgttccat atggcctgca ccccg	45
	<210> 248	
15	<211> 66	
	<212> DNA	
	<213> Secuencia Artificial	
	<220>	
20	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
	<400> 248	
	ccatctcatc cctgcgtgtc tccgactca ctgtgcgtcg cagaagacsg atggccctt	60
	ggtgaa	66
25	<210> 249	
	<211> 64	
	<212> DNA	
	<213> Secuencia Artificial	
30	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
	<400> 249	
	ccatctcatc cctgcgtgtc tccgactca acgcgagtat gaagacsgat gggcccttgg	60
	tgaa	64
35	<210> 250	
	<211> 64	
	<212> DNA	
	<213> Secuencia Artificial	
40	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
	<400> 250	
	ccatctcatc cctgcgtgtc tccgactca cacgctacgt gaagacsgat gggcccttgg	60
45	tgaa	64
	<210> 251	
	<211> 65	
	<212> DNA	
50	<213> Secuencia Artificial	
	<220>	
	<223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
55	<400> 251	
	ccatctcatc cctgcgtgtc tccgactca tgagtca gtaaagacsga tggcccttgg	60
	gtgaa	65

5	<210> 252 <211> 69 <212> DNA <213> Secuencia Artificial	
10	<220> <223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
15	<400> 252 ccatctcatc cctgcgtgtc tccgactcag atctactgac atgatgaaga csgatgggcc	60
20	cttggtgaa  <210> 253 <211> 69 <212> DNA <213> Secuencia Artificial	69
25	<220> <223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
30	<400> 253 ccatctcatc cctgcgtgtc tccgactcag agtagtgatc tcacagaaga csgatgggcc  cttggtgaa	60
35	<210> 254 <211> 51 <212> DNA <213> Secuencia Artificial	
40	<220> <223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
45	<400> 254 cctatcccct gtgtcccttg gcagtctcag atggactgga cctggaggvat c 51  <210> 255 <211> 51 <212> DNA <213> Secuencia Artificial	
50	<220> <223> Descripción de la Secuencia Artificial: oligonucleótido sintético	
55	<400> 255 cctatcccct gtgtcccttg gcagtctcag atggactgga ttggaggat c 51  <210> 256 <211> 51 <212> DNA <213> Secuencia Artificial	
60	<220> <223> Descripción de la Secuencia Artificial: oligonucleótido sintético	

<400> 257  
 cctatcccct gtgtgccttg gcagtctcag atggactgga cctggaggkt c 51  
 5 <210> 258  
 <211> 52  
 <212> DNA  
 <213> Secuencia Artificial  
 10 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 258  
 cctatcccct gtgtgccttg gcagtctcag atggacatac ttgttccac gc 52  
 15 <210> 259  
 <211> 52  
 <212> DNA  
 <213> Secuencia Artificial  
 20 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 259  
 cctatcccct gtgtgccttg gcagtctcag atggacacac ttgctacac ac 52  
 25 <210> 260  
 <211> 52  
 <212> DNA  
 30 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 35 <400> 260  
 cctatcccct gtgtgccttg gcagtctcag atgtctgtct ctttcctcat ct 52  
 <210> 261  
 <211> 51  
 40 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 45 <400> 261  
 cctatcccct gtgtgccttg gcagtctcag atggggtaa ccggccatctt c 51  
 <210> 262  
 50 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 55 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 262  
 cctatcccct gtgtgccttg gcagtctcag atggagttkg grctgagctg g 51  
 60 <210> 263  
 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial  
 65 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 263  
 cctatccct gtgtgcctg gcagtcttag atggagttk ggctkagctg g 51  
 5 <210> 264  
 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial  
 10 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 264  
 cctatccct gtgtgcctg gcagtcttag atggaaactgg ggctccgctg g 51  
 15 <210> 265  
 <211> 51  
 <212> DNA  
 <213> Secuencia Artificial  
 20 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 265  
 cctatccct gtgtgcctg gcagtcttag atggarttg ggctgwgctg g 51  
 25 <210> 266  
 <211> 53  
 <212> DNA  
 30 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 35 <400> 266  
 cctatccct gtgtgcctg gcagtcttag atgaarcayc tgtggttctt cct 53  
 <210> 267  
 <211> 67  
 40 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 45 <400> 267  
 ccatctcatc cctgcgtgtc tccgacttag acgacagctc gaagatgaag acagatggtg 60  
 cagccac 67  
 <210> 268  
 50 <211> 67  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 55 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 268  
 ccatctcatc cctgcgtgtc tccgacttag tcgatcacgt gaagatgaag acagatggtg 60  
 cagccac 67  
 60 <210> 269  
 <211> 67  
 <212> DNA

<213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 5 <400> 269  
ccatctcatc cctgcgtgtc tccgacttag tacacacact gaagatgaag acagatgg 60  
cagccac 67

10 <210> 270  
 <211> 68  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 15 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 270  
ccatctcatc cctgcgtgtc tccgacttag ttagtgtga tgaagatgaa gacagatgg 60  
gcagccac 68

20 <210> 271  
 <211> 57  
 <212> DNA  
 <213> Secuencia Artificial  
 25 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 271  
cctatcccgtgtgccttg gcagtcttag atggacatga gggtscygc tcagctc 57

30 <210> 272  
 <211> 57  
 <212> DNA  
 <213> Secuencia Artificial  
 35 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 272  
cctatcccgtgtgccttg gcagtcttag atggacatga grgtcctcg tcagctc 57

40 <210> 273  
 <211> 54  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

45 <400> 273  
cctatcccgtgtgccttg gcagtcttag atggaaggccc cagcdcgact tctc 54  
 <210> 274  
 <211> 54  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

50 <400> 274  
cctatcccgtgtgccttg gcagtcttag atggaaggccc cagcdcgact tctc 54  
 <210> 274  
 <211> 54  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

55 <400> 274  
cctatcccgtgtgccttg gcagtcttag atgaaacccc cagcgcagact tctc 54

60 <210> 274  
 <211> 54  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 274  
cctatcccgtgtgccttg gcagtcttag atgaaacccc cagcgcagact tctc 54

5       <210> 275  
 <211> 54  
 <212> DNA  
 <213> Secuencia Artificial  
 10      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 15      <400> 275  
 cctatccct gtgtgccttg gcagtctcag atgggtttgc agacccagg ttc 54  
 <210> 276  
 <211> 54  
 <212> DNA  
 <213> Secuencia Artificial  
 20      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 25      <400> 276  
 cctatcccgt gtgtgccttg gcagtctcag atggggtccc aggttcacct cctc 54  
 <210> 277  
 <211> 54  
 <212> DNA  
 <213> Secuencia Artificial  
 30      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 277  
 cctatcccgt gtgtgccttg gcagtctcag atgaggctcc ytgcctagct cctg 54  
 35      <210> 278  
 <211> 64  
 <212> DNA  
 <213> Secuencia Artificial  
 40      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 278  
 ccatctcatc cctgcgtgtc tccgactcag cgttagactag agggcgaa cagagtgacm 60  
 45      gtgg 64  
 <210> 279  
 <211> 64  
 <212> DNA  
 <213> Secuencia Artificial  
 50      <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 279  
 ccatctcatc cctgcgtgtc tccgactcag cgttagactag agggcgaa cagagtgacc 60  
 55      gakg 64  
 <210> 280  
 <211> 66  
 <212> DNA  
 <213> Secuencia Artificial  
 60      <220>

<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 280  
`ccatctcatc cctgcgtgta tccgactca g tctagcgact ataggcgaa aacagagtga` 60

5           `cmgtgg` 66

<210> 281  
<211> 66  
<212> DNA  
<213> Secuencia Artificial

10          <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 281  
`ccatctcatc cctgcgtgta tccgactca g tctagcgact ataggcgaa aacagagtga` 60

15          `ccgakg` 66

<210> 282  
<211> 65  
<212> DNA  
20          <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

25          <400> 282  
`ccatctcatc cctgcgtgta tccgactca g tagcgccgc tagggcgaa acagagtga` 60

`mgtgg` 65

<210> 283  
<211> 65  
30          <212> DNA  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

35          <400> 283  
`ccatctcatc cctgcgtgta tccgactca g tagcgccgc tagggcgaa acagagtga` 60

`cgakg` 65

<210> 284  
40          <211> 64  
<212> DNA  
<213> Secuencia Artificial

<220>  
45          <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 284  
`ccatctcatc cctgcgtgta tccgactca g atagagtact agggcgaa cagagtga` 60

`gtgg` 64

<210> 285  
50          <211> 64  
<212> DNA  
<213> Secuencia Artificial

55          <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 285  
 ccatctcatc cctgcgtgtc tccgacttag atagagtact agggyggaa cagagtgacc 60  
 gakg 64

5 <210> 286  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 286  
 cctatcccgtgtgccttg gcagttctcgatggacctgctcccttccca 55

15 <210> 287  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 287  
 cctatcccgtgtgccttg gcagttctcgatggccggcttcccttccca 55

25 <210> 288  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 288  
 cctatcccgtgtgccttg gcagttctcgatggcctggctcccttccca 55

35 <210> 289  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

40 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 289  
 cctatcccgtgtgccttg gcagttctcgatggcctggatcccttcccttccca 55

45 <210> 290  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

50 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 290  
 cctatcccgtgtgccttg gcagttctcgatggcctggatcccttcccttccca 55

55 <210> 291  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial

60 <220>

<223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
5 <400> 291  
cctatccct gtgtgccttg gcagtctcag atgcccctggg tcatgctcct cctga 55  
<210> 292  
<211> 55  
<212> DNA  
<213> Secuencia Artificial  
10 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
15 <400> 292  
cctatccct gtgtgccttg gcagtctcag atggcctgga ctccctctt tctgt 55  
<210> 293  
<211> 55  
<212> DNA  
20 <213> Secuencia Artificial  
  
<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
25 <400> 293  
cctatccct gtgtgccttg gcagtctcag atggagaaga agaggagacc tgggg 55  
<210> 294  
<211> 55  
30 <212> DNA  
<213> Secuencia Artificial  
  
<220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
35 <400> 294  
cctatccct gtgtgccttg gcagtctcag atggcctgga ccgctctcct tctga 55  
<210> 295  
40 <211> 55  
<212> DNA  
<213> Secuencia Artificial  
  
<220>  
45 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
<400> 295  
cctatccct gtgtgccttg gcagtctcag atggcctgga ccgttctcct cctcg 55  
50 <210> 296  
<211> 55  
<212> DNA  
<213> Secuencia Artificial  
  
<220>  
55 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
  
<400> 296  
cctatccct gtgtgccttg gcagtctcag atggcatgga tccctctt cctcg 55  
60 <210> 297  
<211> 55  
<212> DNA  
<213> Secuencia Artificial  
65 <220>

<223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 297  
 cctatccct gtgtgccttg gcagtctcag atggcctgga tccctctact tctcc 55  
 5  
 <210> 298  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial  
 10  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 298  
 cctatccct gtgtgccttg gcagtctcag atggcctgga yccctctcct gctcc 55  
 15  
 <210> 299  
 <211> 55  
 <212> DNA  
 20 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 25 <400> 299  
 cctatccct gtgtgccttg gcagtctcag atggcatggg ccacactcct gctcc 55  
 <210> 300  
 <211> 55  
 30 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 35 <400> 300  
 cctatccct gtgtgccttg gcagtctcag atggcctgga cccctctcg gctca 55  
 <210> 301  
 40 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 45 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 301  
 cctatccct gtgtgccttg gcagtctcag atggcctggg ttccttcta cttac 55  
 50 <210> 302  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial  
 55 <220>  
 <223> Descripción de la Secuencia Artificial: oligonucleótido sintético  
 <400> 302  
 cctatccct gtgtgccttg gcagtctcag atggcctgga ccccaactcct cttcc 55  
 60 <210> 303  
 <211> 55  
 <212> DNA  
 <213> Secuencia Artificial  
 65 <220>

<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 303  
cctatccct gtgtgcctg gcagtctcag atggcctggg ctccctcgct cctca 55

5 <210> 304  
<211> 55  
<212> DNA  
<213> Secuencia Artificial

10 <220>  
<223> Descripción de la Secuencia Artificial: oligonucleótido sintético

<400> 304  
cctatccct gtgtgcctg gcagtctcag atggcctggg ctccactact tctca 55

15 <210> 305  
<211> 19  
<212> PRT  
20 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

25 <400> 305  
Gly Phe Ala Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser Gly  
1 5 10 15

Gln Pro Lys

<210> 306  
<211> 4  
30 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

35 <400> 306  
Thr Thr Thr Thr  
1

<210> 307  
40 <211> 11  
<212> PRT  
<213> Secuencia Artificial

<220>  
45 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 307  
Arg Gly Tyr Tyr Ser Gly Thr Tyr Val Asp Tyr  
1 5 10

50 <210> 308  
<211> 8  
<212> PRT  
<213> Secuencia Artificial

55 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 308  
Ser Arg Thr Gly Ile Phe Asp Tyr  
1 5

5       <210> 309  
 <211> 8  
 <212> PRT  
 <213> Secuencia Artificial  
 10      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 10      <400> 309  
 Ser Arg Ser Gly Ile Phe Asp Tyr  
 1               5  
 15      <210> 310  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial  
 20      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 20      <400> 310  
 Arg Gly Tyr Tyr Ser Asn Thr Tyr Val Asp Tyr  
 1               5                           10  
 25      <210> 311  
 <211> 7  
 <212> PRT  
 <213> Secuencia Artificial  
 30      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 30      <400> 311  
 Leu Thr Ala Tyr Phe Asp Tyr  
 1               5  
 35      <210> 312  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial  
 40      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 40      <400> 312  
 Arg Phe Tyr Asp Gly Ser Tyr Tyr Phe Asp Tyr  
 1               5                           10  
 45      <210> 313  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial  
 50      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 50      <400> 313  
 Arg Gly Tyr Tyr Ala Asp Thr Tyr Val Asp Tyr  
 1               5                           10  
 55      <210> 314  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

5       <400> 314  
 His Glu Pro Leu Asn Trp Phe Pro Tyr  
 1               5

10      <210> 315  
 <211> 8  
 <212> PRT  
 <213> Secuencia Artificial

15      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

20      15     <400> 315  
 Lys Tyr Gly Gly Gly Phe Asp Tyr  
 1               5

25      <210> 316  
 <211> 6  
 <212> PRT  
 <213> Secuencia Artificial

30      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

35      25     <400> 316  
 Gly Ser Ala Phe Ala Tyr  
 1               5

40      <210> 317  
 <211> 10  
 <212> PRT  
 <213> Secuencia Artificial

45      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

50      45     <400> 317  
 Tyr Tyr Arg Asn Tyr Gly Gly Phe Asp Tyr  
 1               5               10

55      <210> 318  
 <211> 10  
 <212> PRT  
 <213> Secuencia Artificial

60      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

65      55     <400> 318  
 Glu Asp Tyr Tyr Ser Asp Gln Phe Ala Tyr  
 1               5               10

70      <210> 319  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial

75      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

80      65     <400> 319  
 Ser Gln Ser Thr His Val Pro Trp Thr  
 1               5

5        <210> 320  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial  
 10      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 10      <400> 320  
 Gln Gln His Tyr Ser Thr Pro Pro Thr  
 1                5  
 15      <210> 321  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial  
 20      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 20      <400> 321  
 Gln Gln His Phe Ser Thr Pro Pro Thr  
 1                5  
 25      <210> 322  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial  
 30      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 30      <400> 322  
 Gln Gln His Tyr Ser Thr Pro Phe Thr  
 1                5  
 35      <210> 323  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial  
 40      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 40      <400> 323  
 Trp Gln Gly Thr His Phe Pro Gln Thr  
 1                5  
 45      <210> 324  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial  
 50      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 50      <400> 324  
 Gln Gln Tyr Tyr Ser Tyr Pro Arg Thr  
 1                5  
 55      <210> 325  
 <211> 9  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

5      <400> 325  
      Gln Gln Ser Asn Glu Asp Pro Arg Thr  
      1                5

10     <210> 326  
      <211> 9  
      <212> PRT  
      <213> Secuencia Artificial

15     <220>  
      <223> Descripción de la Secuencia Artificial: Péptido sintético

20     15    <400> 326  
      Gln Gln Trp Ser Ser Asn Pro Leu Thr  
      1                5

25     <210> 327  
      <211> 9  
      <212> PRT  
      <213> Secuencia Artificial

30     <220>  
      <223> Descripción de la Secuencia Artificial: Péptido sintético

35     25    <400> 327  
      Gln His Phe Trp Gly Thr Pro Trp Thr  
      1                5

40     <210> 328  
      <211> 9  
      <212> PRT  
      <213> Secuencia Artificial

45     <220>  
      <223> Descripción de la Secuencia Artificial: Péptido sintético

50     <400> 328  
      Gln Gln Arg Ser Ser Tyr Pro Phe Thr  
      1                5

55     <210> 329  
      <211> 9  
      <212> PRT  
      <213> Secuencia Artificial

60     <220>  
      <223> Descripción de la Secuencia Artificial: Péptido sintético

65     <400> 329  
      Leu Gln Tyr Ala Ser Tyr Pro Trp Thr  
      1                5

70     <210> 330  
      <211> 9  
      <212> PRT  
      <213> Secuencia Artificial

75     <220>  
      <223> Descripción de la Secuencia Artificial: Péptido sintético

80     <400> 330  
      Gln Gln Ser Lys Glu Val Pro Leu Thr  
      1                5

5        <210> 331  
 <211> 19  
 <212> PRT  
 <213> Secuencia Artificial  
  
 10      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 10      <400> 331  
 Ile Pro Ser Tyr Ala Ser Ser Arg Gly Tyr Tyr Leu Ile Pro Asp Arg  
 1                    5                    10                    15  
  
 Leu Asp Leu  
  
 15      <210> 332  
 <211> 19  
 <212> PRT  
 <213> Secuencia Artificial  
  
 15      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 20      <400> 332  
 Ile Pro Ser Tyr Val Ser Gly Arg Gly Val Tyr Ile Ile Pro Asp Arg  
 1                    5                    10                    15  
  
 Phe Asp Leu  
  
 25      <210> 333  
 <211> 19  
 <212> PRT  
 <213> Secuencia Artificial  
  
 25      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 30      <400> 333  
 Ala Ser Asp Tyr Asp Ser Ser Arg Gly His Trp Leu Val Tyr Asn Arg  
 1                    5                    10                    15  
  
 Leu Asp Leu  
  
 35      <210> 334  
 <211> 14  
 <212> PRT  
 <213> Secuencia Artificial  
  
 40      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 40      <400> 334  
 Lys Gly Asp Pro Gly His Pro Asn Gly Leu Phe Phe Thr Met  
 1                    5                    10  
  
 45      <210> 335  
 <211> 19  
 <212> PRT  
 <213> Secuencia Artificial  
  
 50      <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
  
 <400> 335

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

**Leu Asp Leu**

<210> 336

<211> 19

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

10

<400> 336

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

**Leu Asp Leu**

<210> 337

<211> 19

15 <212> PRT

<213> Secuencia Artificial

<220>

20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 337

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

**Leu Asp Leu**

25 <210> 338

<211> 19

<212> PRT

<213> Secuencia Artificial

30 <220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 338

Ile Ser Ser Tyr Val Ser Ser Arg Gly Tyr Trp Leu Ile Pro Asp Gly  
 1               5               10               15

**Leu Asp Leu**

35

<210> 339

<211> 19

<212> PRT

<213> Secuencia Artificial

40

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 339

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

45 **Leu Asp Leu**

<210> 340

<211> 10  
<212> PRT  
<213> Secuencia Artificial

5 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 340  
Leu Tyr Asn Ser Val Val Gly Asp Asp Met  
1               5                           10

10 <210> 341  
<211> 19  
<212> PRT  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 341  
Ala Ser Asp Tyr Asp Ser Ser Arg Gly His Trp Leu Val Tyr Asp Arg  
1               5                       10                           15

20 Leu Asp Leu

<210> 342  
<211> 10  
<212> PRT

25 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

30 <400> 342  
Leu Tyr Asn Ser Val Val Gly Asp Asp Ile  
1               5                           10

<210> 343  
<211> 10

35 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

40 <400> 343  
Leu Tyr Asn Ser Leu Val Gly Asp Asp Ile  
1               5                           10

<210> 344  
<211> 12

45 <212> PRT  
<213> Secuencia Artificial

<220>

50 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 344  
Gly Met Pro Gly Ser Thr Ser Gly Asn Ser Asn Ile  
1               5                           10

55 <210> 345  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

5      <400> 345  
 Gly Met Pro Ala Ser Thr Ser Gly Asn Ser Asn Ile  
 1                5    10

10     <210> 346  
 <211> 12  
 <212> PRT  
 10    <213> Secuencia Artificial

15     <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

15     <400> 346  
 Gly Met Pro Gly Ser Thr Ser Gly Asn Ser Asn Ile  
 1                5    10

20     <210> 347  
 <211> 9  
 20    <212> PRT  
 20    <213> Secuencia Artificial

25     <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

25     <400> 347  
 Gly Val Pro Thr Asn Arg Asp Ala Met  
 1                5

30     <210> 348  
 <211> 12  
 <212> PRT  
 30    <213> Secuencia Artificial

35     <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

35     <400> 348  
 Ala Gly Gly Tyr Lys Ser Ser Gly Asp Thr Val Ser  
 1                5    10

40     <210> 349  
 <211> 13  
 <212> PRT  
 40    <213> Secuencia Artificial

45     <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45     <400> 349  
 Gln Gly Glu Phe Ser Cys Arg Asp Phe Asp Cys Thr Val  
 1                5    10

50     <210> 350  
 <211> 12  
 <212> PRT  
 50    <213> Secuencia Artificial

55     <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

55     <400> 350  
 Leu Gly Gly Tyr Lys Thr Thr Asp Gly Ser Ile  
 1                5    10

5 <210> 351  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

10	<400> 351	
	Gln Ser Tyr Tyr His Asn Ser Gly Thr Ser Tyr Ile Thr	
1	5	10

15 <210> 352  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

20	<400> 352	
	Ala Gly Gly Tyr Lys Ser Thr Thr Asp Gly Ser Ala	
1	5	10

25 <210> 353  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

30	<400> 353	
	Gln Ser Tyr Tyr His Asn Ser Gly Asn Ser Tyr Ile Thr	
1	5	10

35 <210> 354  
 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial

40 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

40	<400> 354	
	Gln Ser Tyr Tyr Tyr Gly Ser Gly Thr Ser Tyr Ile Thr	
1	5	10

45 <210> 355  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

50 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

50	<400> 355	
	Ala Gly Gly Tyr Lys Ser Ser Gly Asp Thr Phe Thr	
1	5	10

55 <210> 356  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

5      <400> 356  
Leu Gly Gly Tyr Lys Lys Thr Ile Asp Gly Ser Ala  
1                5                            10

10     <210> 357  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

15     <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

15     <400> 357  
Ala Gly Gly Tyr Lys Ser Ala Ser Asp Gly Ser Ala  
1                5                            10

20     <210> 358  
<211> 13  
<212> PRT  
<213> Secuencia Artificial

25     <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

25     <400> 358  
Gln Gly Glu Phe Ser Cys Asp Ala Gly Val Cys Thr Leu  
1                5                            10

30     <210> 359  
<211> 13  
<212> PRT  
<213> Secuencia Artificial

35     <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

35     <400> 359  
Gln Gly Glu Phe Ser Cys Arg Ser Tyr Asp Cys Thr Val  
1                5                            10

40     <210> 360  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

45     <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

45     <400> 360  
Leu Gln Asp Trp Ser Pro Ser Tyr Ala Asp Val Ala  
1                5                            10

50     <210> 361  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

55     <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

55     <400> 361  
Gln Gln Gly Arg Arg Ser Val Asp Val Asp Asn Val  
1                5                            10

5 <210> 362  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

10	<400> 362
	Ala Gly Gly Tyr Lys Thr Thr Thr Asp Gly Ser Ile
	1                   5                   10

15 <210> 363  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

20	<400> 363
	Gln Gln Gly Tyr Thr Ser Asn Val Asp Asn Val
	1                   5                   10

25 <210> 364  
 <211> 125  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 364  
 Glu Gln Leu Lys Glu Ser Gly Gly Leu Val Thr Pro Gly Gly Thr  
 1                   5                   10                   15

Leu Thr Leu Thr Cys Thr Val Ser Gly Phe Ser Leu Ser Ser Val Ala  
 20                   25                   30

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

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

Arg Phe Thr Val Ser Lys Thr Ser Thr Thr Val Asp Leu Lys Met Thr  
 65                   70                   75                   80

Ser Leu Thr Thr Glu Asp Thr Ala Thr Tyr Phe Cys Ala Arg Ile Pro  
 85                   90                   95

Ser Tyr Val Ser Gly Arg Gly Val Tyr Ile Ile Pro Asp Arg Phe Asp  
 100                105                110

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

35 <210> 365  
 <211> 116  
 <212> PRT  
 <213> Secuencia Artificial

ES 2 666 301 T3

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

5 <400> 365

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

Ala Val Gly Gly Thr Val Thr Ile Ser Cys Gln Ser Ser Glu Ser Val  
20 25 30

Tyr Lys Asn Asn Tyr Leu Ser Trp Tyr Gln Gln Lys Pro Gly His Ser  
35 40 45

Pro Arg Leu Leu Ile Ser Ser Ala Ser Thr Leu Ala Ser Gly Val Pro  
50 55 60

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

Ser Asp Leu Gln Cys Asp Asp Ala Ala Thr Tyr Tyr Cys Ala Gly Gly  
85 90 95

Tyr Lys Ser Ala Ser Asp Gly Ser Ala Phe Gly Gly Thr Glu Val  
100 105 110

Val Val Lys Gly  
115

<210> 366

<211> 116

10 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

15 <400> 366

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

Ala Val Gly Gly Thr Val Thr Ile Ser Cys Gln Ser Ser Glu Ser Val  
20 25 30

Tyr Lys Asn Asn Tyr Leu Ser Trp Tyr Gln Gln Lys Pro Gly His Ser  
35 40 45

Pro Arg Leu Leu Ile Ser Ser Ala Ser Thr Leu Ala Ser Gly Val Pro  
50 55 60

ES 2 666 301 T3

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

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

Tyr Lys Thr Thr Asp Gly Ser Ile Phe Gly Gly Thr Glu Val  
100 105 110

Val Val Lys Gly  
115

<210> 367

<211> 120

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

10

<400> 367

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

Leu Thr Leu Thr Cys Thr Val Phe Gly Phe Asp Val Ser Ser His Ile  
20 25 30

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

Leu Val Asp Ile Gly Lys Ser Ile Lys Trp Tyr Ala Ser Trp Ala Lys  
50 55 60

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

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

Phe Ala Leu Trp Gly Pro Gly Thr Leu Val Thr Val Ser Ser Gly Gln  
100 105 110

Pro Lys Ala Pro Ser Val Phe Pro  
115 120

<210> 368

<211> 120

15 <212> PRT

<213> Secuencia Artificial

<220>

20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 368

ES 2 666 301 T3

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

Gly Thr Val Thr Ile Asn Cys Gln Ala Ser Pro Ser Val Tyr Gly Asn  
20 25 30

Tyr Leu Ser Trp Phe Gln Gln Lys Pro Gly Gln Pro Pro Lys Leu Leu  
35 40 45

Ile Leu Asn Ala Ser Thr Leu Pro Ser Gly Val Ser Ser Arg Phe Lys  
50 55 60

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

Cys Asp Asp Ala Ala Thr Tyr Tyr Cys Leu Gly Ser Tyr Asp Cys Arg  
85 90 95

Ser Asp Asp Cys Asn Val Phe Gly Gly Thr Glu Val Val Val Lys  
100 105 110

Gly Asp Pro Val Ala Pro Thr Val  
115 120

<210> 369

<211> 14

5 <212> PRT

<213> Secuencia Artificial

<220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

10

<400> 369

Ala Gly Thr Asn Tyr Trp Ala Ile Tyr Tyr Gly Met Asp Leu

1 5 10

<210> 370

15 <211> 12

<212> PRT

<213> Secuencia Artificial

<220>

20 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 370

Gln Gln Gly Tyr Lys Ile Thr Asn Ile Glu Asn Val

1 5 10

25 <210> 371

<211> 9

<212> PRT

<213> Secuencia Artificial

30 <220>

<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 371

Cys Ala Arg Gly Ser Ala Phe Ala Tyr

1 5

35 <210> 372

<211> 14

<212> PRT  
 <213> Secuencia Artificial  
 <220>  
 5 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 372  
 Cys Ala Arg Arg Gly Tyr Asp Gly Ser Tyr Tyr Phe Asp Tyr  
 1 5 10  
 10 <210> 373  
 <211> 12  
 <212> PRT  
 <213> Secuencia Artificial  
 15 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 373  
 Cys Ala Arg His Glu Pro Leu Asn Trp Phe Pro Tyr  
 1 5 10  
 20 <210> 374  
 <211> 14  
 <212> PRT  
 <213> Secuencia Artificial  
 25 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 <400> 374  
 Cys Ala Arg Arg Gly Tyr Tyr Ser Asn Thr Tyr Val Asp Tyr  
 30 1 5 10  
 <210> 375  
 <211> 14  
 <212> PRT  
 35 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 40 <400> 375  
 Cys Ala Arg Arg Gly Tyr Tyr Ala Asp Thr Tyr Val Asp Tyr  
 1 5 10  
 <210> 376  
 <211> 14  
 45 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético  
 50 <400> 376  
 Cys Ala Arg Arg Gly Tyr Tyr Ser Gly Thr Tyr Val Asp Tyr  
 1 5 10  
 <210> 377  
 55 <211> 13  
 <212> PRT  
 <213> Secuencia Artificial  
 <220>  
 60 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 377  
 Cys Ala Arg Glu Asp Tyr Tyr Ser Asp Gln Phe Ala Tyr  
 1 5 10

5 <210> 378  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial

10 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 378  
 Cys Leu Tyr Lys Tyr Gly Gly Gly Phe Asp Tyr  
 1 5 10

15 <210> 379  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial

20 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 379  
 Cys Ala Arg Ser Arg Thr Gly Ile Phe Asp Tyr  
 1 5 10

25 <210> 380  
 <211> 11  
 <212> PRT  
 <213> Secuencia Artificial

30 <220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 380  
 Cys Val Arg Ser Arg Ser Gly Ile Phe Asp Tyr  
 1 5 10

35 <210> 381  
 <211> 13  
 <212> PRT  
 40 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

45 <400> 381  
 Cys Ala Arg Tyr Tyr Arg Asn Tyr Gly Gly Phe Asp Tyr  
 1 5 10

<210> 382  
 <211> 10

50 <212> PRT  
 <213> Secuencia Artificial

<220>  
 <223> Descripción de la Secuencia Artificial: Péptido sintético

55 <400> 382  
 Cys Ala Arg Leu Thr Ala Tyr Phe Asp Tyr  
 1 5 10

<210> 383  
 60 <211> 9

<212> PRT  
<213> Secuencia Artificial

5 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 383  
Phe Gln Gly Ser His Val Pro Phe Thr  
1 5

10 <210> 384  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 384  
Leu Gln Tyr Ala Ser Tyr Pro Trp Thr  
1 5

20 <210> 385  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 385  
Gln Gln His Phe Ser Thr Pro Pro Thr  
1 5

30 <210> 386  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 386  
Gln Gln His Tyr Ser Thr Pro Phe Thr  
1 5

40 <210> 387  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

45 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 387  
Gln Gln Trp Ser Ser Asn Pro Leu Thr  
1 5

50 <210> 388  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

55 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 388  
Gln Gln His Tyr Ser Thr Pro Pro Thr  
1 5

5 <210> 389  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

<220>  
10 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 389  
Gln Gln Arg Ser Ser Tyr Pro Phe Thr  
1 5

15 <210> 390  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 390  
Gln Gln Ser Lys Glu Val Pro Leu Thr  
1 5

25 <210> 391  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

30 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 391  
Gln Gln Ser Asn Glu Asp Pro Arg Thr  
1 5

35 <210> 392  
<211> 9  
<212> PRT  
40 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

45 <400> 392  
Gln His Phe Trp Gly Thr Pro Trp Thr  
1 5

<210> 393  
<211> 9

50 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

55 <400> 393  
Gln Gln Tyr Tyr Ser Tyr Pro Arg Thr  
1 5

<210> 394  
60 <211> 9

<212> PRT  
<213> Secuencia Artificial

5 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 394  
Trp Gln Gly Thr His Phe Pro Gln Thr  
1 5

10 <210> 395  
<211> 9  
<212> PRT  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 395  
Ser Gln Ser Thr His Val Pro Trp Thr  
1 5

20 <210> 396  
<211> 4  
<212> PRT  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 396  
Lys Leu Gly Leu  
1

30 <210> 397  
<211> 4  
<212> PRT  
<213> Secuencia Artificial

35 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

40 <400> 397  
Gly Phe Ser Leu  
1

<210> 398  
<211> 5

45 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

50 <400> 398  
Asp Leu Gly Asp Leu  
1 5

<210> 399  
<211> 5

55 <212> PRT  
<213> Secuencia Artificial

<220>

60 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 399  
Asp Leu Gly Asn Leu  
1 5

5 <210> 400  
<211> 4  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 400  
Asp Phe His Leu  
1

15 <210> 401  
<211> 8  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 401  
Gly Ser Leu Gly Thr Leu Pro Leu  
1 5

25 <210> 402  
<211> 4  
<212> PRT  
<213> Secuencia Artificial

30 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 402  
Gly Phe Ala Leu  
1

35 <210> 403  
<211> 12  
<212> PRT  
<213> Secuencia Artificial

40 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 403  
Gly His Asp Asp Gly Tyr Asn Tyr Val Tyr Lys Leu  
1 5 10

45 <210> 404  
<211> 4  
<212> PRT  
<213> Secuencia Artificial

50 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 404  
Gly Phe Thr Leu  
1

55 <210> 405  
<211> 13

<212> PRT  
<213> Secuencia Artificial

5 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 405  
Leu Ala Gly Tyr Asp Cys Thr Thr Gly Asp Cys Phe Ala  
1                   5                   10

10 <210> 406  
<211> 13  
<212> PRT  
<213> Secuencia Artificial

15 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 406  
Leu Gly Gly Tyr Asp Cys Asp Asn Gly Asp Cys Phe Thr  
1                   5                   10

20 <210> 407  
<211> 13  
<212> PRT  
<213> Secuencia Artificial

25 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 407  
Leu Gly Thr Tyr Asp Cys Arg Arg Ala Asp Cys Asn Thr  
1                   5                   10

30 <210> 408  
<211> 12  
<212> PRT  
35 <213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

40 <400> 408  
Gln Ser Thr Leu Tyr Ser Ser Thr Asp Glu Ile Val  
1                   5                   10

<210> 409  
<211> 9  
45 <212> PRT  
<213> Secuencia Artificial

<220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

50 <400> 409  
Gln Cys Ser Tyr Val Asn Ser Asn Thr  
1                   5

<210> 410  
55 <211> 13  
<212> PRT  
<213> Secuencia Artificial

<220>

60 <223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 410  
Leu Gly Ser Tyr Asp Cys Arg Ser Asp Asp Cys Asn Val  
1 5 10

5 <210> 411  
<211> 11  
<212> PRT  
<213> Secuencia Artificial

10 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 411  
Leu Gly Ala Tyr Asp Asp Ala Ala Asp Asn Ser  
1 5 10

15 <210> 412  
<211> 13  
<212> PRT  
<213> Secuencia Artificial

20 <220>  
<223> Descripción de la Secuencia Artificial: Péptido sintético

<400> 412  
Leu Gly Thr Tyr Asp Cys Asn Ser Ala Asp Cys Asn Val  
1 5 10

25

## REIVINDICACIONES

1. Un método para obtener secuencias de nucleótidos o secuencias de aminoácidos de cadenas pesadas o ligeras de inmunoglobulinas que se unen específicamente a un antígeno, que comprende:

- 5       (a) proporcionar secuencias de ácido nucleico que codifican cadenas de inmunoglobulina o regiones variables de las mismas de al menos un animal de una especie seleccionada;
- 10      (b) obtener una población de inmunoglobulinas policlonales de un animal de dicha especie seleccionada de la etapa (a) que se une específicamente a un antígeno, y obtener información de los espectros de masas de fragmentos peptídicos de cadenas pesadas o cadenas ligeras o regiones variables de las mismas de dicha población;
- 15      (c) correlacionar la información de espectros de masas de los fragmentos peptídicos con la información pronosticada de los espectros de masas de las secuencias de ácidos nucleicos, dicha información pronosticada de los espectros de masas derivada de secuencias de aminoácidos pronosticadas codificadas por dichas secuencias de ácidos nucleicos, para identificar secuencias de nucleótidos o secuencias de aminoácidos de cadenas de inmunoglobulina (o regiones variables de las mismas) que comprenden los fragmentos peptídicos; y
- 20      (d) seleccionar entre las secuencias de nucleótidos o las secuencias de aminoácidos identificadas de cadenas de inmunoglobulina (o regiones variables de las mismas) basándose en la cobertura de la secuencia de aminoácidos de las cadenas de inmunoglobulina o fragmentos de las mismas por los fragmentos peptídicos, para obtener secuencias de nucleótidos o secuencias de aminoácidos de cadenas pesadas o ligeras de inmunoglobulinas que se unen específicamente al antígeno.

2. El método de la reivindicación 1, en donde dicho al menos un animal de la etapa (a) es un animal expuesto al antígeno y/u opcionalmente

en donde las secuencias de ácido nucleico son secuencias de ácidos nucleicos expresadas y/u opcionalmente

25      en donde las secuencias de ácido nucleico que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) se obtienen a partir de dicho al menos un animal mediante:

- (1) aislamiento de moléculas de ácido nucleico de glóbulos blancos de dicho al menos un animal; y
- (2) amplificación de la cadena de inmunoglobulina (o región variable de la misma) que codifica las moléculas de ácido nucleico utilizando cebadores específicos para secuencias de polinucleótido adyacentes a dicha cadena de inmunoglobulina (o región variable de la misma) que codifican moléculas de ácido nucleico, y
- 30      (3) obtención de secuencias de nucleótido de dichas moléculas de ácido nucleico amplificadas que codifican cadenas de inmunoglobulina (o regiones variables de las mismas) y/u opcionalmente

en donde las moléculas de ácido nucleico que codifican polipéptidos son moléculas de ARN y dicha etapa de amplificación incluye una etapa de transcripción inversa inicial y/u opcionalmente

35      en donde dichas secuencias de polinucleótido adyacentes a dichas moléculas de ácido nucleico que codifican la cadena de inmunoglobulina (o región variable de la misma) se seleccionan del grupo que consiste en genes de inmunoglobulina que flanquean ADN genómico, secuencias de polinucleótido que codifican regiones constantes de cadenas de inmunoglobulina y secuencias de polinucleótido que codifica regiones marco de cadenas de inmunoglobulina.

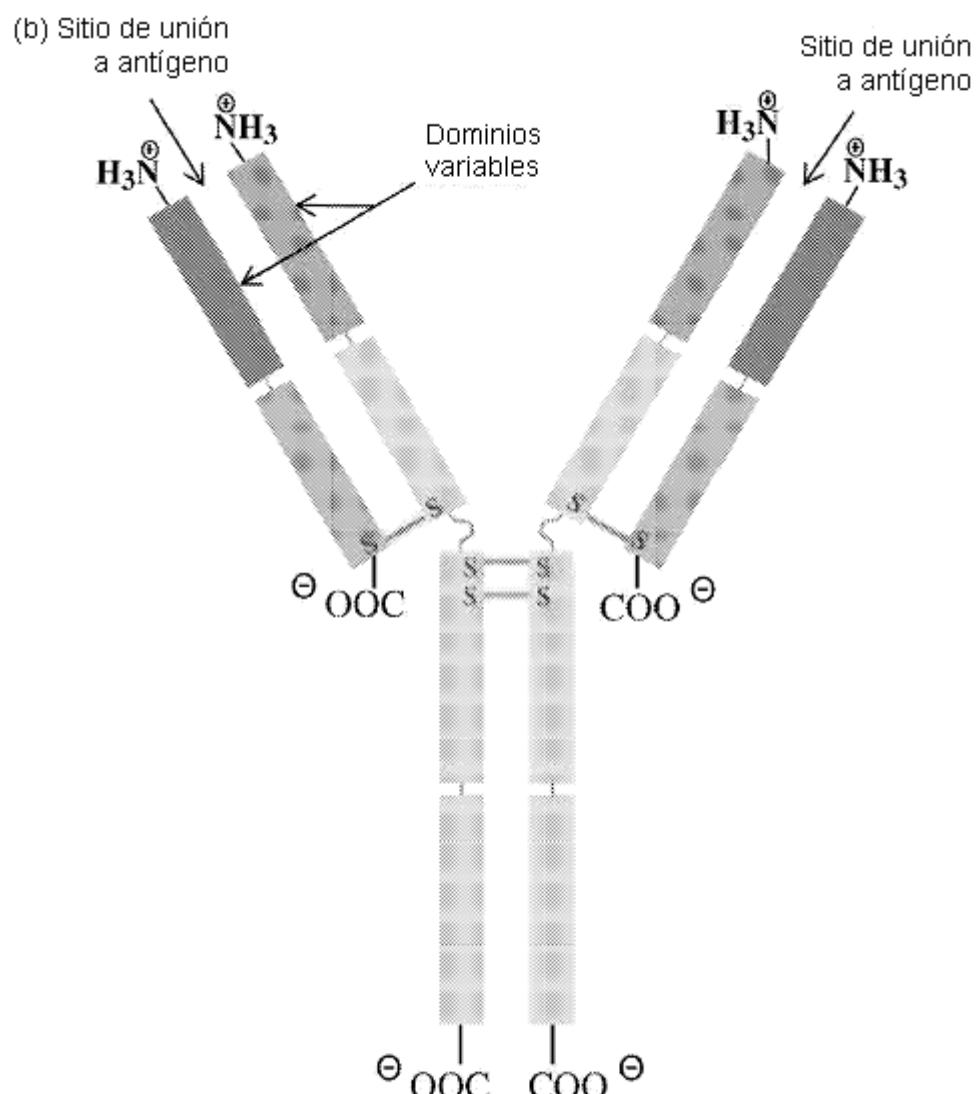
40      3. El método de la reivindicación 1, en donde la información del espectro de masas pronosticado se obtiene utilizando un método que comprende las etapas de:

- (i) realizar una digestión teórica de secuencias de aminoácidos pronosticadas codificadas por secuencias de nucleótidos de las secuencias de ácido nucleico con una o más proteasas y/o uno o más reactivos de escisión química de proteínas para generar fragmentos peptídicos virtuales; y
- (ii) crear espectros de masas pronosticados de dichos fragmentos peptídicos virtuales.

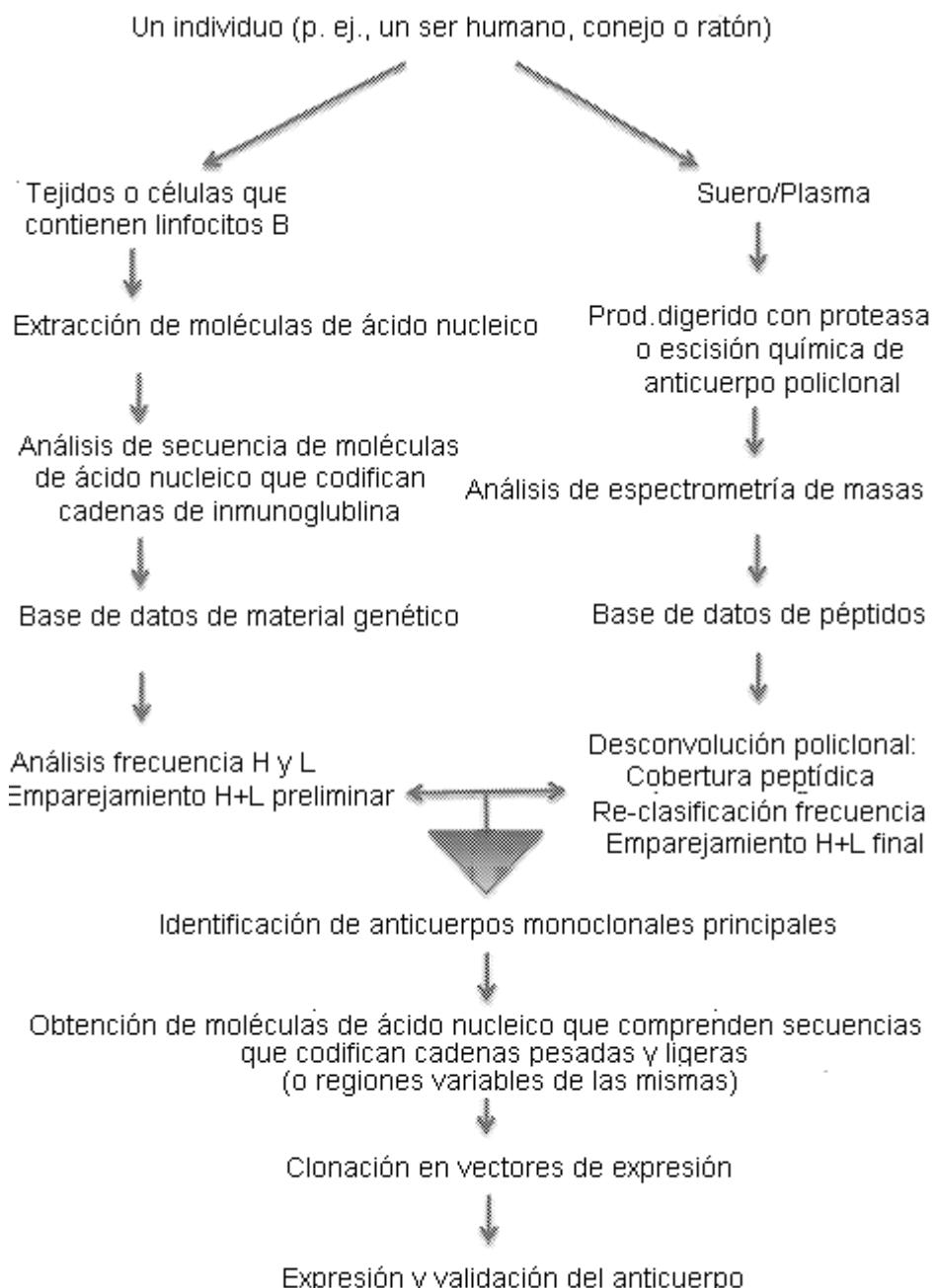
45      4. El método de la reivindicación 1, en donde las secuencias de ácido nucleico de la etapa (a), las secuencias de aminoácidos pronosticadas y los espectros de masas pronosticados derivados de dichas secuencias de ácido nucleico se ubican dentro de una base de datos de material genético.

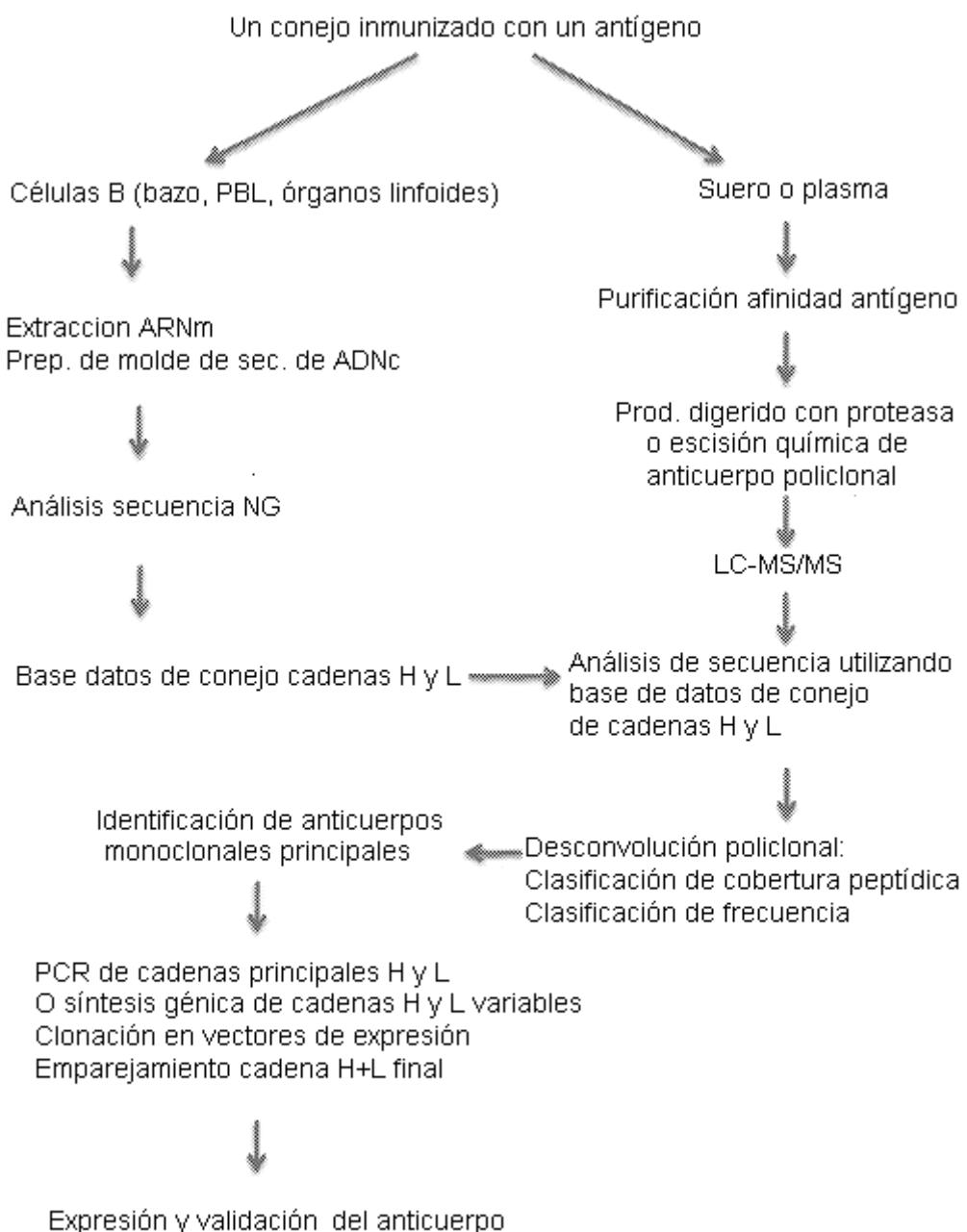
50      5. El método de la reivindicación 1, en donde dicha población polyclonal de inmunoglobulinas de la etapa (b) se obtiene de una muestra de fluido corporal o fracción de la misma de un animal y/u opcionalmente dicho fluido corporal se selecciona del grupo que consiste en sangre, fluido cerebroespinal, líquido sinovial, líquido peritoneal, secreciones de la mucosa, lágrimas, secreciones nasales, saliva, leche y secreciones genitourinarias.

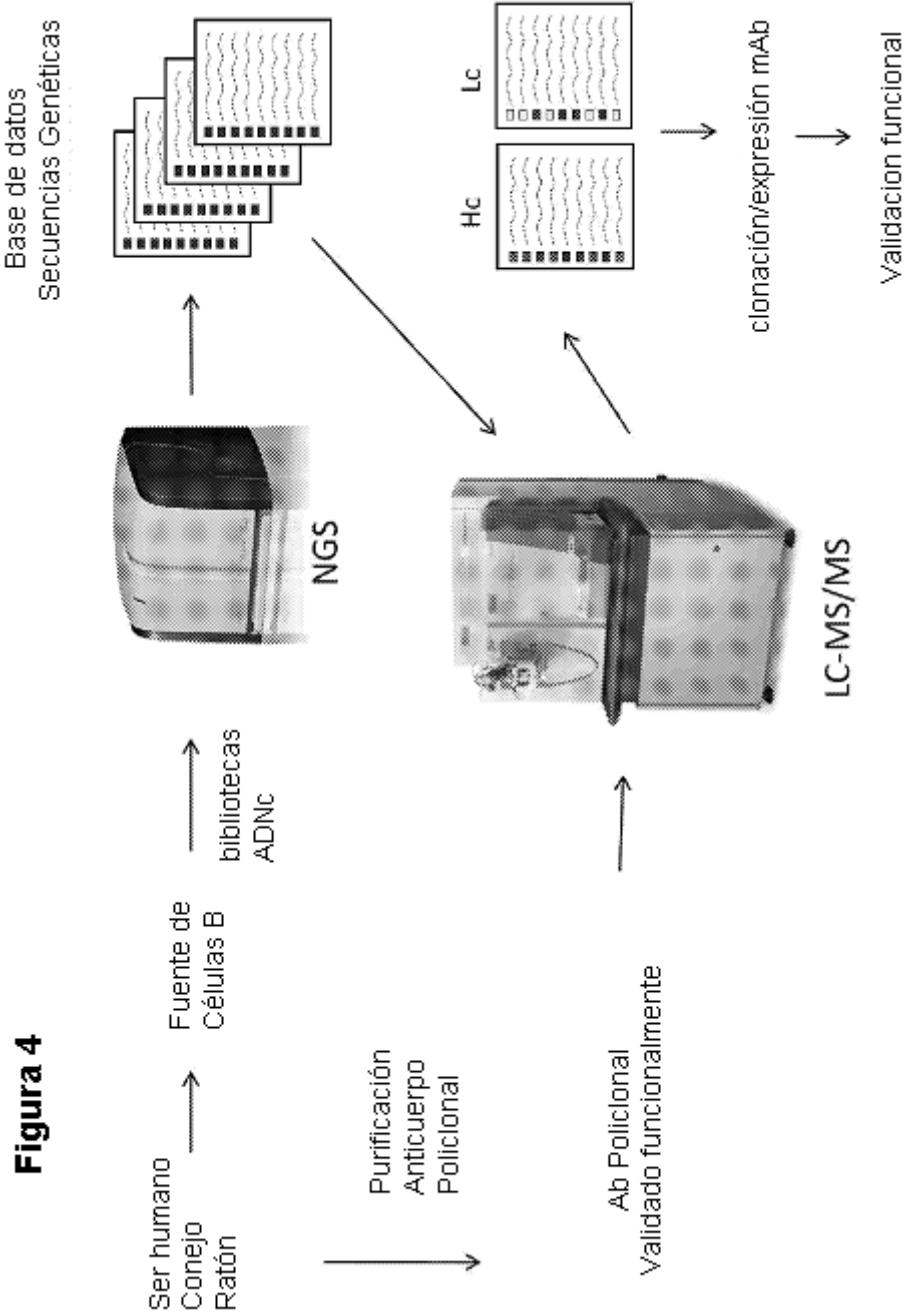
6. El método de la reivindicación 5, en donde el animal es un animal previamente expuesto al antígeno, y/u opcionalmente en donde el animal previamente expuesto al antígeno es un animal previamente inmunizado con el antígeno.
7. El método de la reivindicación 5, en donde el animal es el mismo que dicho al menos un animal de la etapa (a).
- 5 8. El método de la reivindicación 5, en donde dicha población policlonal de inmunoglobulinas de la etapa (b) se purifica de dicha muestra de fluido corporal o fracciones de la misma y/u opcionalmente en donde la purificación se logra por purificación con proteína A o proteína G o purificación por afinidad basada en el antígeno.
9. El método de la reivindicación 1, en donde dicha población policlonal de inmunoglobulinas de la etapa (b) se obtiene del medio de células cultivadas in vitro, y/u opcionalmente en donde la información del espectro de masas de los fragmentos peptídicos de la etapa (b) se obtiene de la población policlonal de inmunoglobulinas mediante la digestión de la población con una o más proteasas y/o uno o más reactivos de escisión química de proteínas para generar fragmentos, y someter los fragmentos generados a análisis de espectrometría de masas.
- 10 10. El método de la reivindicación 1, en donde la selección en la etapa (d) se basa en la cobertura de la secuencia de aminoácidos de las secuencias de CDR3 de las inmunoglobulinas por los fragmentos peptídicos.
- 15 11. El método de la reivindicación 1 o la reivindicación 10, en donde la selección en la etapa (d) se basa independiente o adicionalmente en la cobertura de la secuencia de aminoácidos de las regiones variables de las inmunoglobulinas por los fragmentos peptídicos.
- 20 12. Método de la reivindicación 1, en donde la selección en la etapa (d) se basa adicionalmente en al menos un parámetro seleccionado del grupo que consiste en el número de péptidos únicos mapeados, la compartición del espectro, el recuento de péptidos totales, el recuento de péptidos únicos, la frecuencia de las secuencias de ácidos nucleicos codificantes y la relación clonal.
13. El método de la reivindicación 1, en donde el animal es un ser humano.
14. El método de la reivindicación 1, en donde el animal es un conejo o un ratón.



**Figura 1**

**Figura 2**

**Figura 3**



**Cadenas pesadas**

ref	total_péptidos	cadena	cdr3_recuento	cdr3	SEQ ID NO:
G623FKB01BJ0L7	153	pesada	2	RGYYSGTYVVDY	307
<b>G623FKB01AF34N</b>	<b>148</b>	<b>pesada</b>	<b>244</b>	<b>SRTGIFDY</b>	<b>308</b>
<b>G623FKB01BSNEU</b>	<b>131</b>	<b>pesada</b>	<b>15</b>	<b>SRSGIFDY</b>	<b>309</b>
<b>G623FKB01A2ZY4</b>	<b>130</b>	<b>pesada</b>	<b>58</b>	<b>RGYYSNTYVVDY</b>	<b>310</b>
<b>G623FKB01A0MXT</b>	<b>126</b>	<b>pesada</b>	<b>68</b>	<b>LTAIFYFDY</b>	<b>311</b>
G623FKB01AXQN3	126	pesada	6	RFYDGSSYFFDY	312
G623FKB01AC4H8	122	pesada	19	RGYYADTYVVDY	313
<b>G623FKB01AB7YA</b>	<b>84</b>	<b>pesada</b>	<b>13</b>	<b>HEPLNNWFPY</b>	<b>314</b>
G623FKB01AUXQW	69	pesada	45	KYGGGFDY	315
G623FKB01BO141	52	pesada	8	GSFAFAY	316
<b>G623FKB01BBIXE</b>	<b>49</b>	<b>pesada</b>	<b>5</b>	<b>YRNYGGGFDY</b>	<b>317</b>
G623FKB01AL5DX	43	pesada	15	EDYYSDQFAY	318

**Cadenas ligeras**

ref	total_péptidos	cadena	NGS_recuento	cdr3	SEQ ID NO:
G623FKB01A3GC7	235	Ligera	8	SQSTTHVPWWT	319
G623FKB01AXJ1C	199	Ligera	1068	QQHYSTPPPT	320
G623FKB01ATC8J	153	Ligera	112	QQHFSTPPPT	321
G623FKB01BCUK2	132	Ligera	597	QQHYSTPFT	322
<b>G623FKB01BY43F</b>	<b>74</b>	<b>Ligera</b>	<b>532</b>	<b>WQGTHFPQT</b>	<b>323</b>
<b>G623FKB01BAZ8V</b>	<b>64</b>	<b>Ligera</b>	<b>6</b>	<b>QQYYSPRT</b>	<b>324</b>
<b>G623FKB01A580V</b>	<b>39</b>	<b>Ligera</b>	<b>20</b>	<b>QQSNEPDPT</b>	<b>325</b>
G623FKB01A312C	39	Ligera	384	QQWSSNPLT	326
<b>G623FKB01BADIC</b>	<b>31</b>	<b>Ligera</b>	<b>258</b>	<b>QHFWGTPWT</b>	<b>327</b>
G623FKB01ADDWX	31	Ligera	219	QQRSSYPFT	328
G623FKB01A2982	23	Ligera	891	LOYASYPWT	329
<b>G623FKB01BPZWW</b>	<b>38</b>	<b>Ligera</b>	<b>19</b>	<b>QQSKEVPLT</b>	<b>330</b>

*Cursiva negrita:* cadenas de anticuerpo que son reactivas con el antígeno

**Figura 5**

ID LC	ID Hc	BSNEU		BO141		AOMXT		BBIXE		AXQN3		AB77A	
		A	B	A	B	A	B	A	B	A	B	A	B
BY43F	A	0.1991	0.229	0.0995	0.0932	0.0921	0.0895	0.1193	0.102	0.1548	0.1429	0.1796	0.1863
	B	0.9938	1.2748	0.2205	0.1463	0.1588	0.1899	0.5913	2.3349	0.0819	0.183	0.1681	0.1131
A3GC7	A	0.6476	0.4431	0.0954	0.0882	0.155	0.1128	0.1514	0.1829	0.1717	0.1595	0.3717	0.192
	B	0.571	0.5998	0.2424	0.0955	0.1599	0.1177	0.1499	0.1755	0.1035	0.0986	0.6849	0.1495
BAZ8V	A	0.6732	0.7344	0.0882	0.0744	0.1114	0.1048	0.4041	0.8467	0.1023	0.0671	0.1645	0.1409
	B	0.7187	0.6232	0.1125	0.1171	0.1907	0.1074	0.1949	0.2536	0.1141	0.0937	0.1411	0.0746
A312C	A	0.0961	0.0897	0.0814	0.0959	0.0869	0.0849	0.2561	0.1673	0.1375	0.1143	0.1255	0.1243
	B	0.1122	0.0985	0.2757	0.1145	0.1526	0.1026	0.2335	0.1849	0.1215	0.1104	0.1061	0.1562
placa 3													
A580V	A	3.47	3.0952	0.0851	0.0929	1.2682	1.7622	0.2069	0.2057	0.1099	0.1162	0.1188	0.175
	B	1.5397	1.3374	0.085	0.107	1.0411	1.8156	0.2299	0.2735	0.0855	0.1234	0.0958	0.1281
BP2WW	A	0.1257	0.1044	0.1008	0.1091	0.0894	0.0842	1.1936	1.2218	0.1372	0.1115	0.1123	0.124
	B	0.162	0.1011	0.2664	0.1034	0.176	0.1706	1.4229	0.9012	0.0883	0.1091	0.0803	0.1166
ADDWX	A	0.092	0.0975	0.0775	0.1021	0.2347	0.1038	0.3545	0.2561	0.1478	0.1201	0.1084	0.1312
	B	0.1022	0.0911	0.0684	0.0896	0.1448	0.1014	0.2455	0.5667	0.1025	0.1033	0.1026	0.0754
AXJ1C	A	0.0929	0.0835	0.0838	0.0994	0.0983	0.1008	0.2138	0.3213	0.1731	0.1144	0.1113	0.1104
	B	0.104	0.1181	0.1993	0.2078	0.1369	0.3932	0.1526	0.3997	0.2355	0.1435	0.1263	0.1915
placa 5													
BCUK2	A	0.1451	0.1077	0.1041	0.1548	0.1076	0.1561	0.353	0.2947	0.1246	0.1243	0.1462	0.1731
	B	0.1074	0.0872	0.1016	0.0952	0.1507	0.1442	0.3181	0.3213	0.1058	0.1341	0.0987	0.1175
ATC8J	A	0.0832	0.0987	0.0996	0.1377	0.1399	0.1315	0.3191	0.2757	0.1496	0.096	0.1022	0.1378
	B	0.1029	0.0872	0.2011	0.0915	0.1117	0.1118	0.2807	0.2755	0.1078	0.0796	0.0789	0.1487
BADIC	A	3.6013	2.0746	0.0854	0.1078	3.0399	2.3754	0.187	0.3191	0.1331	0.0846	0.1008	0.1468
	B	1.0088	1.6872	0.0883	0.1117	1.7897	1.2741	0.1918	0.2086	0.0892	0.1111	0.1075	0.0772
A2982	A	0.1187	0.1052	0.1109	0.1093	0.1003	0.0997	0.2348	0.4317	0.139	0.1361	0.12	0.0998
	B	0.1547	0.0901	0.112	0.1175	0.1186	0.1065	0.1875	0.3154	0.1201	0.1146	0.1422	0.1878
placa 7													
AS297	A	0.1615	0.1045	0.121	0.1232	0.1307	0.1299	0.4285	0.4476	0.143	0.1361	0.169	0.1675
	B	0.1197	0.084	0.0982	0.101	0.1362	0.1288	0.3966	0.3443	0.1103	0.1664	0.1162	0.1344
9106													
Control	1.6003	1.0512	1.4815	1.4683									
Zario													
solo	0.1702	0.1141	0.1707	0.0959	0.1136	0.1081	0.0882	0.0993	0.0822	0.1045	0.0924	0.1662	

**Figura 6**

ID LC	ID Hc	Placa 1		A22Y4		AC4H8		BJOL7		AL5DX		AUJXQW		AF34N	
		A	B	A	B	A	B	A	B	A	B	A	B	A	B
BY43F	A	0.1422	0.0906	0.0956	0.0965	0.1142	0.0951	0.125	0.1081	0.124	0.1346	0.1486	0.148		
	B	0.3013	0.6843	0.1651	0.1841	0.124	0.1012	0.1195	0.0961	0.09B2	0.1439	0.1278	0.1596		
A3GC7	A	0.0958	0.194	0.2652	0.4371	0.4326	0.4052	0.2398	0.233	0.1992	0.1144	0.2087	0.1812		
	B	0.0953	0.089	0.5074	0.4602	0.44	0.4557	0.2844	0.2952	0.3203	0.1075	0.1821	0.1748		
BAZ8V	A	0.0966	0.084	0.2187	0.3583	0.5084	0.3416	0.4487	0.4084	0.3246	0.1517	0.1331	0.1267		
	B	0.0937	0.0759	0.1215	0.3744	0.4007	0.3472	0.3713	0.4181	0.2366	0.0809	0.1075	0.0982		
A312C	A	0.0973	0.0973	0.1037	0.0944	0.1731	0.1139	0.2151	0.1307	0.1041	0.1071	0.102	0.1016		
	B	0.0899	0.1232	0.194	0.1527	0.1216	0.1237	0.1262	0.0902	0.1538	0.1891	0.1267	0.145		
Placa 3															
A580V	A	0.1256	0.1403	0.1009	0.1162	0.1183	0.0843	0.091	0.0967	0.1121	0.1369	1.6113	4.0514		
	B	0.0965	0.0923	0.0925	0.1119	0.151	0.116	0.1443	0.095	0.078	0.1416	4.4636	4.7646		
BPZWW	A	0.1009	0.084	0.1061	0.1346	0.1929	0.1668	0.1219	0.1021	0.147	0.1143	0.1262	0.1294		
	B	0.1098	0.1563	0.1862	0.1062	0.1875	0.1935	0.1004	0.1742	0.1604	0.1031	0.0813	0.1238		
ADDWX	A	0.0917	0.0898	0.0737	0.1389	0.2034	0.1038	0.1136	0.1091	0.1574	0.08	0.1027	0.1345		
	B	0.0941	0.0766	0.0868	0.1144	0.1279	0.0779	0.1884	0.1056	0.0974	0.0851	0.1016	0.0867		
AJX1C	A	0.0996	0.0985	0.088	0.1245	0.1463	0.1464	0.153	0.1366	0.1798	0.1065	0.1297	0.1081		
	B	0.1232	0.0857	0.127	0.1102	0.1855	0.139	0.1621	0.1028	0.1296	0.1331	0.1085	0.1783		
Placa 5															
BCUJK2	A	0.1191	0.1299	0.1049	0.0994	0.1008	0.1397	0.1856	0.1667	0.1412	0.1217	0.1951	0.1944		
	B	0.1017	0.0901	0.0908	0.09	0.1329	0.1093	0.1191	0.1134	0.0902	0.1369	0.1233	0.1771		
ATC8J	A	0.1139	0.1023	0.1015	0.1093	0.1077	0.1066	0.0904	0.1016	0.104	0.1135	0.1038	0.1455		
	B	0.1116	0.0864	0.1698	0.0999	0.1153	0.1032	0.0848	0.1091	0.0921	0.0871	0.0887	0.1607		
BADIC	A	0.1118	0.0916	0.0875	0.0922	0.0869	0.0862	0.0886	0.1082	0.1101	0.0988	4.4611	4.4264		
	B	0.1211	0.1007	0.1083	0.0994	0.1223	0.1087	0.0998	0.1298	0.1193	0.11	3.7921	4.049		
A2982	A	0.108	0.0961	0.0886	0.1053	0.1523	0.1132	0.1099	0.1466	0.2152	0.1184	0.1058	0.1696		
	B	0.1896	0.1141	0.1493	0.1354	0.1549	0.1358	0.1705	0.1357	0.193	0.1744	0.1836	0.3142		
Placa 7															
AS297	A	0.2515	0.0969	0.1444	0.1882	0.1441	0.1392	0.1377	0.1882	0.2288	0.1873	0.2154	0.2061		
	B	0.1415	0.11	0.1125	0.1112	0.1225	0.1437	0.1244	0.1287	0.1193	0.1486	0.1215	0.1323		
9106	Control	1.6764	1.4255	1.9828	1.6362										
	Zario solo	0.1395	0.109	0.256	0.1528	0.1541	0.1128	0.1023	0.1168	0.0921	0.1133	0.1056	0.1895		

Figura 6 continuación

RT-PCR de Esplenocitos de Ratón

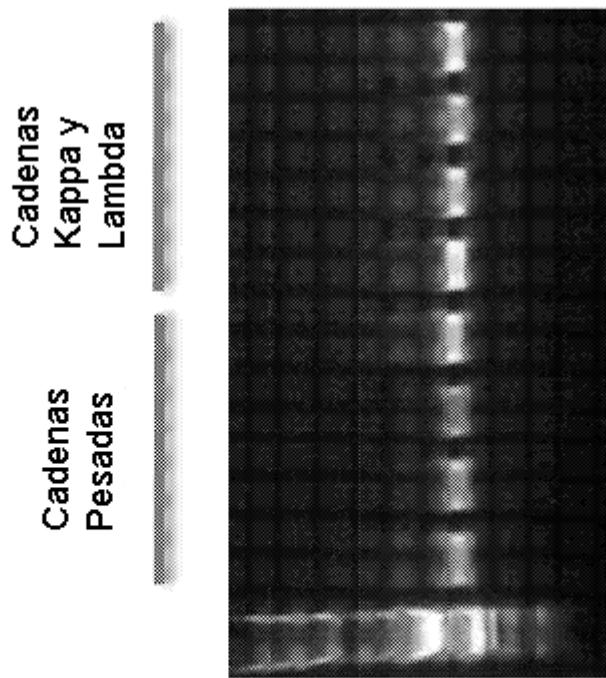


Figura 7

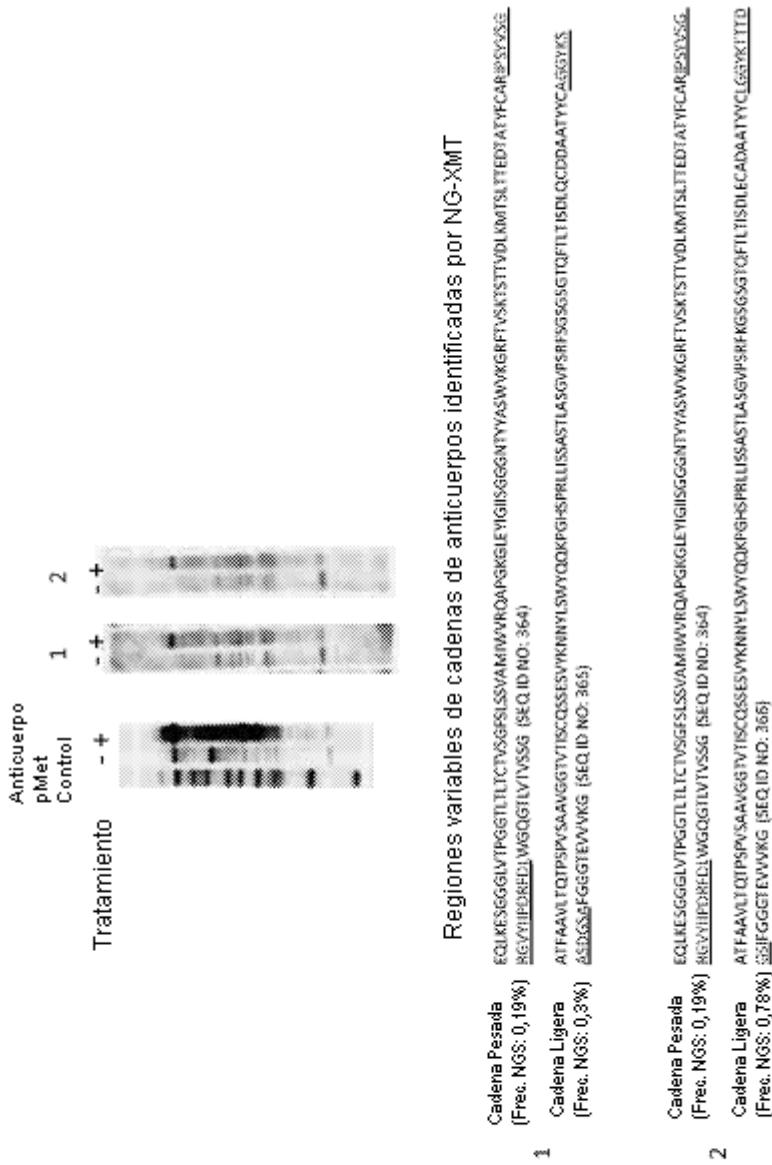
3126(pMet)				
ID Secuencia NGS	Cadena	Secuencia CDR3	SEQ ID NO:	Frec NGS (max 9,12%)
GXRYQP201BFGR	cadena pesada	IPSYASSRGYYLIPDRLDL	331	0.03%
<b>GXRYQP201BIQD2</b>	<b>cadena pesada</b>	<b>IPSYVSGRGVYIIPDRFDL</b>	<b>332</b>	<b>0.19%</b>
GXRYQP201AGL7B	cadena pesada	ASDYDSSRGHVLVYNRLDL	333	0.08%
<b>GXRYQP201A97DZ</b>	<b>cadena pesada</b>	<b>KGDPGHPNGLFFT</b>	<b>334</b>	<b>0.05%</b>
GXRYQP201ARCKH	cadena pesada	IPSYVSSRGYYLIPDGLDL	335	0.02%
GXRYQP201AYV7	cadena pesada	IPSYVSSRGYYLVPDGLDL	336	0.01%
GXRYQP201BRYT5	cadena pesada	IPSYVSSRGYYLIPDRLDL	337	0.05%
GXRYQP201AKBWL	cadena pesada	ISSYVSSRGYWLIIPDGLDL	338	0.13%
GXRYQP201ANGRW	cadena pesada	ISSYVSSRGYYLIPDGLDL	339	0.01%
GXRYQP201BGVA8	cadena pesada	LYNSVVGDDM	340	0.55%
GXRYQP201AGFK4	cadena pesada	ASDYDSSRGHVLVYDRLDL	341	0.01%
GXRYQP201BJROT	cadena pesada	LYNSVVGDDI	342	0.03%
GXRYQP201A8DBE	cadena pesada	LYNSLVGDDI	343	0.22%
GXRYQP201B2QN3	cadena pesada	GMPG5TSGNSNI	344	0.59%
GXRYQP201A0CZK	cadena pesada	GMPASTSGNSNI	345	0.17%
GXRYQP201AZOWE	cadena pesada	GMPG5TSGNSNI	346	0.59%
GXRYQP201A1C3B	cadena pesada	GVPTNRDAM	347	9.12%
3126(pMet)				
ID Secuencia NGS	Cadena	Secuencia CDR3		Frec NGS (max 3,69%)
GXRYQP201AAKYU	cadena ligera	AGGYKSSGDTVS	348	0.14%
GXRYQP201AG5FC	cadena ligera	QGEFSCRDFDCTV	349	0.11%
<b>GXRYQP201A291T</b>	<b>cadena ligera</b>	<b>LGGYKTTTDGSI</b>	<b>350</b>	<b>0.78%</b>
GXRYQP201A09ZW	cadena ligera	QSYYHN5GTSYIT	351	0.89%
GXRYQP201A3ZHW	cadena ligera	AGGYKSTTDGSA	352	0.07%
GXRYQP201A12IR	cadena ligera	QSYYHNSGNSYIT	353	0.19%
GXRYQP201A5GBW	cadena ligera	QSYYYGSGTSYIT	354	0.04%
GXRYQP201BP3WS	cadena ligera	AGGYKSSGDTFT	355	0.01%
GXRYQP201AY70W	cadena ligera	LGGYKKTIIDGSA	356	0.05%
<b>GXRYQP201BRIWK</b>	<b>cadena ligera</b>	<b>AGGYKSASDGSA</b>	<b>357</b>	<b>0.30%</b>
GXRYQP201AYFKS	cadena ligera	QGEFSCDAGVCTL	358	0.07%
<b>GXRYQP201ALDF5</b>	<b>cadena ligera</b>	<b>QGEFSCRSDGTV</b>	<b>359</b>	<b>0.18%</b>
GXRYQP201AATVT	cadena ligera	LQDWSPSYADVA	360	0.06%
GXRYQP201APNW9	cadena ligera	QQGRRSVDVDNV	361	0.05%
GXRYQP201AT2TB	cadena ligera	AGGYKTTTDGSI	362	0.02%
<b>GXRYQP201AHXJZ</b>	<b>cadena ligera</b>	<b>QQGYTY5NVNDNV</b>	<b>363</b>	<b>3.69%</b>

En cursiva negrita están las cadenas de anticuerpo específicas del antígeno

(Se sintetizaron un total de 6 cadenas pesadas (cursiva) y 5 cadenas ligeras (cursiva negrita), incluyendo las cadenas de anticuerpo de mayor abundancia basándose en la frecuencia de NGS)

**Figura 8**

**Figura 9**  
**Clones de NG-XMT positivos para Western**



El emparejamiento de cadenas Pesada (9,12%) y Ligera (3,69%) de mayor frecuencia NGS no produjo anticuerpos específicos del antígeno

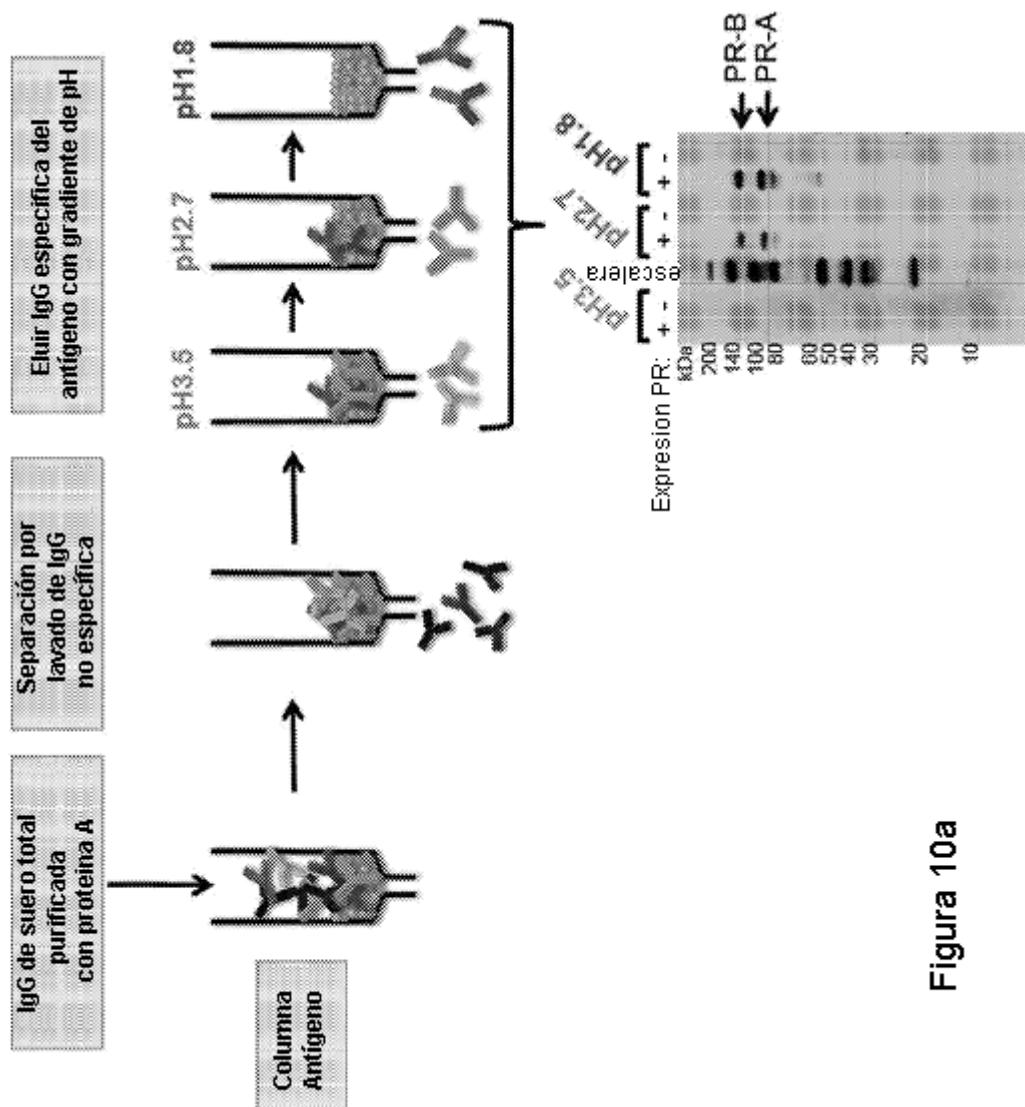
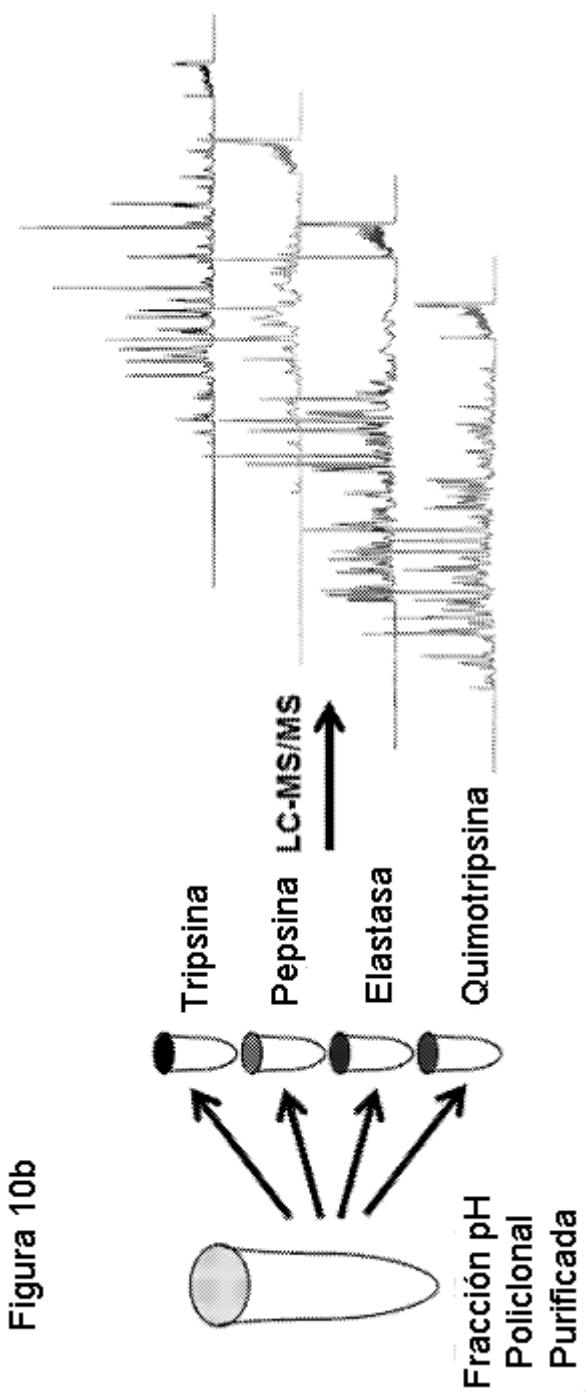


Figura 10a



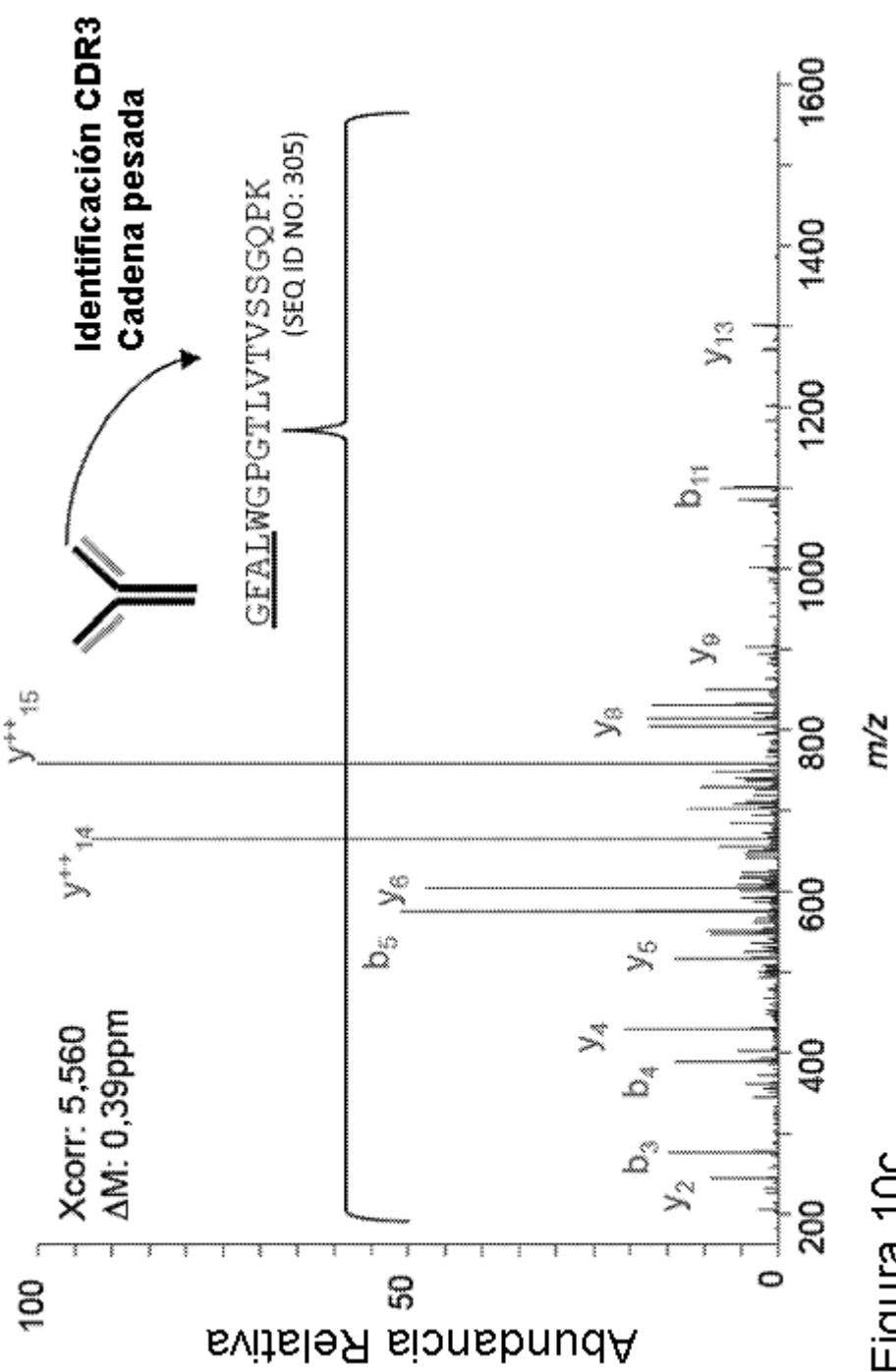


Figura 10c

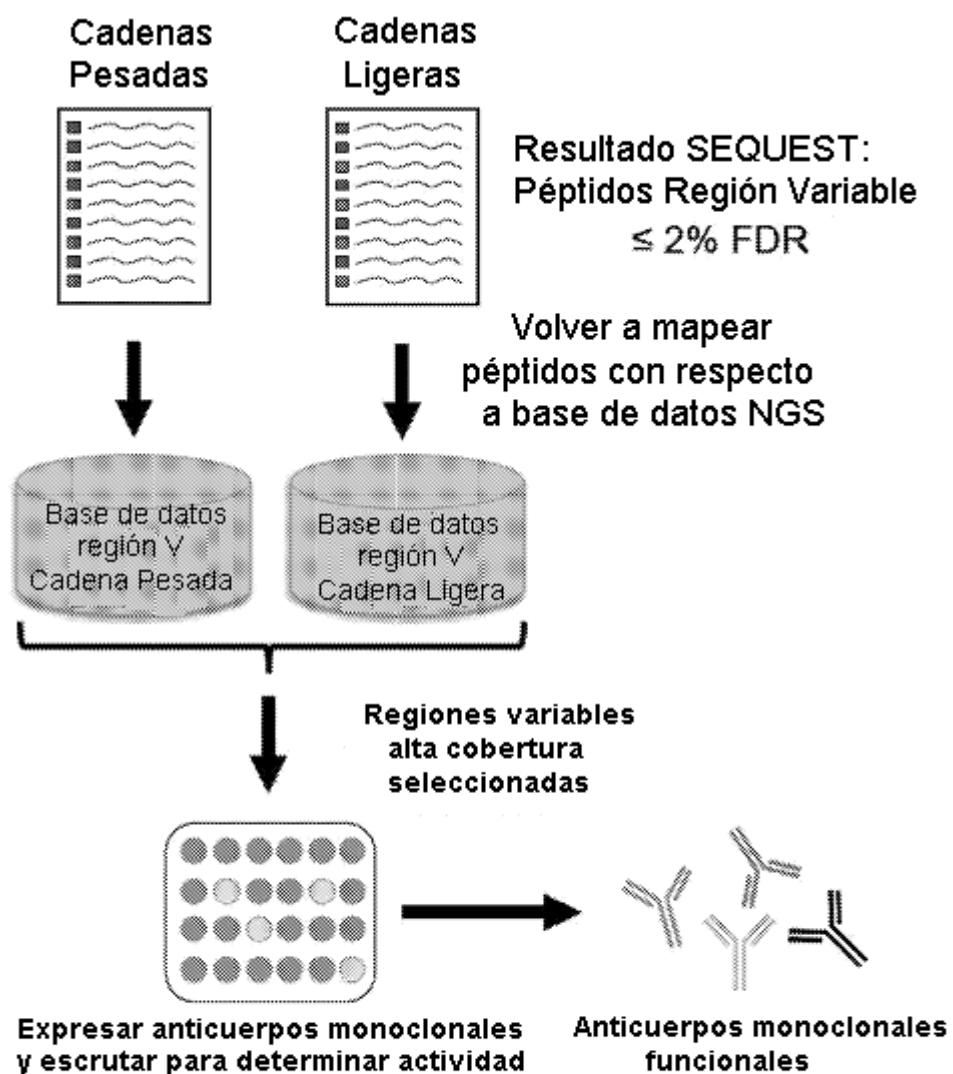


Figura 10d

Cadena Pesada: Cobertura = 77% Péptidos Totales = 50

OSVEESGCCRLVPPGTPPLTLCTVFGFDVSSHIMSWVRQAPGKGLEWIGLVDIGKSIKWYA  
SWAKGRFTISKTSSTIYLKLTRPTTGDTATYFCCSRGFALNWPGTIVTVSSGQPKAPSVP  
(SEO ID NO: 367)

Cadena Kappa: Cobertura = 65% Péptidos Totales = 24

AQVLTQTVSPVSAAVGGTVTINQASPSVYGNYLISWFQQKPGQPPKKLILINASTLPSGVSSRFKGSGSGTHFTLTISDVQCDAAATYYCLGSYDCRSDDCNVFGGGTEVVVVKGDPVAPTV  
(SEQ ID NO: 368)

Quimotripsina Pepsina CDR  
Elastasa Tripsina

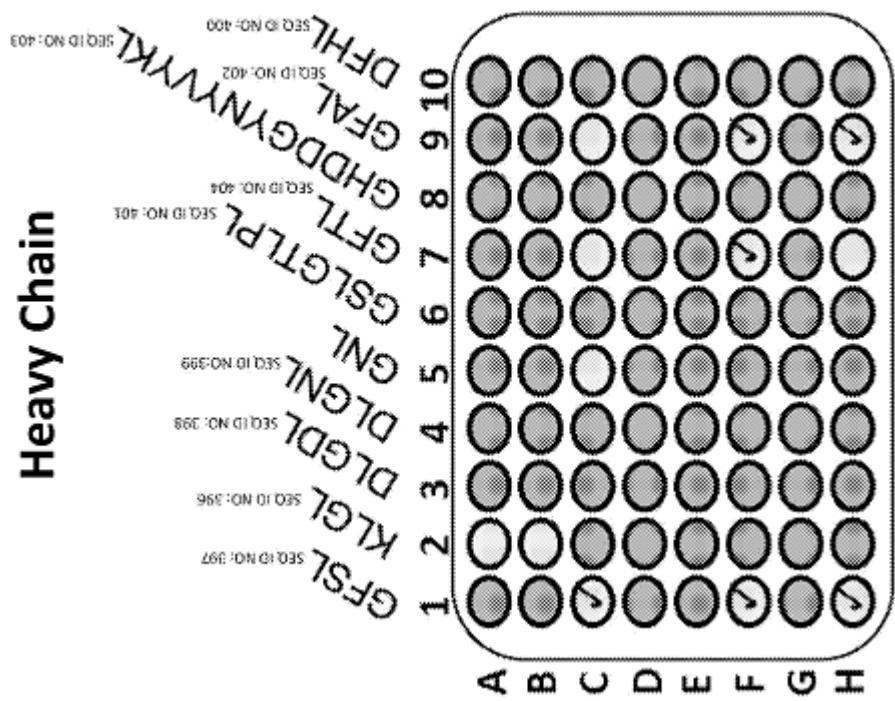
Figura 10e

**Heavy Chain**

Figura 11a

Cadena Ligera

SEQ ID NO:	405 LAGYDCTTGDCFA	406 LGYYDCDNGDCFT	407 LGTYDCRRADCNT	408 QSTLYSSTDEIV	409 QCSYVNSNT	410 LGSYDCRSSDCNV	411 LGAYDDAADNS	412 LGTYDCNSADCNV
A	LAGYDCTTGDCFA	LGYYDCDNGDCFT	LGTYDCRRADCNT	QSTLYSSTDEIV	QCSYVNSNT	LGSYDCRSSDCNV	LGAYDDAADNS	LGTYDCNSADCNV
B								
C								
D								
E								
F								
G								
H								



ELISA+  
✓ Transferencia Western+

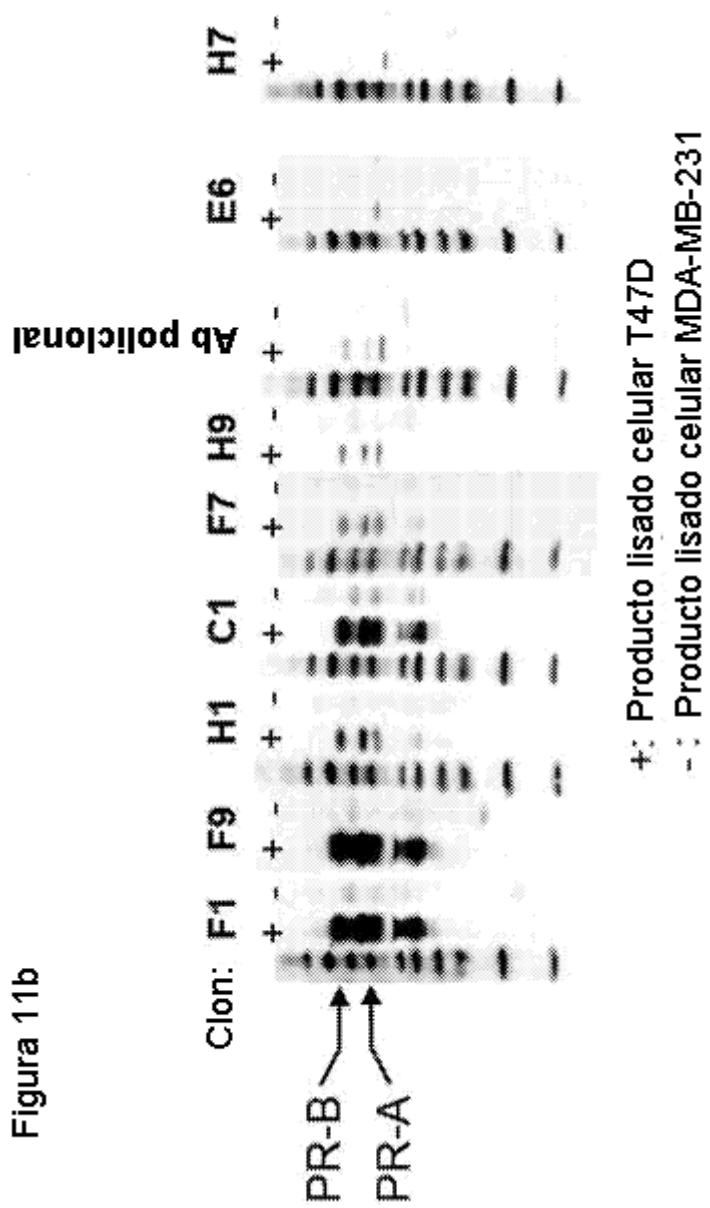


Figura 11b

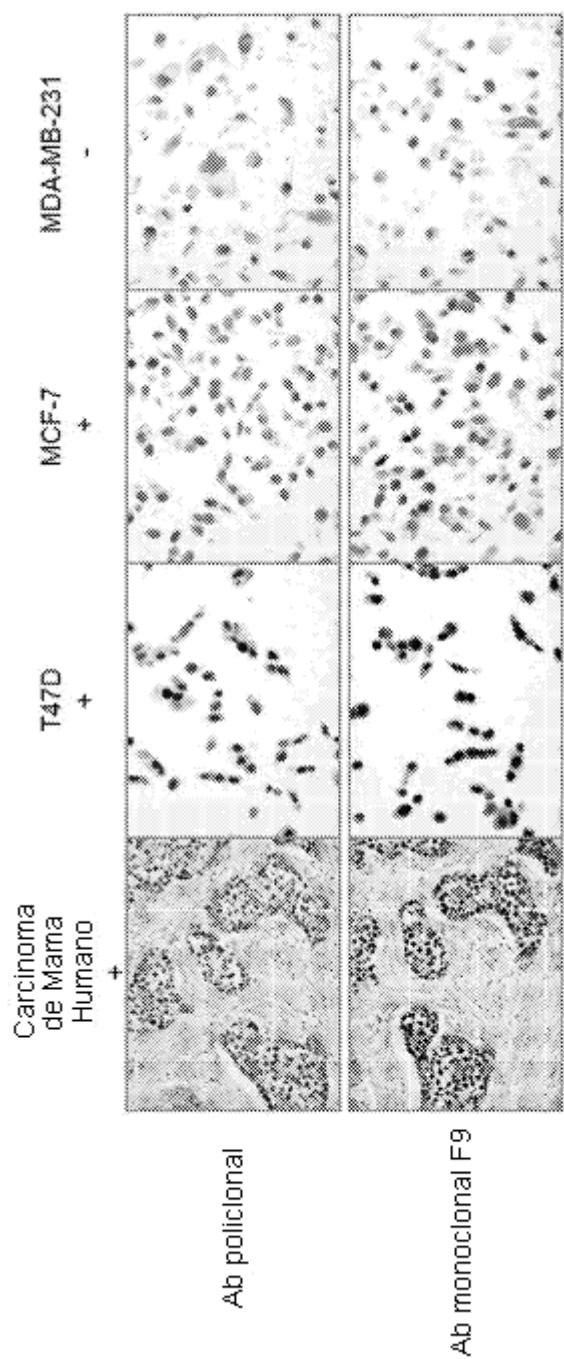
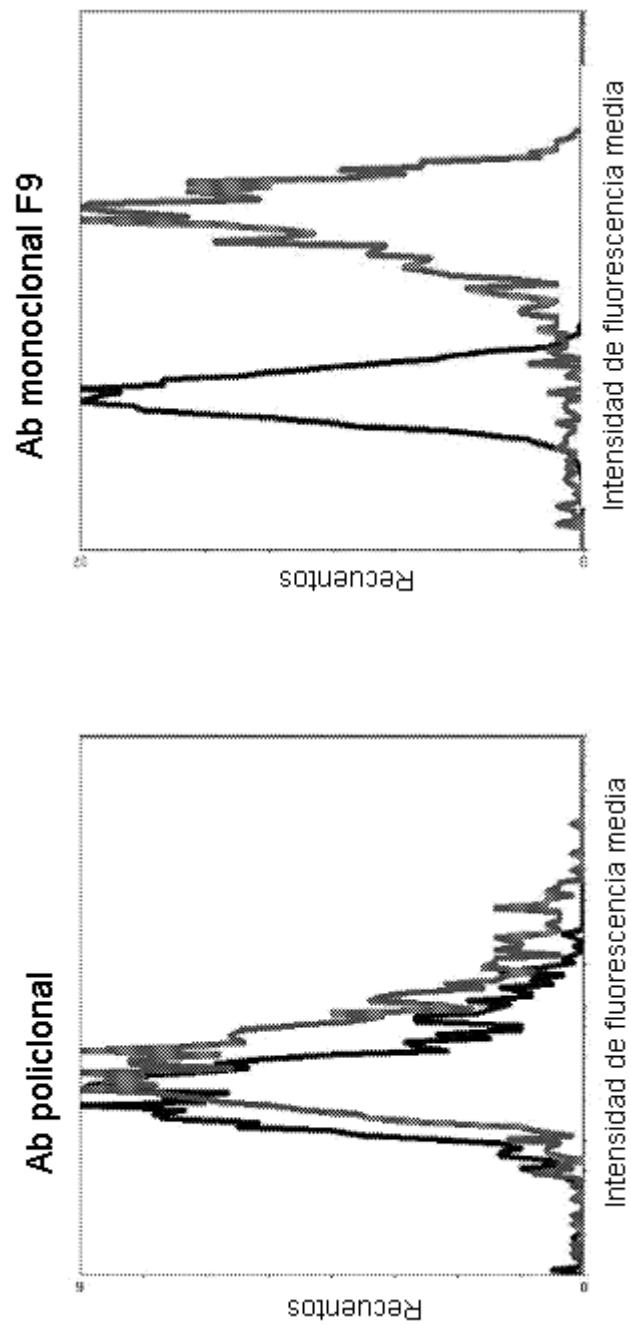


Figura 11c

Figura 11d



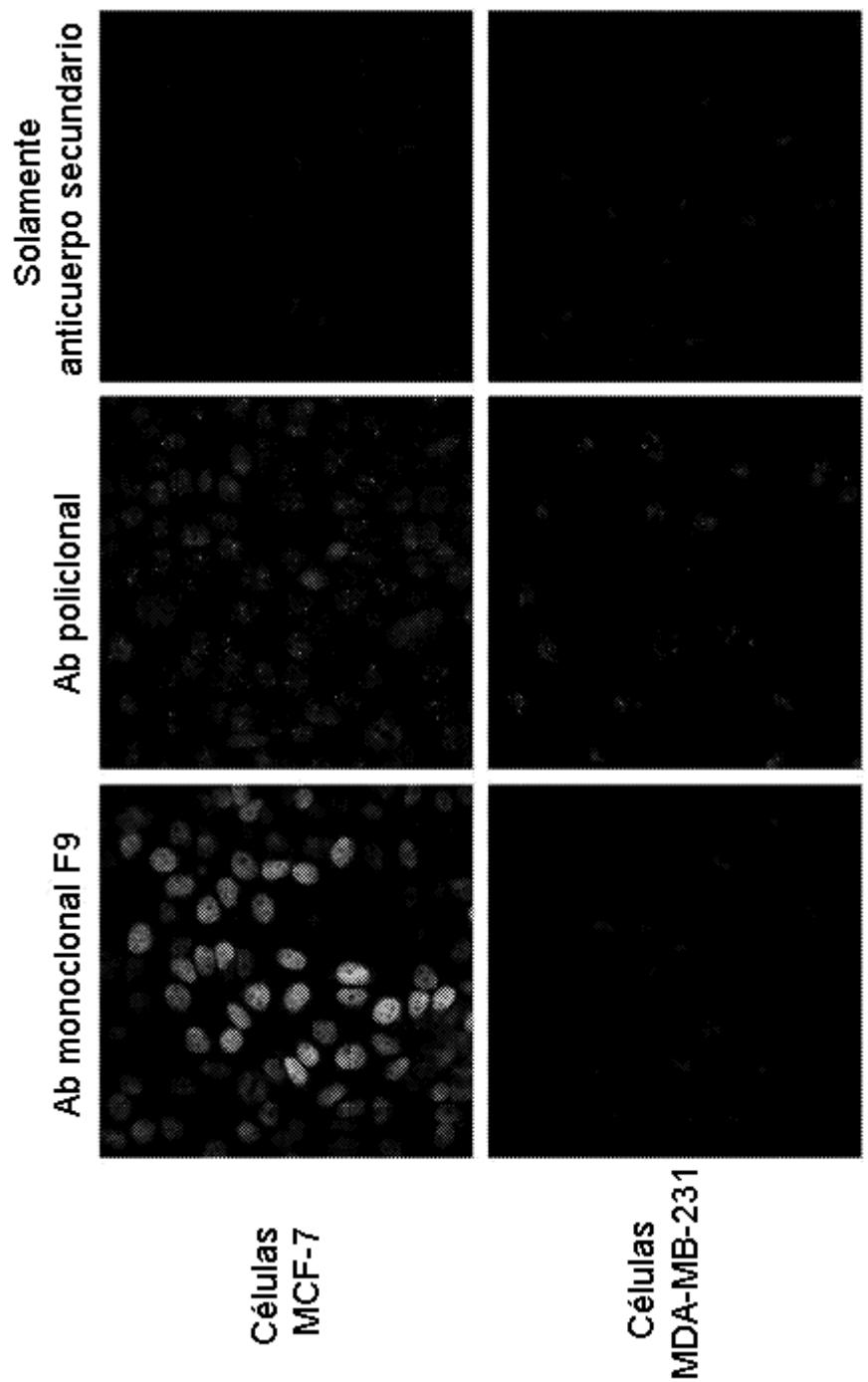
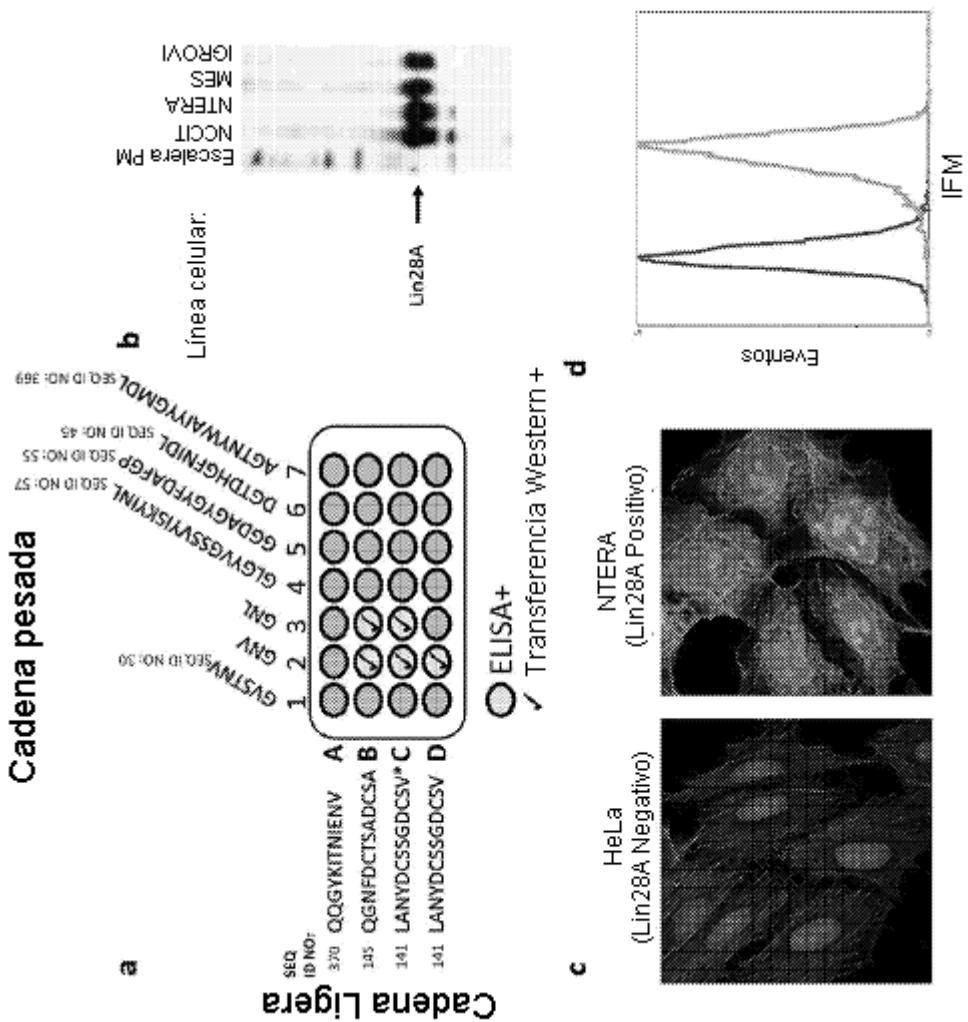
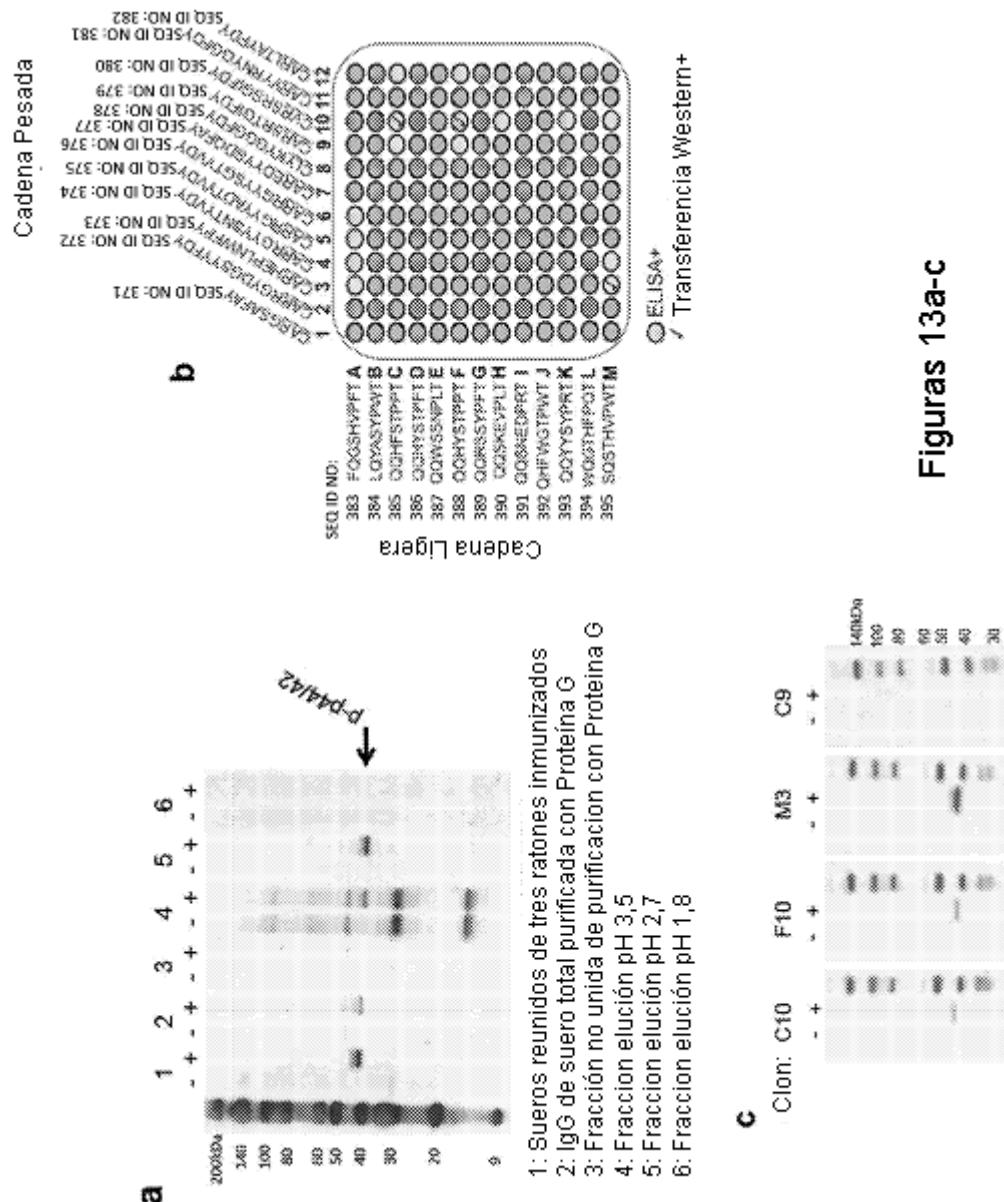


Figura 11e



Figuras 12a-d



- 1: Sueros reunidos de tres ratones inmunizados
- 2: IgG de suero total purificado con Proteína G
- 3: Fracción no unida de purificación con Proteína G
- 4: Fracción elución pH 3,5
- 5: Fracción elución pH 2,7
- 6: Fracción elución pH 1,8

**Figuras 13a-c**