

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 716 013**

51 Int. Cl.:

**C12N 5/10** (2006.01)

**C07K 16/12** (2006.01)

**C07K 16/00** (2006.01)

**G01N 33/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.04.2012 PCT/US2012/000221**

87 Fecha y número de publicación internacional: **01.11.2012 WO12148497**

96 Fecha de presentación y número de la solicitud europea: **27.04.2012 E 12775920 (7)**

97 Fecha y número de publicación de la concesión europea: **28.11.2018 EP 2702146**

54 Título: **Identificación de polinucleótidos asociados a una muestra**

30 Prioridad:

**28.04.2011 US 201161517976 P**

**24.08.2011 US 201161575652 P**

**16.02.2012 US 201261599870 P**

**08.03.2012 US 201261608571 P**

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

**07.06.2019**

73 Titular/es:

**THE BOARD OF TRUSTEES OF THE LELAND  
STANFORD JUNIOR UNIVERSITY (50.0%)**

**1705 El Camino Real**

**Palo Alto, CA 94306-1106, US y**

**DEPARTMENT OF VETERANS AFFAIRS (50.0%)**

72 Inventor/es:

**ROBINSON, WILLIAM H.;**

**TAN, YANN CHONG y**

**SOKOLOVE, JEREMY**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

ES 2 716 013 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**

Identificación de polinucleótidos asociados a una muestra

Antecedentes

5 Producir anticuerpos monoclonales terapéuticos a partir de fuentes humanas es un desafío biológico y técnico. Hasta la fecha, se han descrito varias metodologías, que incluyen la generación de hibridomas humanos, el uso de ratones transgénicos que expresan inmunoglobulinas humanas y el uso de bibliotecas de presentación de fagos de inmunoglobulina humana.

10 Los hibridomas humanos pueden ser difíciles de generar porque los asociados de fusión del mieloma humano, a diferencia de sus homólogos de ratón, son ineficientes para generar hibridomas. Los hibridomas humanos también tienen una tendencia a perder espontáneamente los genes de anticuerpos expresados después de un cultivo prolongado. La transformación del virus de Epstein-Barr (EBV) inmortaliza las células B, pero solo las fracciones extremadamente pequeñas de todas las células B transformadas con EBV son maduras por afinidad o reconocen el antígeno diana. La generación de hibridomas incluye típicamente cribas grandes para obtener anticuerpos monoclonales terapéuticos. Ninguno de los anticuerpos monoclonales terapéuticos actualmente aprobados por la F.D.A. de los Estados Unidos se crearon a través de la generación de hibridomas humanos o la transformación EBV de las células B, atestiguando las dificultades técnicas y los desafíos planteados por estos métodos.

20 Las bibliotecas de presentación de fagos de secuencias de anticuerpos humanos representan otro método para producir anticuerpos monoclonales humanos terapéuticos. Este método utiliza tecnología de ADN recombinante para expresar aleatoriamente secuencias de cadenas pesadas y livianas de anticuerpos humanos para permitir la detección de combinaciones que se unen al antígeno diana. Sin embargo, esta estrategia no produce anticuerpos madurados por afinidad, y los anticuerpos producidos de esta manera generalmente se unen al antígeno con baja afinidad y aidez. Luego se necesitan sucesivos pasos de mutación y selección/proyección para generar anticuerpos de alta afinidad. En otros métodos, las bibliotecas de anticuerpos se expresan y seleccionan en un sistema in vitro sin células. Los sistemas de expresión in vitro actuales, aunque son buenos para expresar dominios variables de anticuerpos individuales, son ineficientes para expresar anticuerpos de múltiples dominios tales como moléculas de anticuerpos de cadena única (scFv). Esto se debe tanto a la estructura de las bibliotecas de anticuerpos scFv actuales como a las condiciones de reacción de los sistemas de expresión in vitro actuales. El documento US2010/ 099103 describe bibliotecas de ARN de visualización in vitro mejoradas para permitir la expresión y selección fiables de moléculas de anticuerpo scFv. Esto tiene varias ventajas, que incluyen proporcionar una biblioteca de anticuerpos scFv de visualización in vitro mejorada que contiene un enlazador interdominio optimizado para una expresión mejorada; proporcionando una biblioteca de anticuerpos scFv de visualización in vitro mejorada que contiene códigos de barras cortos de ácido nucleico; proporcionar cebadores para generar las bibliotecas de anticuerpos scFv de visualización in vitro mejoradas; proporcionar cebadores para tratar las regiones CDR3 de las regiones variables de la cadena pesada de las moléculas de anticuerpo scFv en dichas bibliotecas; y métodos para hacer las bibliotecas de visualización in vitro mejoradas.

40 Otra forma de producir anticuerpos monoclonales humanos terapéuticos es mediante la creación o el uso de ratones transgénicos que posean un repertorio de anticuerpos humanos. Cuando se inmunizan, tales ratones producen anticuerpos que se dirigen al antígeno inmunizante, y luego se pueden generar hibridomas para la producción de anticuerpos monoclonales humanos terapéuticos. Dichos ratones transgénicos están patentados y no están comúnmente disponibles para su uso en la generación de anticuerpos humanos.

45 Por lo tanto, los inventores han identificado la necesidad de composiciones, kits y métodos que, por ejemplo, puedan producir grandes cantidades de anticuerpos humanos madurados por afinidad, evitando la necesidad de una humanización laboriosa y prolongada de un anticuerpo, o la necesidad de realizar extensas pruebas de detección. Las composiciones, kits y métodos descritos en este documento abordan esta necesidad. Además, las composiciones, los kits y los métodos descritos en el presente documento son ampliamente aplicables fuera del espacio del anticuerpo humano y se pueden usar en varias aplicaciones diferentes que incluyen, por ejemplo, la coincidencia de dos o más polinucleótidos de interés que se derivan de una sola muestra y están presentes en una biblioteca de polinucleótidos.

Resumen

50 La invención proporciona una biblioteca de polinucleótidos que comprende una pluralidad de composiciones, en donde la biblioteca comprende ADNc que codifican pares afines de regiones variables de cadenas pesadas y livianas de inmunoglobulina y en donde cada composición está presente en un recipiente separado, cada composición comprende polinucleótidos que comprenden moléculas de ADNc derivadas de un único plasmoblasto que comprenden moléculas de ADNc que codifican un par análogo de región variable de cadena pesada de inmunoglobulina y región variable de cadena liviana de inmunoglobulina del plasmoblasto, y una región de identificación de la muestra unida a las moléculas de ADNc, en donde la secuencia de nucleótidos de la región de identificación de la muestra es distinta de la secuencia de nucleótidos de la región de identificación de la muestra de las otras composiciones presentes en otros contenedores separados en la biblioteca.

La invención también proporciona un método para producir uno o más polinucleótidos de interés, que comprende: obtener la biblioteca de la invención; y amplificar la biblioteca de polinucleótidos con un conjunto de cebadores para producir uno o más polinucleótidos de interés, en donde el polinucleótido de interés es el par análogo de regiones variables de cadena liviana y cadena pesada de inmunoglobulina.

- 5 En el presente documento también se describe una composición que comprende un polinucleótido, en donde el polinucleótido comprende una primera región y una segunda región, en donde la primera región comprende una región de inmunoglobulina variable expresada de células B y la segunda región comprende al menos una región de identificación, y en donde la primera región está acoplado a la segunda región.

10 Como se describe en el presente documento, la región de inmunoglobulina variable comprende una región VDJ de una secuencia de nucleótidos de inmunoglobulina IgG aislada de una célula B humana activada mayor o igual a 8  $\mu\text{m}$  de diámetro, y en donde el extremo 5' de la región de inmunoglobulina está acoplado al extremo 3' de la región de identificación. En algunos aspectos, la composición está comprendida en una familia clonal.

15 Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B activada. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es un plasmoblasto. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una única célula B. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una única célula B activada. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B activada única localizada en la sangre de un sujeto. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B activada humana. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula plasmática. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B específica de antígeno. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de mamífero. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de ratón. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto con una enfermedad o condición de interés. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto que se recupera o se recupera de una enfermedad o condición de interés. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto administrado con al menos un antígeno de interés. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto administrado con al menos un antígeno de interés y un adyuvante. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B ubicada en la sangre de un sujeto. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B ubicada en la médula ósea de un sujeto. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B ubicada en el bazo de un sujeto. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B ubicada en al menos un ganglio linfático de un sujeto. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B ubicada en el tejido linfoide de un sujeto. Como se describe en el presente documento, la región de inmunoglobulina se aísla de una célula B ubicada en el intestino de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de aproximadamente 8-20  $\mu\text{m}$ . En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que es 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, o más de 20  $\mu\text{m}$  de diámetro. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que es aproximadamente 60, 70, 80, 90, 100, 120, 130, 140, 150, 200, 250, 300, 350, o más de 350  $\mu\text{m}^2$  en el área. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que es aproximadamente 250, 268, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000 o superior a 4000  $\mu\text{m}^3$  en volumen. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de 10% o más en tamaño que el diámetro mediano de una célula B de control en reposo. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de 15% o más en tamaño que el diámetro mediano de una célula B de control en reposo. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de 20% o más en tamaño que el diámetro medio de una célula B en reposo de control. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada capaz de secretar inmunoglobulina. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B en la fase de brecha 1 (G1), síntesis(s), brecha 2 (G2) o mitosis (M) del ciclo celular. En algunos aspectos, la región de inmunoglobulina que se aísla de una célula B no está en la fase 0 (G0) del ciclo celular. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B caracterizada por tener un FSC mayor que 1.2x del promedio FSC de linfocitos B en reposo por citometría de flujo. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B caracterizada por tener una media FSC entre 0.7 - 1.15x de la media FSC de monocitos humanos mediante citometría de flujo. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B positiva para CD19. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B positiva para CD38. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B positiva para CD27. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B negativa para CD20. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B CD19<sup>+</sup>CD20<sup>-</sup>CD27<sup>+</sup>CD38<sup>hi</sup>.

60 En algunos aspectos, el extremo 5' de la región de inmunoglobulina está acoplado al extremo 3' de la región de identificación.

En algunos aspectos, la región de inmunoglobulina variable comprende una región VDJ de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una región VJ de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una región V, D y/o J de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una cadena pesada y/o liviana de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una secuencia de inmunoglobulina IgG, IgM, IgD, IgE o IgA. En algunos aspectos, la región de inmunoglobulina variable comprende una secuencia de inmunoglobulina IgG1, IgG2, IgG3 o IgG4 humana. En algunos aspectos, la región de inmunoglobulina variable comprende una secuencia de inmunoglobulina IgG1, IgG2a, IgG2b o IgG3 de ratón. En algunos aspectos, la región de inmunoglobulina tiene una longitud de aproximadamente 200-2000 nucleótidos. En algunos aspectos, la región de inmunoglobulina es menor que 200, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, o más de 2000 nucleótidos de longitud.

En algunos aspectos, la región de identificación comprende una pluralidad de regiones de identificación. En algunos aspectos, la región de identificación comprende una pluralidad de regiones de identificación, y en donde cada región de identificación en la pluralidad tiene una secuencia distinta. En algunos aspectos, la región de identificación comprende al menos una región de identificación de la muestra y al menos una región de identificación de la placa. En algunos aspectos, la región de identificación comprende una secuencia distinta de la secuencia de la región de inmunoglobulina. En algunos aspectos, la región de identificación tiene una longitud de aproximadamente 2 a 100 nucleótidos. En algunos aspectos, la región de identificación es 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100 o más de 100 nucleótidos de longitud. En algunos aspectos, la región de identificación tiene una longitud de 2 a 1000 nucleótidos. En algunos aspectos, la región de identificación tiene una longitud igual o superior a 100 nucleótidos. En algunos aspectos, la región de identificación comprende una secuencia de nucleótidos no codificante contigua. En algunos aspectos, la región de identificación comprende una secuencia de nucleótidos no codificante. En algunos aspectos, la región de identificación comprende una secuencia de nucleótidos no contigua, no codificante. En algunos aspectos, la longitud de la secuencia de la región de identificación es menor que la longitud de la secuencia de la región de inmunoglobulina.

En algunos aspectos, una composición descrita en el presente documento puede incluir una tercera región, en donde la tercera región comprende una región adaptadora. En algunos aspectos, la tercera región comprende una región adaptadora, y en donde la tercera región está situada entre la primera región y la segunda región. En algunos aspectos, la tercera región comprende una región adaptadora, y en donde la región adaptadora comprende al menos un nucleótido G situado en su extremo 3'.

En algunos aspectos, la región de identificación tiene una longitud de 2 a 100 nucleótidos y tiene una secuencia distinta de la secuencia de la región de inmunoglobulina, y en donde la región del adaptador comprende al menos un nucleótido G en su extremo 3' y está ubicada a 3' de la región de identificación de la muestra y 5' de la región de inmunoglobulina, y en donde la región variable de inmunoglobulina ha sufrido hipermutación y se diferencia de la secuencia de la línea germinal de una célula B ingenua.

Como se describe en este documento, la composición está presente en un recipiente. En algunos aspectos, una pluralidad de las composiciones está presente en un recipiente. Como se describe en el presente documento, una pluralidad de las composiciones está presente en un solo pozo de una placa única que comprende una pluralidad de pozos.

Como se describe en el presente documento, la composición se encuentra en una biblioteca de composiciones, en donde cada composición está presente en un contenedor separado, en donde cada composición comprende un polinucleótido que comprende una primera región y una segunda región, en donde la primera región comprende una región de inmunoglobulina variable expresada en el plasmoblasto y la segunda región comprende una región de identificación, en donde la primera región está acoplada a la segunda región, en donde la secuencia de nucleótidos de cada región de identificación o es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en la biblioteca, y en donde las últimas secuencias de nucleótidos de una pluralidad de regiones de inmunoglobulina variables en la biblioteca comparten al menos 80-99% de identidad de secuencia.

Como se describe en el presente documento, la composición está comprendida en una biblioteca que comprende una pluralidad de composiciones de polinucleótidos, en donde cada composición está presente en un recipiente separado, en donde cada composición comprende un polinucleótido, en donde el polinucleótido comprende una primera región y una segunda región, en donde la primera la región comprende una región de inmunoglobulina variable expresada en células B y la segunda región comprende una región de identificación, en donde la primera región está acoplada a la segunda región, y en donde la secuencia de nucleótidos de cada región de identificación es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en cada contenedor separado en la biblioteca.

En el presente documento también se describe una biblioteca de composiciones de polinucleótidos que comprende una pluralidad de composiciones de polinucleótidos, en donde cada composición está presente en un recipiente separado, en donde cada composición comprende un polinucleótido, en donde el polinucleótido comprende una primera región y una segunda región, en donde la primera región comprende un la región de inmunoglobulina variable



expresada de células B y la segunda región comprende una región de identificación, en donde la primera región está acoplada a la segunda región, y en donde la secuencia de nucleótidos de cada región de identificación es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en cada contenedor en la biblioteca.

5 En este documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido de la pluralidad está presente en un recipiente separado, en donde cada polinucleótido de la pluralidad comprende una primera región y una segunda región, en donde la primera región comprende una célula B expresada la región de inmunoglobulina variable y la segunda región comprende una región de identificación, en donde la primera región está acoplada a la segunda región, en donde la secuencia de nucleótidos de cada región de identificación es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en la biblioteca, y en donde al menos dos regiones de inmunoglobulina variables en la pluralidad compartan al menos 80-99% de identidad de secuencia.

10 En el presente documento también se describe una biblioteca de polinucleótidos que comprende una familia clonal de polinucleótidos, en donde cada polinucleótido en la familia comprende una primera región y una segunda región, en donde la primera región comprende una región de inmunoglobulina variable expresada de células B y la segunda región comprende una región de identificación, en donde la primera región está acoplada a la segunda región, en donde la secuencia de nucleótidos de cada región de identificación es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en la familia, y en donde cada una de las regiones de inmunoglobulina variables en la familia exhiben al menos 80-99% de identidad de secuencia. En algunos aspectos, la biblioteca comprende una pluralidad de familias clonales.

15 En el presente documento también se describe una familia clonal de secuencias de inmunoglobulina en donde cada secuencia en la familia está acoplada a una región de identificación. Como se describe en este documento, cada región de identificación es distinta de las otras regiones de identificación. Como se describe en el presente documento, las secuencias de inmunoglobulina comprenden secuencias de inmunoglobulina de cadena pesada. Como se describe en el presente documento, las secuencias de inmunoglobulina comprenden secuencias de inmunoglobulina de cadena liviana. Como se describe en el presente documento, las secuencias de inmunoglobulina comprenden secuencias de inmunoglobulina de cadena pesada y cadena liviana. Como se describe en el presente documento, una o más de las regiones de identificación comprenden una secuencia de inmunoglobulina de cadena liviana. Como se describe en el presente documento, una o más de las regiones de identificación comprenden una secuencia de inmunoglobulina de cadena pesada.

20 También se describe en este documento un conjunto de dos o más de las familias clonales descritas en este documento.

También se describe en este documento un conjunto de 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más de las familias clonales descritas en este documento.

25 En el presente documento también se describe una familia clonal de secuencias de inmunoglobulina en donde cada secuencia en la familia está acoplada operativamente a al menos una secuencia de nucleótidos contigua. Como se describe en el presente documento, las secuencias de inmunoglobulina comprenden secuencias de inmunoglobulina de cadena pesada y la al menos una secuencia de nucleótidos contigua comprende una secuencia de inmunoglobulina de cadena liviana. Como se describe en el presente documento, las secuencias de inmunoglobulina comprenden secuencias de inmunoglobulina de cadena liviana y la al menos una secuencia de nucleótidos contigua comprende una secuencia de inmunoglobulina de cadena pesada.

30 En el presente documento también se describe un método para producir una familia clonal de secuencias de inmunoglobulinas que comprende obtener una pluralidad de secuencias de inmunoglobulinas, cada una con regiones V, D y/o J y cada una acoplada a una región de identificación; y agrupar dos o más secuencias de la pluralidad para producir la familia clonal en donde cada secuencia en la familia clonal es una versión mutada de la misma secuencia de inmunoglobulina de línea germinal que tiene una región V, D y/o J o la secuencia de inmunoglobulina de línea germinal que tiene la región V, D, y/o J.

En algunas divulgaciones, cada región de identificación es distinta de las otras regiones de identificación.

35 En el presente documento también se describe un método para producir una familia clonal de secuencias de inmunoglobulina que comprende obtener una pluralidad de secuencias de inmunoglobulina, cada una con regiones V, D y/o J y cada una acoplada a una región de identificación, y en donde cada región de identificación es distinta de la otra regiones de identificación; eliminar una o más regiones de identificación; y agrupar dos o más secuencias de la pluralidad para producir la familia clonal en donde cada secuencia en la familia clonal es una versión mutada de la misma secuencia de inmunoglobulina de línea germinal que tiene una región V, D y/o J o la secuencia de inmunoglobulina de línea germinal que tiene la región V, D, y/o J.

40 En el presente documento también se describe un método para identificar un segundo ADNc acoplado a una primera región de identificación que comprende seleccionar un primer ADNc acoplado a la primera región de identificación e

identificar el segundo ADNc basándose en la identidad compartida de la región de identificación acoplada a cada ADNc.

5 En el presente documento también se describe un método para producir una cola 3' en una segunda secuencia de nucleótidos que comprende obtener una primera secuencia de nucleótidos y poner en contacto la primera secuencia de nucleótidos con una transcriptasa reversa ARNasa H<sup>-</sup> termoestable que tiene actividad de cambio de plantilla a menos de 50°C, en donde el contacto produce la cola 3' y la segunda secuencia de nucleótidos. En algunos aspectos, la primera secuencia de nucleótidos se pone en contacto a menos de 50, 49, 48, 47, 46, 45, 44, 43, 42, o menos de 42°C. En algunos aspectos, la primera secuencia de nucleótidos se pone en contacto a 42°C. En algunos aspectos, la primera secuencia de nucleótidos se pone en contacto a 45.5°C. En algunos aspectos, la transcriptasa es un Virus de la Leucemia Murina de Moloney (MMLV) RNasa H<sup>-</sup>transcriptasa reversa. En algunos aspectos, la transcriptasa es SuperScript III.

15 En el presente documento también se describe un método para determinar la secuencia de origen natural de una primera secuencia de interés que comprende obtener una pluralidad de secuencias relacionadas con la primera secuencia y cada una acoplada a una primera región de identificación, en donde cada primera región de identificación es idéntica, y en donde una o más de las secuencias en la pluralidad es distinta de la secuencia natural; y comparar las secuencias en la pluralidad para determinar la secuencia que ocurre naturalmente en la primera secuencia de interés. En algunas descripciones, la pluralidad de secuencias comprende secuencias de inmunoglobulina. En algunas descripciones, la pluralidad de secuencias comprende secuencias de inmunoglobulina. En algunas descripciones, la pluralidad de secuencias se acopla cada una a una segunda región de identificación y cada segunda región de identificación es idéntica. En algunas descripciones, la primera secuencia de interés es una secuencia de inmunoglobulina. En algunas divulgaciones, la pluralidad de secuencias son secuencias de inmunoglobulina.

25 En el presente documento también se describe una composición que comprende un polinucleótido que comprende una primera región y una segunda región, en donde la primera región comprende una región de inmunoglobulina variable derivada de células B y la segunda región comprende una región de identificación, y en donde la primera región está acoplada a la segunda región.

30 En el presente documento también se describe una biblioteca de composiciones de polinucleótidos que comprende una pluralidad de composiciones de polinucleótidos, en donde cada composición está presente en un recipiente separado, en donde cada composición comprende un polinucleótido que comprende una región de inmunoglobulina variable derivada de células B y una región de identificación, en donde la región de inmunoglobulina variable está acoplado a la región de identificación, en donde la secuencia de nucleótidos de cada región de identificación es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en cada contenedor separado en la biblioteca.

35 En el presente documento también se describe un método para producir una composición polinucleotídica, que comprende: obtener un polinucleótido que comprende una primera región, en donde la primera región comprende una región de inmunoglobulina variable expresada de células B asociada con un sujeto; y generar la composición polinucleotídica que comprende la primera región y una segunda región acoplado la primera región a la segunda región, en donde la segunda región comprende una región de identificación.

40 En algunos aspectos de la divulgación, obtener el polinucleótido comprende obtener una célula B asociada con el sujeto y procesar la célula para preparar el polinucleótido. En algunos aspectos de la divulgación, obtener el polinucleótido comprende recibir el polinucleótido directa o indirectamente de un tercero que ha procesado una célula B asociada con el sujeto para preparar el polinucleótido. En algunos aspectos de la divulgación, obtener el polinucleótido comprende recibir el polinucleótido directa o indirectamente de un tercero que ha solubilizado una célula B asociada con el sujeto para preparar el polinucleótido. En algunos aspectos de la divulgación, obtener el polinucleótido comprende obtener una célula B utilizando un citómetro de flujo. En algunos aspectos de la divulgación, obtener el polinucleótido comprende obtener una célula B utilizando un dispositivo de movilización de microfluidos

50 En algunos aspectos de la divulgación, la región de inmunoglobulina variable comprende una región VDJ de una secuencia de nucleótidos de inmunoglobulina IgG aislada de una célula B humana activada mayor o igual a 8 µm de diámetro, y en donde el extremo 5' de la región de inmunoglobulina está acoplado hasta el extremo 3' de la región de identificación. En algunos aspectos de la divulgación, la composición está comprendida en una familia clonal.

55 En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B activada. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es un plasmoblasto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B única. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una única célula B activada. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B activada única localizada en la sangre de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B activada humana. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B de memoria.

En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula plasmática. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B, y en donde la célula B es una célula B específica de antígeno. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de mamífero. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B humana. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de ratón. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto con una enfermedad o condición de interés. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto que se recupera o se recupera de una enfermedad o condición de interés. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto administrado con al menos un antígeno de interés. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B de un sujeto administrado con al menos un antígeno de interés y un adyuvante. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B ubicada en la sangre de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B ubicada en la médula ósea de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B ubicada en el bazo de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B ubicada en al menos un ganglio linfático de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B ubicada en el tejido linfóide de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B ubicada en el intestino de un sujeto. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de aproximadamente 8-20  $\mu\text{m}$ . En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que es 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, o más de 20  $\mu\text{m}$  de diámetro. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que es aproximadamente 60, 70, 80, 90, 100, 120, 130, 140, 150, 200, 250, 300, 350, o más de 350  $\mu\text{m}^2$  en el área. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que es aproximadamente 250, 268, 300, 400, 500, 600, 700, 800, 900, 1000, 2000, 3000, 4000 o superior a 4000  $\mu\text{m}^3$  en volumen. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de 10% o más en tamaño que el diámetro mediano de una célula B de control en reposo. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de 15% o más en tamaño que el diámetro mediano de una célula B de control en reposo. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada que tiene un diámetro de 20% o más en tamaño que el diámetro medio de una célula B en reposo de control. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B activada capaz de secretar inmunoglobulina. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B en la fase de brecha 1 (G1), síntesis(s), brecha 2 (G2) o mitosis (M) del ciclo celular. En algunos aspectos, la región de inmunoglobulina que se aísla de una célula B no está en la fase 0 (G0) del ciclo celular. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B caracterizada por tener un FSC mayor que 1.2x del promedio FSC de linfocitos B en reposo por citometría de flujo. En algunos aspectos, la región de inmunoglobulina se aísla de una célula B caracterizada por tener una media FSC entre 0.7 - 1.15x de la media FSC de monocitos humanos mediante citometría de flujo. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B positiva para CD19. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B positiva para CD38. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B positiva para CD27. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B negativa para CD20. En algunos aspectos, la región de inmunoglobulina se aísla de una única célula B CD19<sup>+</sup>CD20<sup>-</sup>CD27<sup>+</sup>CD38<sup>hi</sup>.

En algunos aspectos, el extremo 5' de la región de inmunoglobulina está acoplado al extremo 3' de la región de identificación.

En algunos aspectos, la región de inmunoglobulina variable comprende una región VDJ de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una región VJ de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una región V, D y/o J de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una cadena pesada y/o liviana de una secuencia de nucleótidos de inmunoglobulina. En algunos aspectos, la región de inmunoglobulina variable comprende una secuencia de inmunoglobulina IgG, IgM, IgD, IgE o IgA. En algunos aspectos, la región de inmunoglobulina variable comprende una secuencia de inmunoglobulina IgG1, IgG2, IgG3 o IgG4 humana. En algunos aspectos, la región de inmunoglobulina variable comprende una secuencia de inmunoglobulina IgG1, IgG2a, IgG2b o IgG3 de ratón. En algunos aspectos, la región de inmunoglobulina tiene una longitud de aproximadamente 200-2000 nucleótidos. En algunos aspectos, la región de inmunoglobulina es menor que 200, 200, 300, 400, 500, 600, 700, 800, 900, 1000, 1100, 1200, 1300, 1400, 1500, 1600, 1700, 1800, 1900, 2000, o más de 2000 nucleótidos de longitud.

En algunos aspectos, la región de identificación comprende una pluralidad de regiones de identificación. En algunos aspectos, la región de identificación comprende una pluralidad de regiones de identificación, y en donde cada región de identificación en la pluralidad tiene una secuencia distinta. En algunos aspectos, la región de identificación comprende al menos una región de identificación de la muestra y al menos una región de identificación de la placa. En algunos aspectos, la región de identificación comprende una secuencia distinta de la secuencia de la región de inmunoglobulina. En algunos aspectos, la región de identificación tiene una longitud de aproximadamente 2 a 100 nucleótidos. En algunos aspectos, la región de identificación es 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100 o más de 100 nucleótidos de longitud. En algunos aspectos, la región de identificación tiene una longitud de 2 a 1000 nucleótidos. En algunos aspectos, la región de identificación tiene una longitud igual o superior a 100 nucleótidos. En algunos aspectos, la región de identificación comprende una secuencia

de nucleótidos no codificante contigua. En algunos aspectos, la región de identificación comprende una secuencia de nucleótidos no codificante. En algunos aspectos, la región de identificación comprende una secuencia de nucleótidos no contigua, no codificante. En algunos aspectos, la longitud de la secuencia de la región de identificación es menor que la longitud de la secuencia de la región de inmunoglobulina.

5 En algunos aspectos, una composición descrita en el presente documento puede incluir una tercera región, en donde la tercera región comprende una región adaptadora. En algunos aspectos, la tercera región comprende una región adaptadora, y en donde la tercera región está situada entre la primera región y la segunda región. En algunos aspectos, la tercera región comprende una región adaptadora, y en donde la región adaptadora comprende al menos un nucleótido G situado en su extremo 3'.

10 En algunos aspectos, la región de identificación tiene una longitud de 2 a 100 nucleótidos y tiene una secuencia distinta de la secuencia de la región de inmunoglobulina, y en donde la región del adaptador comprende al menos un nucleótido G en su extremo 3' y está ubicada a 3' de la región de identificación de la muestra y 5' de la región de inmunoglobulina, y en donde la región variable de inmunoglobulina ha sufrido hipermutación y se diferencia de la secuencia de la línea germinal de una célula B ingenua.

15 En algunos aspectos, la composición está presente en un recipiente. En algunos aspectos, una pluralidad de las composiciones está presente en un recipiente. En algunos aspectos, una pluralidad de las composiciones está presente en un solo pozo de una placa única que comprende una pluralidad de pozos.

20 En algunos aspectos, la composición está en una biblioteca de composiciones, en donde cada composición está presente en un recipiente separado, en donde cada composición comprende un polinucleótido que comprende una primera región y una segunda región, en donde la primera región comprende una región de inmunoglobulina variable expresada de células B y la segunda región comprende una región de identificación, en donde la primera región está acoplada a la segunda región, en donde la secuencia de nucleótidos de cada región de identificación o es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en la biblioteca, y en donde el último nucleótido las secuencias de una pluralidad de regiones de inmunoglobulina variables en la biblioteca comparten al menos 80-99% de identidad de secuencia.

30 En algunos aspectos, la composición está comprendida en una biblioteca que comprende una pluralidad de composiciones de polinucleótidos, en donde cada composición está presente en un recipiente separado, en donde cada composición comprende un polinucleótido, en donde el polinucleótido comprende una primera región y una segunda región, en donde la primera región comprende una región de inmunoglobulina variable expresada en células B y la segunda región comprende una región de identificación, en donde la primera región está acoplada a la segunda región, y en donde la secuencia de nucleótidos de cada región de identificación es distinta de la secuencia de nucleótidos de las otras regiones de identificación presentes en cada contenedor separado en la biblioteca.

35 En el presente documento también se describe un método para producir una composición polinucleotídica, que comprende: obtener una célula B asociada con un sujeto; aislamiento de polinucleótidos de la célula que comprende una región de inmunoglobulina variable expresada de células B; y generar la composición polinucleotídica que comprende la región de inmunoglobulina variable y una región de identificación mediante el acoplamiento de la región de inmunoglobulina variable a la región de identificación.

40 En el presente documento también se describe un método para producir una composición polinucleotídica, que comprende: obtener un polinucleótido que comprende una región de inmunoglobulina variable derivada de células B asociada con un sujeto; y generar la composición polinucleotídica que comprende la región de inmunoglobulina variable y una región de identificación mediante el acoplamiento de la región de inmunoglobulina variable a la región de identificación.

45 Como se describe en este documento, obtener el polinucleótido comprende obtener una célula B y procesar la célula para preparar el polinucleótido. En algunos aspectos, obtener el polinucleótido comprende recibir el polinucleótido directa o indirectamente de un tercero que ha procesado una célula B para preparar el polinucleótido.

50 En el presente documento también se describe un método para producir dos o más composiciones polinucleotídicas, que comprende: obtener una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos asociados con una pluralidad de muestras obtenidas de uno o más sujetos, en donde uno o más polinucleótidos comprenden una región de inmunoglobulina variable expresada en células B, en donde cada muestra está asociada con una célula B, y en donde cada polinucleótido asociado con cada muestra está presente en un contenedor separado; y generar dos o más composiciones polinucleotídicas que comprenden cada una un polinucleótido de la pluralidad de polinucleótidos y una región de identificación mediante el acoplamiento del polinucleótido a la región de identificación, en donde la secuencia de cada región de identificación es distinta de la secuencia de las regiones de identificación acopladas a los otros polinucleótidos en la biblioteca.

55 Como se describe en el presente documento, obtener la biblioteca de polinucleótidos comprende obtener una pluralidad de células B y procesar las células para preparar la biblioteca de polinucleótidos. En algunos aspectos,

obtener la biblioteca de polinucleótidos comprende recibir la biblioteca de polinucleótidos directa o indirectamente de un tercero que ha procesado una pluralidad de células B para preparar la biblioteca de polinucleótidos.

5 En el presente documento también se describe un método para producir dos o más composiciones polinucleotídicas, que comprende: obtener una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos asociados con una pluralidad de muestras obtenidas de uno o más sujetos, en donde uno o más polinucleótidos comprenden una variable derivada de células B región de inmunoglobulina, y en donde cada polinucleótido asociado con cada muestra está presente en un recipiente separado; y generar dos o más composiciones polinucleotídicas que comprenden cada una un polinucleótido de la pluralidad de polinucleótidos y una región de identificación mediante el acoplamiento del polinucleótido a la región de identificación, en donde la secuencia de cada región de identificación es distinta de la secuencia de las regiones de identificación acopladas a los otros polinucleótidos en la biblioteca.

10 En algunos aspectos, obtener la biblioteca de polinucleótidos comprende obtener una pluralidad de células B y procesar las células para preparar la biblioteca de polinucleótidos. En algunos aspectos, obtener la biblioteca de polinucleótidos comprende recibir la biblioteca de polinucleótidos directa o indirectamente de un tercero que ha procesado una pluralidad de células B para preparar la biblioteca de polinucleótidos.

15 En el presente documento también se describe una biblioteca de composiciones de polinucleótidos que comprende una pluralidad de composiciones de polinucleótidos, en donde cada composición está presente en un recipiente separado, en donde cada composición comprende una única región de ADNc derivada de muestra que comprende el nucleótido C en el extremo 3' de la región de ADNc y una región de adaptador de identificación de muestra que comprende una región de identificación de muestra acoplada a una región de adaptador, en donde la secuencia de nucleótidos de la región de identificación de muestra de cada región de adaptador de identificación de muestra es distinta de la secuencia de nucleótidos de la región de identificación de muestra de la otra identificación de muestra regiones adaptadoras presentes en cada contenedor separado en la biblioteca, en donde la región adaptadora comprende el nucleótido G en el extremo 3' de la región adaptadora, y en donde la región adaptadora de identificación de muestras está unida a la región de ADNc mediante la unión entre C y G .

20 25 En algunos aspectos, la región de ADNc comprende un polinucleótido de ARN hibridado con un polinucleótido de ADN. En algunos aspectos, la región de ADNc comprende un polinucleótido de ARNm hibridado a un polinucleótido de ADNc. En algunos aspectos, la región de ADNc comprende al menos una C en el extremo 3' y en donde la región del adaptador comprende al menos una G en el extremo 3'.

30 En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una región de identificación de muestra, una región de adaptador y una única región de ADNc derivada de muestra, en donde el extremo 3' de la región de identificación de muestra está acoplado a la 5' extremo de la región del adaptador, en donde la región de ADNc está acoplada al extremo 3' de la región del adaptador, en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única, y en donde la región de identificación de la muestra es de doble cadena. En algunos aspectos, cada polinucleótido comprende una pluralidad de regiones de identificación de muestras.

35 40 En este documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una región de cebador universal, una región de identificación de muestra, una región de adaptador y una región de amplicón de una sola muestra, en donde el extremo 3' de la región de cebador universal está acoplada al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, en donde la región del amplicón está acoplada operativamente a la región del adaptador, en donde la secuencia la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

45 50 En algunos aspectos, el extremo 5' de la región del amplicón está acoplada al extremo 3' de la región del adaptador, en donde la región del cebador universal comprende la secuencia CACGACCGGTGCTCGATTTAG, y en donde la región del adaptador comprende al menos un G. En algunos aspectos, la secuencia de la región del cebador universal no es completamente complementaria a ningún exón de gen humano, y en donde la región del cebador universal tiene una estructura secundaria mínima que no interfiere con la región del adaptador. En algunos aspectos, la región del cebador universal es la secuencia CACGACCGGTGCTCGATTTAG. En algunos aspectos, la región del amplicón comprende una región de ADNc que comprende una secuencia de nucleótidos de ADNc. En algunos aspectos, la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única difiere en al menos 1 nucleótido de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra. En algunos aspectos, la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única difiere en al menos 2 nucleótidos de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más

muestras distintas de la primera muestra. En algunos aspectos, la región de identificación de la muestra tiene al menos 10 nucleótidos de longitud. En algunos aspectos, la región de identificación de la muestra tiene al menos 1 nucleótido de longitud. En algunos aspectos, la secuencia de cada región de identificación de muestra se selecciona de las Tablas 2 y 7. En algunos aspectos, la secuencia de la región adaptadora comprende al menos un nucleótido G en su extremo 3'. En algunos aspectos, la región de amplicón comprende una secuencia de amplicón de cadena pesada de inmunoglobulina, una secuencia de amplicón de cadena liviana de inmunoglobulina, una secuencia de amplicón alfa del receptor de células T, o una secuencia de amplicón beta del receptor de células T.

En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende la secuencia 5'-ABCD-3', en donde A es una región de cebador universal, en donde B es una región de identificación de muestra, en donde C es una región adaptadora, en donde D es una región de amplicón de una sola muestra, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

En este documento también se describe un polinucleótido que comprende una región de cebador universal, una región de identificación de muestra, una región de adaptador y una región de amplicón de una sola muestra, en donde el extremo 3' de la región de cebador universal está acoplada al extremo 5' de la muestra región de identificación, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, y en donde la región del amplicón está acoplada operativamente a la región del adaptador.

En algunos aspectos, el extremo 5' de la región del amplicón está acoplada al extremo 3' de la región del adaptador, en donde la región del cebador universal comprende CACGACCGGTGCTCGATTTAG, y en donde la región del adaptador comprende al menos una G.

En el presente documento también se describe un polinucleótido que comprende la secuencia 5'-A-B-C-D-3', en donde A es una región de cebador universal, en donde B es una región de identificación de muestra, en donde C es una región adaptadora, y en donde D es una región de amplicón de una sola muestra.

En este documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una primera región de identificación de placa, una región de cebador universal, una región de identificación de muestra, una región adaptadora y una región de amplicón de una sola muestra, en donde el extremo de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, en donde la primera región de identificación de la placa está acoplada operativamente a la región del cebador universal, en donde la región del amplicón está acoplada operativamente a la región del adaptador, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

En algunos aspectos, la secuencia de la primera región de identificación de placa de cada polinucleótido de un primer conjunto de muestras individuales es distinta de la secuencia de la primera región de identificación de placa de los otros polinucleótidos en la biblioteca de uno o más conjuntos de muestra única distinta del primer conjunto de muestras individuales. En algunos aspectos, la secuencia de la primera región de identificación de placa de cada polinucleótido del primer conjunto de muestras individuales difiere en al menos 1 nucleótido de la secuencia de la primera región de identificación de la placa de los otros polinucleótidos en la biblioteca de uno o más conjuntos de muestras individuales distintos del primer conjunto de muestras individuales.

En algunos aspectos, la secuencia de la primera región de identificación de placa de cada polinucleótido del primer conjunto de muestras únicas difiere en al menos 2 nucleótidos de la secuencia de la primera región de identificación de placa de los otros polinucleótidos en la biblioteca de uno o más conjuntos de muestras individuales distintos del primer conjunto de muestras individuales. En algunos aspectos, la primera región de identificación de la placa tiene una longitud de al menos 10 nucleótidos. En algunos aspectos, la secuencia de la primera región de identificación de la placa se selecciona de las Tablas 3 y 6. En algunos aspectos, el extremo 3' de la primera región de identificación de la placa se acopla al extremo 5' de la región del cebador universal, en donde la 5 el extremo de la región del amplicón está acoplada al extremo 3 de la región del adaptador, en donde la región del cebador universal comprende CACGACCGGTGCTCGATTTAG, en donde la región del adaptador comprende al menos un G, en donde cada polinucleótido comprende además una segunda región de identificación de la placa, una primera secuenciación región, y una segunda región de secuenciación, en donde el extremo 5' de la segunda región de identificación de placa está acoplada al extremo 3' de la región de amplicón, en donde el extremo 3' de la primera región de secuenciación está acoplada al extremo 5' de la región primera región de identificación de la placa, y en donde el extremo 5' de la segunda región de secuenciación está acoplado al extremo 3' de la segunda región de identificación de la placa. En algunos aspectos, la secuencia de la segunda región de identificación de la placa es idéntica a la secuencia de la primera

región de identificación de la placa en cada polinucleótido. En algunos aspectos, la secuencia de la segunda región de identificación de placa de cada polinucleótido de un primer conjunto de muestras individuales es distinta de la secuencia de la segunda región de identificación de placa de los otros polinucleótidos en la biblioteca de uno o más conjuntos de muestra única distinta del primer conjunto de muestras individuales. En algunos aspectos, la secuencia de la segunda región de identificación de placa de cada polinucleótido del primer conjunto de muestras únicas difiere en al menos 1 nucleótido de la secuencia de la segunda región de identificación de placa de los otros polinucleótidos en la biblioteca de uno o más conjuntos de muestras individuales distintos del primer conjunto de muestras individuales. En algunos aspectos, la secuencia de la segunda región de identificación de placa de cada polinucleótido del primer conjunto de muestras únicas difiere en al menos 2 nucleótidos de la secuencia de la segunda región de identificación de placa de los otros polinucleótidos en la biblioteca de una o más muestras individuales conjuntos distintos del primer conjunto de muestras individuales. En algunos aspectos, la segunda región de identificación de la placa tiene una longitud de al menos 10 nucleótidos. En algunos aspectos, la secuencia de la segunda región de identificación de placa se selecciona de las Tablas 3 y 6. En algunos aspectos, la primera región de secuenciación comprende GAGAGACTGACAGCGTATCGCCTCCCTCGCGCCATCAG. En algunos aspectos, la segunda región de secuenciación comprende CTATGCGCCTTGCCAGCCCGCTCAG.

En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende la secuencia 5'-A-B-C-D-E-3', en donde A es una región de identificación de placa, en donde B es una región de cebador universal, en donde C es una región de identificación de muestra, en donde D es una región adaptadora, en donde E es una región de amplicón de una sola muestra, y en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinto de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

En el presente documento también se describe un polinucleótido que comprende una primera región de identificación de placa, una región de cebador universal, una región de identificación de muestra, una región adaptadora y una región de amplicón de una sola muestra, en donde el extremo 3' de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, donde la primera región de identificación de la placa está acoplada operativamente a la región del cebador universal, y en donde la región de amplicón está acoplada operativamente a la región del adaptador.

En algunos aspectos, el extremo 3' de la primera región de identificación de la placa está acoplado al extremo 5' de la región del cebador universal, en donde el extremo 5' de la región del amplicón está acoplada al extremo 3' de la región del adaptador, en donde la región del cebador universal comprende CACGACCGGTGCTCGATTAG, en donde la región del adaptador comprende al menos un G, en donde cada polinucleótido comprende además una segunda región de identificación de placa, una primera región de secuenciación, y una segunda región de secuenciación, en donde el extremo 5' de la segunda región de identificación de placa es acoplado al extremo 3' de la región del amplicón, en donde el extremo 3' de la primera región de secuenciación está acoplado al extremo 5' de la primera región de identificación de placa, y en donde el extremo 5' de la segunda región de secuenciación está acoplado al extremo 3' de la segunda región de identificación de placa.

En el presente documento también se describe un polinucleótido que comprende la secuencia 5'-A-B-C-D-E-3', en donde A es una región de identificación de la placa, en donde B es una región de cebador universal, en donde C es una región de identificación de la muestra, en donde D es una región adaptadora, y en donde E es una región de amplicón de una sola muestra.

En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una primera región de sitio de restricción, una región de cebador universal, una región de identificación de muestra, una región de adaptador, una región de amplicón de una sola muestra y un segundo sitio de restricción región, en donde el extremo 3' de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, en donde el primer la región del sitio de restricción está acoplada operativamente a la región del cebador universal, en donde la región del amplicón está acoplada operativamente a la región del adaptador, en donde la segunda región del sitio de restricción está acoplada operativamente a la región del amplicón, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinto de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

En algunos aspectos, la primera región del sitio de restricción comprende uno o más sitios de restricción. En algunos aspectos, la primera región del sitio de restricción comprende uno o más sitios de restricción seleccionados del grupo que consiste en: NheI, XhoI, BstBI, EcoRI, SacII, BbvCI, PspXI, AgeI, ApaI, KpnI, Acc65I, XmaI, BstEII, DraIII, PaeI, FseI, AsiSI y AscI. En algunos aspectos, la segunda región del sitio de restricción comprende uno o más sitios de

- restricción. En algunos aspectos, la segunda región del sitio de restricción comprende uno o más sitios de restricción seleccionados del grupo que consiste en: NheI, XhoI, BstBI, EcoRI, SacII, BbvCI, PspXI, AgeI, ApaI, KpnI, Acc65I, XmaI, BstEII, DraIII, PaeI, FseI, AsiSI y AscI. En algunos aspectos, el extremo 3' de la primera región del sitio de restricción está acoplado al extremo 5' de la región del cebador universal, en donde el extremo 3' de la región del adaptador está acoplado al extremo 5' de la región del amplicón, en donde el extremo 3' de la región del amplicón está acoplada al extremo 5' de la segunda región del sitio de restricción, en donde el extremo 3' de la región del amplicón está acoplada al extremo 5' de la segunda región del sitio de restricción, en donde la región del cebador universal comprende CACGACCGGTGCTCGATTTAG, y en donde la región del adaptador comprende al menos un G.
- 10 En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende la secuencia 5'-ABCDEF-3', en donde A es una primera región de sitio de restricción, en donde B es una región de cebador universal, en donde C es una identificación de muestra la región, en donde D es una región adaptadora, en donde E es una región de amplicón de una sola muestra, en donde F es una segunda región del sitio de restricción, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.
- 15 En el presente documento también se describe un polinucleótido para la inserción en un vector, que comprende una primera región de sitio de restricción, una región de cebador universal, una región de identificación de muestra, una región de adaptador, una región de amplicón de una sola muestra y una segunda región de sitio de restricción, en donde el extremo 3' de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, en donde la primera región del sitio de restricción está acoplada operativamente a la región del cebador universal, en donde la región amplicón está acoplada operativamente a la región adaptadora, y en donde la segunda región del sitio de restricción está acoplada operativamente a la región amplicón.
- 20 En algunos aspectos de la divulgación, el extremo 3' de la primera región del sitio de restricción está acoplado al extremo 5' de la región del cebador universal, en donde el extremo 3' de la región del adaptador está acoplado al extremo 5' de la región del amplicón, en donde el extremo 3' de la región del amplicón está acoplada al extremo 5' de la segunda región del sitio de restricción, en donde la región del cebador universal comprende CACGACCGGTGCTCGATTTAG, y en donde la región del adaptador comprende al menos una G.
- 25 En el presente documento también se describe un polinucleótido para la inserción en un vector, que comprende la secuencia 5'-ABCDEF-3', en donde A es una primera región de sitio de restricción, en donde B es una región de cebador universal, en donde C es una región de identificación de muestra, en donde D es una región adaptadora, en donde E es una región de amplicón de una sola muestra, y en donde F es una segunda región de sitio de restricción.
- 30 En el presente documento también se describe una molécula adaptadora de polinucleótido, que comprende una región de cebador universal, una región de identificación de muestra y una región de adaptador, en donde el extremo 3' de la región de cebador universal está acoplada al extremo 5' de la región de identificación de muestra, y en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador. En algunos aspectos, la región del cebador universal comprende CACGACCGGTGCTCGATTTAG, y en donde la región del adaptador comprende al menos un G.
- 35 En el presente documento también se describe un cebador polinucleotídico, que comprende una región de cebador universal y una región de identificación de placa, y en donde el extremo 3' de la región de identificación de placa está acoplado al extremo 5' de la región de cebador universal. En algunos aspectos, la región del cebador universal comprende CACGACCGGTGCTCGATTTAG, en donde el cebador comprende además una región de secuenciación, y en donde el extremo 3' de la región de secuenciación está acoplado al extremo 5' de la región de identificación de la placa.
- 40 En el presente documento también se describe un vector que comprende un polinucleótido descrito en el presente documento. En algunos aspectos, el vector comprende una pluralidad de polinucleótidos. En algunos aspectos, el vector se selecciona del grupo que consiste en: pEE6.4 y pEE12.4
- 45 En el presente documento también se describe una célula huésped aislada que comprende un vector descrito en este documento o un polinucleótido descrito en este documento. En algunos aspectos, la célula huésped se selecciona del grupo que consiste en: células CHO, células CHO-K1, células CHO-S, células NS0, células CHO que son células CHO dhfr-, CHO-dhfr-, DUKX-B11 y células DG44 CHO.
- 50 En el presente documento también se describe un método para producir uno o más polinucleótidos de interés, que comprende: obtener una biblioteca de ADNc que comprende una pluralidad de ADNc asociados con una pluralidad de muestras obtenidas de uno o más sujetos, en donde cada ADNc está asociado con una sola muestra en la pluralidad



de muestras, y en donde cada ADNc asociado con cada muestra está presente en un contenedor separado; y agregar una molécula adaptadora al ADNc asociado con cada muestra para producir uno o más polinucleótidos de interés, en donde la molécula adaptadora comprende una región de identificación de muestra y una región adaptadora, en donde el extremo 3' de la región de identificación de muestra está acoplado a la 5' final de la región del adaptador, y en donde la secuencia de la región de identificación de la muestra de cada molécula del adaptador es distinta de la secuencia de la región de la identificación de la muestra de las otras moléculas del adaptador agregadas a cada ADNc en la biblioteca.

Como se describe aquí, el método incluye además permitir que el extremo 3' de la región del adaptador se una al extremo 3' de cada ADNc en la biblioteca para producir uno o más polinucleótidos de interés. Como se describe en este documento, obtener la biblioteca de ADNc comprende obtener la pluralidad de muestras y procesar las muestras para preparar la biblioteca de ADNc. Como se describe en el presente documento, la molécula adaptadora comprende además una región de cebador universal, en donde el extremo 3' de la región de cebador universal está acoplada al extremo 5' de la región de identificación de la muestra. Como se describe en el presente documento, cada región de ADNc comprende un polinucleótido de ARNm hibridado con un polinucleótido de ADNc. Como se describe en el presente documento, cada muestra comprende una célula. Como se describe en el presente documento, la célula es una célula B. Como se describe en el presente documento, la célula B es un plasmoblasto. Como se describe en el presente documento, cada muestra comprende una pluralidad de células. Como se describe en este documento, obtener la biblioteca de ADNc comprende recibir la biblioteca de ADNc directa o indirectamente de un tercero que ha procesado la pluralidad de muestras para preparar la biblioteca de ADNc. Como se describe aquí, el adaptador se agrega al fusiónamiento del adaptador a la cola '3 de un ADNc generado durante una reacción de transcripción reversa. Como se describe en este documento, cada ADNc comprende al menos un nucleótido C, en donde C está ubicado en el extremo 3' de cada ADNc, en donde la región del adaptador comprende al menos un nucleótido G, en donde G está ubicado en el extremo 3' de la región del adaptador y en donde la región del adaptador está unida a cada ADNc a través de la unión entre G y C. Como se describe en este documento, la molécula adaptadora es de cadena sencilla, y comprende además incorporar la molécula adaptadora en cada ADNc permitiendo que una enzima produzca la molécula adaptadora doble cadena. Como se describe en el presente documento, la molécula adaptadora se incorpora a cada ADNc para producir el polinucleótido de interés mediante una transcriptasa reversa MMLV H<sup>-</sup>

En el presente documento también se describe un método para producir uno o más polinucleótidos de interés para la secuenciación, que comprende: obtener una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una región de cebador universal, una región de identificación de la muestra, una región adaptadora y un amplicón región de una sola muestra, en donde el extremo 3' de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, y en donde la región de amplicón está acoplada operativamente a la región adaptadora, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distinta de la primera muestra única; y amplificar la biblioteca de polinucleótidos con un conjunto de cebadores para producir uno o más polinucleótidos de interés para la secuenciación, en donde el uno o más polinucleótidos de interés para la secuenciación comprende una primera región de secuenciación, una región de identificación de la primera placa, una región de cebador universal, una región de identificación de la muestra, una región adaptadora, una región de amplicón de una sola muestra y una segunda región de secuenciación, en donde el extremo 3' de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' la región de identificación de la muestra está acoplada al extremo 5' de la región del adaptador, en donde la primera región de identificación de la placa está acoplada operativamente a la región del cebador universal, en donde la región del amplicón está acoplada operativamente a la región del adaptador, en donde la primera región de secuenciación es ubicado en el extremo 5' del polinucleótido de interés, y en donde la segunda región de secuenciación está ubicada en el extremo 3' del polinucleótido de interés.

Como se describe en el presente documento, el método incluye además la secuenciación de uno o más polinucleótidos de interés. En algunos aspectos, el método incluye además la secuenciación de uno o más polinucleótidos de interés con la secuenciación 454. Como se describe en el presente documento, el método incluye además la secuenciación de uno o más polinucleótidos de interés con la secuenciación SMRT. Como se describe en el presente documento, el método incluye además la secuenciación de uno o más polinucleótidos de interés con la secuenciación SOLiD. Como se describe en este documento, el método incluye además la secuenciación de uno o más polinucleótidos de interés con la secuenciación SOLEXA. Como se describe en el presente documento, el método incluye además la secuenciación de uno o más polinucleótidos de interés con la secuenciación tSMS. Como se describe en el presente documento, el conjunto de cebadores se selecciona de entre los cebadores que se muestran en las Tablas 1 y 5. Como se describe en el presente documento, la obtención de la biblioteca de polinucleótidos comprende la preparación de la biblioteca de polinucleótidos en un laboratorio. Como se describe en este documento, obtener la biblioteca de polinucleótidos comprende recibir la biblioteca de polinucleótidos directa o indirectamente de un tercero que ha preparado la biblioteca de polinucleótidos.

En el presente documento también se describe un método para analizar datos de secuenciación, que comprende: obtener un conjunto de datos asociado con una pluralidad de polinucleótidos, en donde el conjunto de datos comprende datos de secuenciación para la pluralidad de polinucleótidos, en donde cada polinucleótido en la pluralidad de polinucleótidos comprende una región de identificación de muestra, y en donde cada región de identificación de muestra en cada polinucleótido es única para una sola muestra, en donde la secuencia de la región de identificación de muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de muestra de los otros polinucleótidos en la pluralidad de polinucleótidos de una o más muestras distintas de la primera muestra única; y analizar el conjunto de datos para hacer coincidir polinucleótidos con regiones idénticas de identificación de la muestra, en donde una coincidencia indica que los polinucleótidos se originaron en la misma muestra.

Como se describe aquí, cada polinucleótido en la pluralidad de polinucleótidos comprende además una primera región de identificación de la placa, en donde cada combinación de cada primera región de identificación de la placa y la región de identificación de la muestra en cada polinucleótido es única para una sola muestra, en donde la secuencia de la primera identificación de la placa la región de cada polinucleótido de un primer conjunto de muestras únicas es distinta de la secuencia de la primera región de identificación de placas de los otros polinucleótidos en la pluralidad de polinucleótidos de uno o más conjuntos de muestras individuales distinta del primer conjunto de muestras individuales, y que comprende además analizar el conjunto de datos para hacer coincidir polinucleótidos con regiones idénticas de identificación de la primera placa y regiones idénticas de identificación de la muestra, en donde una coincidencia entre ambas regiones indica que los polinucleótidos se originaron en la misma muestra. Como se describe aquí, obtener el conjunto de datos comprende obtener la pluralidad de polinucleótidos y secuenciar la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. Como se describe aquí, obtener el conjunto de datos comprende recibir el conjunto de datos directa o indirectamente de un tercero que ha secuenciado la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. Como se describe aquí, el conjunto de datos se almacena en un medio de almacenamiento electrónico. Como se describe en este documento, la muestra única es una sola célula. Como se describe en el presente documento, la muestra única comprende una sola célula. Como se describe en el presente documento, la muestra única comprende una única célula B. Como se describe en el presente documento, la muestra única comprende una pluralidad de células B. Como se describe en el presente documento, que comprende además seleccionar uno o más polinucleótidos para la clonación.

En el presente documento también se describe un método para identificar un segundo polinucleótido de interés con base en la selección de un primer polinucleótido de interés, que comprende: obtener un conjunto de datos asociado con una pluralidad de polinucleótidos, en donde el conjunto de datos comprende datos de secuenciación para la pluralidad de polinucleótidos, en donde cada polinucleótido en la pluralidad de polinucleótidos comprende una región de identificación de muestra, y en donde cada región de identificación de muestra en cada polinucleótido es única a una muestra única, asociando así cada polinucleótido en la pluralidad de polinucleótidos con una muestra única distinta, en donde la secuencia de la región de identificación de muestra de cada polinucleótido de una primera muestra única es distinto de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la pluralidad de polinucleótidos de una o más muestras distintas de la primera muestra única; y seleccionando un primer polinucleótido de interés asociado con una primera muestra única del conjunto de datos e identificando un segundo polinucleótido de interés en la primera muestra única en base a la región de identificación de la muestra del primer polinucleótido de interés.

Como se describe aquí, cada polinucleótido en la pluralidad de polinucleótidos comprende además una primera región de identificación de la placa, en donde cada combinación de cada primera región de identificación de la placa y la región de identificación de la muestra en cada polinucleótido es única para una sola muestra, en donde la secuencia de la primera identificación de la placa la región de cada polinucleótido de un primer conjunto de muestras únicas es distinta de la secuencia de la primera región de identificación de placas de los otros polinucleótidos en la pluralidad de polinucleótidos de uno o más conjuntos de muestras individuales distinta del primer conjunto de muestras individuales, y que comprende además identificar un segundo polinucleótido de interés en la primera muestra única en base a la región de identificación de la muestra y la primera región de identificación de la placa del primer polinucleótido de interés. Como se describe en el presente documento, la primera muestra única comprende una célula B. Como se describe en el presente documento, la primera muestra única comprende una pluralidad de células B. Como se describe en el presente documento, la primera muestra única comprende una célula B, en donde el primer polinucleótido de interés comprende una secuencia de nucleótidos de cadena pesada de anticuerpo, y en donde el segundo polinucleótido de interés comprende una secuencia de nucleótidos de cadena liviana de anticuerpo. Como se describe en el presente documento, la primera muestra única comprende una célula B, en donde el primer polinucleótido de interés comprende una secuencia de nucleótidos de cadena liviana de anticuerpo, y en donde el segundo polinucleótido de interés comprende una secuencia de nucleótidos de cadena pesada de anticuerpo. Como se describe aquí, obtener el conjunto de datos comprende obtener la pluralidad de polinucleótidos y secuenciar la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. Como se describe aquí, obtener el conjunto de datos comprende recibir el conjunto de datos directa o indirectamente de un tercero que ha secuenciado la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. Como se describe aquí, el conjunto de datos se almacena en un medio de almacenamiento electrónico.

En este documento también se describe un método para producir uno o más polinucleótidos de interés para la clonación, que comprende: obtener una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una región de cebador universal, una región de identificación de la muestra, una región adaptadora y un amplicón región de una sola muestra, en donde el extremo 3' de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador, y en donde la región de amplicón está acoplada operativamente a la región adaptadora, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única; y amplificar la biblioteca de polinucleótidos con un conjunto de cebadores para producir uno o más polinucleótidos de interés para la clonación, en donde el uno o más polinucleótidos de interés para la clonación comprende una primera región de sitio de restricción, una región de cebador universal, una región de identificación de muestra, una región adaptadora, una región de amplicón de una sola muestra y una segunda región de sitio de restricción, en donde el extremo 3' de la región del cebador universal está acoplado al extremo 5' de la región de identificación de la muestra, en donde el extremo 3' de la identificación de la muestra la región está acoplada al extremo 5' de la región del adaptador, en donde la región del amplicón está acoplada operativamente a la región del adaptador, en donde la primera región del sitio de restricción está ubicada en el extremo 5' del polinucleótido de interés, y en donde la segunda región del sitio de restricción está situada en el extremo 3' del polinucleótido de interés.

Como se describe aquí, obtener la biblioteca de polinucleótidos comprende preparar la biblioteca de polinucleótidos en un laboratorio. Como se describe en este documento, obtener la biblioteca de polinucleótidos comprende recibir la biblioteca de polinucleótidos directa o indirectamente de un tercero que ha preparado la biblioteca de polinucleótidos.

En el presente documento también se describe un método para producir una molécula de interés, que comprende: obtener una célula huésped que comprende un polinucleótido que comprende una región de identificación de muestra, una región de adaptador y una región de amplicón de una sola muestra, en donde el extremo 3' de la identificación de muestra la región está acoplada al extremo 5' de la región del adaptador, y en donde la región del amplicón está acoplada operativamente a la región del adaptador; y cultivar la célula huésped en condiciones suficientes para producir la molécula de interés. Como se describe en el presente documento, la obtención de la célula huésped comprende la preparación de la célula huésped que comprende el polinucleótido en un laboratorio. Como se describe en este documento, obtener la célula huésped comprende recibir la célula huésped que comprende el polinucleótido directa o indirectamente de un tercero que ha preparado la célula huésped. Como se describe aquí, la molécula de interés es una proteína. Como se describe en el presente documento, la molécula de interés es un anticuerpo. Como se describe en el presente documento, la molécula de interés es un anticuerpo monoclonal humano. Como se describe en el presente documento, que comprende además recoger la molécula de interés.

En el presente documento también se describe un kit que comprende un polinucleótido, una biblioteca de polinucleótidos, un vector o una célula huésped descrita en el presente documento e instrucciones de uso.

En el presente documento también se describe un método para enlazar y codificar con barras una pluralidad de secuencias de polinucleótidos no contiguas de interés, comprendiendo dicho método: (a) proporcionar una pluralidad de moléculas de ADNc; (b) unir físicamente las moléculas de ADNc de interés; y (c) agregar una secuencia de código de barras a los ADNc de interés antes, durante o después del enlace físico.

En algunos aspectos de la divulgación, el enlace físico es por ligación. En algunos aspectos de la divulgación, el enlace físico es por recombinación. En algunos aspectos de la divulgación, el enlace físico comprende el uso de una secuencia de superposición y extensión. En algunos aspectos de la divulgación, la secuencia del código de barras está ubicada en uno o ambos extremos de los ADNc físicamente unidos. En algunos aspectos de la divulgación, la secuencia del código de barras se ubica entre los ADNc físicamente unidos. En algunos aspectos de la divulgación, la ligación se realiza por fusionamiento y ligación de extremos compatibles. En algunos aspectos de la divulgación, los fines compatibles son un sitio de restricción. En algunos aspectos de la divulgación, la ligación se realiza por ligación de extremo romo. En algunos aspectos de la divulgación, la secuencia de superposición-extensión se agrega durante el curso de la amplificación utilizando un cebador que comprende la cola de superposición-extensión. En algunos aspectos de la divulgación, la secuencia de superposición-extensión se agrega durante el curso de la transcripción reversa utilizando un cebador que comprende la cola de superposición-extensión. En algunos aspectos de la divulgación, la secuencia de superposición-extensión se agrega por fusionamiento de un adaptador a la cola 3' de un ADNc generado durante una reacción de transcripción reversa. En algunos aspectos de la divulgación, la secuencia del código de barras se agrega por ligación. En algunos aspectos de la divulgación, la ligación se realiza por fusionamiento y ligación de extremos compatibles. En algunos aspectos, los extremos compatibles son un sitio de restricción. En algunos aspectos de la divulgación, la ligación se realiza por ligación de extremo romo de un adaptador que comprende la secuencia de código de barras. En algunos aspectos de la divulgación, la secuencia del código de barras se agrega durante el curso de una reacción de amplificación utilizando un cebador que comprende la secuencia del código de barras. En algunos aspectos de la divulgación, la secuencia del código de barras se agrega durante el curso de una reacción de transcripción reversa utilizando un cebador que comprende la secuencia del código de barras. En algunos aspectos de la divulgación, la secuencia del código de barras se agrega al recocer un adaptador

en la cola 3' de un ADNc generado durante una reacción de transcripción reversa. En algunos aspectos de la divulgación, el extremo '3 del ADNc comprende al menos un nucleótido C, y en donde el extremo 3' del adaptador comprende al menos un nucleótido G, y en donde el adaptador está unido a cada ADNc a través de la unión entre C y G. Como se describe en este documento, el adaptador es de una sola hebra, y además comprende incorporar el adaptador en cada ADNc permitiendo que una enzima haga que el adaptador sea de doble cadena. Como se describe en este documento, el adaptador se incorpora a cada ADNc mediante una transcriptasa reversa MMLV H<sup>-</sup>. Como se describe aquí, la secuencia de superposición-extensión comprende una secuencia de código de barras. Como se describe en el presente documento, las secuencias de polinucleótidos de interés comprenden cadenas pesadas y livianas de anticuerpos. Como se describe en el presente documento, que comprende además (d) agregar una región de secuenciación a los ADNc de interés antes, durante o después del enlace físico. Como se describe en este documento, la región de secuenciación se agrega con un adaptador. Como se describe en el presente documento, que comprende además (e) la secuenciación de las moléculas de ADNc de interés físicamente unidas utilizando una plataforma de secuenciación NextGen. Como se describe en este documento, la plataforma de secuenciación NextGen es la secuenciación 454. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SMRT. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SOLiD. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SOLEXA. Como se describe aquí, la plataforma de secuenciación de NextGen es la secuenciación tSMS. Como se describe aquí, la pluralidad de moléculas de ADNc proviene de muestras individuales contenidas en una placa con al menos 6 pozos, al menos 12 pozos, al menos 24 pozos, al menos 48 pozos, al menos 96 pozos, al menos 384 pozos, al menos 1536 pozos, o más pozos. Como se describe en el presente documento, la pluralidad de moléculas de ADNc proviene de muestras individuales contenidas en al menos una, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, veinte, treinta, cuarenta, cincuenta, setenta y cinco, cien, o más placas con al menos 96 pozos cada una.

En el presente documento también se describe un método para enlazar y codificar con barras una pluralidad de muestras que contienen secuencias de polinucleótidos de interés, comprendiendo dicho método: (a) distribuir las muestras en una pluralidad de recipientes; (b) sintetizar secuencias polinucleotídicas de interés usando plantillas de la muestra, en donde dicha síntesis da como resultado la adición de una secuencia de código de barras; y (c) efectuar el enlace de las secuencias de polinucleótidos de interés sintetizadas en la etapa (b).

Como se describe en el presente documento, cada muestra comprende una célula. Como se describe en el presente documento, la célula es una célula B. Como se describe en el presente documento, la célula B es un plasmoblasto, una célula B de memoria o una célula plasmática. Como se describe en el presente documento, cada muestra comprende una pluralidad de células. Como se describe en el presente documento, las secuencias de polinucleótidos de interés comprenden cadenas pesadas y livianas de anticuerpos. Como se describe en el presente documento, dicha síntesis comprende una amplificación por RT-PCR. Como se describe en el presente documento, dicha amplificación de RT-PCR se realiza en un solo paso. Como se describe en el presente documento, dicho enlace del polinucleótido de interés se realiza durante el curso de una amplificación por RT-PCR utilizando un cebador de superposición-extensión. Como se describe en el presente documento, que comprende además (d) agregar una región de secuenciación a las secuencias de polinucleótidos de interés antes, durante o después de la adición o enlace de la secuencia de códigos de barras. Como se describe en este documento, la región de secuenciación se agrega con un adaptador. Como se describe en el presente documento, que comprende además (e) la secuenciación de las secuencias de polinucleótidos enlazadas de interés utilizando una plataforma de secuenciación NextGen. Como se describe en este documento, la plataforma de secuenciación NextGen es la secuenciación 454. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SMRT. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SOLiD. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SOLEXA. Como se describe aquí, la plataforma de secuenciación de NextGen es la secuenciación tSMS. Como se describe aquí, la pluralidad de muestras son muestras individuales contenidas en una placa con al menos 6 pozos, al menos 12 pozos, al menos 24 pozos, al menos 48 pozos, al menos 96 pozos, al menos 384 pozos, al menos 1536 pozos, o más pozos. Como se describe en el presente documento, la pluralidad de muestras son muestras individuales contenidas en al menos uno, dos, tres, cuatro, cinco, seis, siete, ocho, nueve, diez, veinte, treinta, cuarenta, cincuenta, setenta y cinco, cien, doscientas, quinientas o más placas con al menos 96 pozos cada una.

En el presente documento también se describe un método para enlazar y codificar con barras una pluralidad de secuencias de polinucleótidos no contiguas de interés, comprendiendo dicho método: (a) distribuir células en una pluralidad de recipientes para obtener una o más células aisladas; (b) amplificar secuencias de polinucleótidos de interés usando plantillas de dicho una o más células aisladas, en donde dicha amplificación da como resultado la adición de una secuencia de código de barras; y (c) efectuar el enlace de las secuencias de polinucleótidos de interés amplificadas en la etapa (b).

Como se describe en el presente documento, las secuencias de nucleótidos de interés comprenden cadenas pesadas y livianas de anticuerpos. Como se describe en el presente documento, dicha amplificación comprende una amplificación por RT-PCR. Como se describe en el presente documento, dicha amplificación de RT-PCR se realiza en un solo paso. Como se describe en el presente documento, dicho enlace del nucleótido de interés se realiza durante el curso de una amplificación por RT-PCR utilizando un cebador de superposición-extensión. Como se describe en el presente documento, que comprende además (d) agregar una región de secuenciación a las secuencias de

polinucleótidos de interés antes, durante o después de la adición o enlace de la secuencia de códigos de barras. Como se describe en este documento, la región de secuenciación se agrega con un adaptador. Como se describe en el presente documento, que comprende además (e) la secuenciación de las secuencias de polinucleótidos enlazadas de interés utilizando una plataforma de secuenciación NextGen. Como se describe en este documento, la plataforma de secuenciación NextGen es la secuenciación 454. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SMRT. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SOLiD. Como se describe aquí, la plataforma de secuenciación NextGen es la secuenciación SOLEXA. Como se describe aquí, la plataforma de secuenciación de NextGen es la secuenciación tSMS. Como se describe aquí, la una o más células están contenidas en una placa con al menos 6 pozos, al menos 12 pozos, al menos 24 pozos, al menos 48 pozos, al menos 96 pozos, al menos 384 pozos, al menos 1536 pozos, o más pozos. Como se describe aquí, la una o más células están contenidas en al menos una, dos, tres, cuatro, cinco, seis, siete, ocho, diez, veinte, treinta, cuarenta, cincuenta, setenta y cinco, cien o más placas con al menos 96 pozos cada una.

En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende la secuencia 5'-ABCD-3', en donde A es una región de identificación de muestra (secuencia de código de barras), en donde B es una primera región de ADNc de una muestra única, en donde C es una región enlazadora, en donde D es una segunda región de ADNc de la misma muestra única, y en donde la secuencia de la región de identificación de muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de muestra de la otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende la secuencia 5'-ABCD-3', en donde A es una primera región de ADNc de una sola muestra, en donde B es una región enlazadora, en donde C es una segunda región de ADNc de la misma muestra única, en donde D es una región de identificación de muestra (secuencia de código de barras), y en donde la secuencia de la región de identificación de muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de muestra de la otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

En el presente documento también se describe una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende la secuencia 5'-ABC-3', en donde A es una primera región de ADNc de una muestra única, en donde B es una región enlazadora que comprende una región de identificación de muestra (secuencia de código de barras), en donde C es una segunda región de ADNc de la misma muestra única, y en donde la secuencia de la región de identificación de muestra de cada polinucleótido de una primera muestra única es distinta de la secuencia de la región de identificación de muestra de los otros polinucleótidos en la biblioteca de una o más muestras distintas de la primera muestra única.

En algunos aspectos, dicha primera región de ADNc comprende una cadena pesada de anticuerpo y dicha segunda región de ADNc comprende una cadena liviana de anticuerpo. En algunos aspectos, la biblioteca comprende al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 3000, al menos 10 000, al menos 30 000, al menos 100 000, al menos 300 000, al menos 1 000 000, al menos 3 000 000, al menos 10 000 000, al menos 30 000 000, o más miembros. En algunos aspectos, la biblioteca comprende al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 10 000, al menos 30 000 o más genes del transcriptoma completo de una muestra de células. En algunos aspectos, la biblioteca comprende al menos 1, al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 10 000, al menos 100 000, al menos 1 000 000, al menos 10 000 000, al menos 100 000 000 o más de las diferentes especies de anticuerpos presentes en la sangre de un individuo. En algunos aspectos, los anticuerpos se expresan mediante plasmoblastos, células plasmáticas, células B de memoria, células plasmáticas de larga vida, otras células de linaje B o combinaciones de las mismas.

Breve descripción de las varias vistas de los dibujos

Estas y otras características, aspectos y ventajas se entenderán mejor con respecto a la siguiente descripción y los dibujos adjuntos, donde:

Figura 1. Diferenciación de células B. Las células B ingenuas maduras son CD19<sup>+</sup> y pueden activarse para proliferar y diferenciarse en la exposición antigénica en tejidos linfoides secundarios como los ganglios linfáticos y el bazo. Proliferan y se diferencian en los focos extrafoliculares o en los centros germinales. Al diferenciarse las células B en los focos extrafoliculares suele diferenciarse para convertirse en células plasmáticas de vida corta y, por lo general, residen en el tejido linfoide secundario del que se originaron. Las células B que se diferencian en los centros germinales pueden convertirse en células B de memoria, que pueden estimularse aún más para diferenciarse mediante el desafío antigénico posterior, o convertirse en plasmoblastos que tienen el potencial de convertirse en células plasmáticas de larga vida. Estos plasmoblastos pueden ingresar a la circulación y son transportados a los diversos tejidos en los que residen las células plasmáticas de larga vida, como la médula ósea, los tejidos de la mucosa (para las células de plasma IgA<sup>+</sup>) y los tejidos inflamados. Algunas células plasmáticas en tránsito también están presentes en la sangre. Todos los tipos de células mencionados anteriormente también se pueden encontrar en la circulación en la sangre.

Figura 2. Esquema de la secuenciación, clonación y expresión de alto rendimiento de genes pareados de células clasificadas individuales. Las poblaciones de células deseadas se clasifican por células individuales basándose en su expresión de marcadores de superficie celular en placas de PCR de 96 pozos. Durante la transcripción reversa, las moléculas adaptadoras de ADN con código de barras (ID de muestra) se agregan a la 1ª hebra de ADNc sintetizada utilizando la propiedad de cambio de plantilla de las transcriptasas reversas MMLV-H. Los productos de RT de cada placa se agrupan por separado y se realizan 2 rondas de PCR para amplificar regiones específicas de amplicón (amplicones). La PCR se realiza con cebadores con 5' que flanquean códigos de barras para agregar regiones de identificación de placa (ID de placa) a los amplicones. Los amplicones se envían para secuenciación 454 (a). Las secuencias obtenidas se agrupan por ID de placa e ID de muestra antes del ensamblaje de secuencia. Los amplicones de la misma célula se emparejan utilizando ID de placa e ID de muestra y los clones deseados seleccionados para clonación y expresión (b). Los amplicones específicos se pueden amplificar de cada placa combinada de amplicones utilizando cebadores de clonación que son complementarios a la identificación de la muestra de ese amplicón en particular. Los cebadores de clonación también agregan regiones de sitio de restricción (RS) que luego se usan para insertar el clon en vectores de expresión de mamíferos para la expresión y selección en sentido descendente. En este ejemplo, los amplicones son genes de cadenas pesadas y livianas de inmunoglobulina (Ig) que codifican el anticuerpo. Los anticuerpos expresados se pueden usar para la detección corriente abajo (c). 5' RS y 3' RS: 5' y 3' sitios de restricción respectivamente. HC y LC: cadena pesada y cadena liviana respectivamente.

Figura 3. Esquema de la transcripción reversa y PCR para agregar ID de muestra e ID de placa a los amplicones de Ig. La transcripción reversa (RT) se realizó con Superscript II o Superscript III, que son transcriptasas reversas MMLVH. Estas transcriptasas tienen una actividad de colas en 3' y agregan un par de citosinas al extremo 3' de los ADNc de la 1ª hebra recién sintetizados. Un oligonucleótido que termina con -GGG (una región adaptadora) puede complementar el par de bases con esto y la transcriptasa reversa luego cambia la plantilla al oligonucleótido y continúa con la transcripción, lo que resulta en la incorporación del adaptador de ID de muestra al extremo 3' del ADNc (a). Como el adaptador de ID de muestra contiene una secuencia invariable en 5' (región del cebador universal), se pueden usar cebadores directos complementarios a esta secuencia para las PCR posteriores. La primera PCR se realizó con Cebador largo de avance 1, Cebador corto de avance 1 y GSP1. El Cebador largo de avance 1 tiene una región flanqueante 5' que contiene un código de barras ID de placa y el 454 Cebador A de titanio para la secuenciación 454, que se incorporaron al amplicón. El Cebador corto de avance 1 tiene una Tm similar al primer GSP1 y se incluyó para aumentar ligeramente la eficiencia de la PCR. Cada GSP1 (cebador específico de gen 1) tiene una secuencia específica de gen complementaria y se usó para amplificar un gen específico. Aquí, los cebadores específicos del gen son para las cadenas livianas kappa y lambda y la cadena pesada gamma para amplificar estos genes. Se muestran las secuencias para los cebadores (b). La segunda PCR es una PCR anidada y se realizó con Cebador de avance 2, Cebador largo GSP2 y RV Primer2. Las secuencias para los cebadores son las que se muestran. El Primer GSP2 largo tiene una región flanqueante 5' que contiene un código de barras ID de placa y el Primer B de titanio 454 para la secuenciación 454, que se incorporaron al amplicón. Los cebadores largos GSP2 amplifican de nuevo las cadenas livianas kappa y lambda y la cadena pesada gamma. El RV Primer2 tiene una Tm similar a Cebador de avance 2 y se incluyó para aumentar ligeramente la eficiencia de la PCR. Se muestran las secuencias de los cebadores. Después de las RT y 2 PCR, cada amplicón tendrá 454 cebadores de titanio A y B para la secuenciación de 454, dos ID de placa idénticas, cada una identificando que el amplicón proviene de una placa particular de 96 pozos clasificada por células y una ID de muestra que determina su posición en la placa de 96 pozos (c).

Figura 4. Amplificación exitosa de células B ordenadas unicelulares utilizando cebadores de secuenciación y clonación. Una placa de 96 pozos de células B clasificadas de una sola célula se sometió a transcripción reversa, se agruparon y se amplificaron como se muestra en los esquemas de la Figura 1a y la Figura 2. Se visualizaron las bandas para la cadena liviana kappa, la cadena liviana lambda y la cadena pesada gamma en un gel de agarosa en los tamaños esperados: ~600 pb para kappa y lambda y ~700 pb para gamma (a). El cromatograma de ADN de la secuenciación de Sanger de la cadena pesada gamma del extremo 5' mostró una secuencia 'variable' correspondiente a múltiples ID de muestra para la placa combinada (b). La secuenciación de Sanger de la cadena liviana kappa desde el extremo 3' mostró una secuencia 'variable' después de la región constante y comenzando en la unión VJ correspondiente a múltiples cadenas livianas (c). Se usó un par de cebadores de clonación específico para el pozo A1 para amplificar la cadena liviana kappa. La secuenciación de Sanger mostró que, a diferencia de (c), solo se amplificó una secuencia limpia (d). Todos los resultados son representativos de los otros genes de inmunoglobulina amplificada. M: escalera de ADN de 100pb, K: cadena liviana kappa, L: cadena liviana lambda, G: cadena pesada gamma.

Figura 5. Los pacientes con CCP+ RA tienen porcentajes de plasmoblastos de sangre periférica que se correlacionan con la actividad de la enfermedad y autoanticuerpos anti-citrulina secretados. Los plasmoblastos son CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>CD38<sup>++</sup> y se clasificaron primero en células CD3<sup>-</sup>, luego se clasificaron como se muestra (a). La sangre periférica se obtuvo de pacientes con AR y los plasmoblastos se representaron como un porcentaje del total de PBMC. Una prueba de Mann-Whitney mostró que los pacientes con CCP+ RA poseen porcentajes de plasmoblastos significativamente más altos (p<0.05) que los pacientes con CCP- RA (b). Los porcentajes de plasmoblastos se correlacionaron significativamente con el índice de actividad de la enfermedad clínica (CDAI) en pacientes con CCP+ mediante regresión lineal. Se realizó una regresión lineal en porcentajes de plasmoblastos transformados logarítmicamente para lograr la normalidad del conjunto de datos (c). Los plasmoblastos CCP+ también se clasificaron de forma simulada o se sometieron a una selección de depleción de plasmoblastos con > 95% de eliminación de plasmoblastos (d) y se cultivaron durante 7 días en RPMI suplementado con FBS al 10%. El sobrenadante se recogió

y se analizó para determinar la reactividad del péptido anticitrulinado con Luminex. Las intensidades medias de fluorescencia para la reactividad del anticuerpo contra cada péptido se representaron como reactividades peptídicas (e).

5 Figura 6. Estrategia para la selección y cribado de clones para neutralizar anticuerpos. Las secuencias de anticuerpos LC y HC emparejados se obtienen a partir del análisis bioinformático de los amplicones secuenciados en 454 y se agrupan en familias clonales en función de su uso de LC y HC V(D)J. El clon o los clones que se producen a la frecuencia más alta en cada familia clonal se clonarán, expresarán y seleccionarán selectivamente para determinar su unión al antígeno diana de interés utilizando ELISA. Los clones representativos de toda la familia clonal que segregan anticuerpos de unión se clonarán y expresarán para la selección de anticuerpos neutralizantes. Cada diagrama de anticuerpos representa un clon.

15 Figura 7. Caracterización de las secuencias V (D) J de la cadena pesada de inmunoglobulina y las familias clonales derivadas de sujetos humanos individuales. La sangre se obtuvo de humanos con las siguientes afecciones: (i) osteomielitis crónica por *Staphylococcus aureus* en un humano que no estaba tomando antibióticos (debido al incumplimiento) pero cuyo sistema inmunitario suprimió la infección y previno la infección fulminante durante varios meses sin antibióticos; (ii) un humano con bacteriemia por *Staphylococcus aureus* aguda y fulminante que requiere transferencia a la unidad de cuidados intensivos y tratamiento agresivo con antibióticos por vía intravenosa; (iii) 3 humanos con artritis reumatoide activa crónica (AR) (con puntuaciones de actividad de la enfermedad (DAS) > 5); (iv) un ser humano 7 días después de recibir la vacuna trivalente contra la influenza (Fluzone, Sanofi); y (v) un humano con adenocarcinoma de pulmón metastásico que se esperaba que expirara, pero luego de la quimioterapia entró en un estado de no progresión a largo plazo. En todos los casos, los pacientes humanos mostraron elevaciones en sus niveles de plasmoblastos de sangre periférica (de 1.5 a 6% de las células B de sangre periférica son plasmoblastos [CD20<sup>+</sup>CD19<sup>+</sup>CD38<sup>++</sup>CD27<sup>+</sup>], con niveles en humanos normales de 0.1-0.2% de células B de sangre periférica), lo que indica una respuesta inmune activada. Los plasmoblastos se clasificaron en una sola célula en placas de 96 pozos, y se realizó el código de barras y la secuenciación 454 del ADNc de inmunoglobulina expresado como se describe en las Figuras 2 y 3. Se utilizó el análisis bioinformático para emparejar las inmunoglobulinas de las cadenas pesada y liviana expresadas por plasmoblastos individuales. Se presentan diagramas de sectores del porcentaje de uso de V(D)J de la cadena pesada para los pacientes individuales -cada cuña representa el porcentaje de plasmoblastos que expresan una reorganización de la secuencia V(D)J de la cadena pesada-.

30 Figura 8. La agrupación de las secuencias V(D)J de la cadena pesada de inmunoglobulina de sujetos humanos demuestra familias clonales y subfamilias clonales. Los conjuntos de datos de secuencia de la cadena pesada de inmunoglobulina generados en los estudios descritos en la Figura 7 se sometieron a agrupamiento jerárquico utilizando el programa Clustal. La agrupación jerárquica produjo árboles evolutivos que representan la respuesta de anticuerpos en cada humano individual.

35 Figura 9. Esquema de RT y PCR para agregar identificadores de muestra e identificadores de placa a cualquier amplicón y utilidad descendente. Las muestras individuales de las células individuales o las células múltiples se transcriben por separado en pozos. La transcripción reversa agrega una ID de muestra y una región de cebador universal 5' a todos los ADNc de la primera cadena como se ha descrito anteriormente (a). El ADNc de todos los pozos de una placa se combinan y se ven en 2 rondas de PCR. El 1<sup>er</sup> PCR usa Cebador corto de avance 1, Cebador largo de avance 1 como cebadores directos y agrega un 454 Cebador A de titanio para la secuenciación de 454 y una placa de identificación en el extremo 5' de la secuencia. El Cebador corto de avance 1 tiene una Tm similar al primer GSP1 y se incluyó para aumentar la eficiencia de la PCR. Cada cebador GSP1 tiene una secuencia específica de gen y puede amplificar específicamente ese gen. Se muestran las secuencias de los cebadores. Tenga en cuenta que independientemente de qué gen se amplifica, los cebadores directos permanecen constantes (b). La segunda PCR es una PCR anidada. Cebador de avance 2 es el cebador directo, y los cebadores reversos son Cebador largo GSP2 y cebador Rv 2. Long GSP2 es específico de un gen y solo amplifica un gen específico. También agrega el 454 Primer de titanio B para la secuenciación 454 y una placa de identificación en el extremo 3' del amplicón. RV Primer2 tiene una Tm similar a Cebador de avance 2 y se incluyó para aumentar la eficiencia de la PCR. Se muestran las secuencias de los cebadores. Después de RT y 2 PCR, los amplicones de todas las placas se agrupan y se secuencian en 454. La combinación de ID de placa y ID de muestra permite la identificación de secuencias que se originan en la misma muestra. Esto permite la comparación de secuencias entre múltiples muestras. Las secuencias del mismo origen también se pueden expresar en los pares para obtener la proteína exacta de la célula original, como el receptor de las células y otros isotipos de Ig como IgM, IgE e IgA (c).

55 Figura 10. Cebadores específicos de genes para la transcripción reversa (RT-GSP) de las cadenas pesada y liviana de inmunoglobulina. Se usaron RT-GSP en lugar de oligo (dT) como cebadores en la transcripción reversa de los genes de las cadenas pesada y liviana. El ADNc se amplificó luego por PCR y se visualizó en un gel de agarosa. Los cebadores RT-GSP IgKC\_v3 (a), IgLC\_v5, IgLC\_v6, IgLC\_v7 e IgLC\_v8 en los carriles 1-4 respectivamente (b), IgHGC\_v10, IgHGC\_v11, IgHGC\_v13 e IgGC\_v15 en los carriles 1-4 respectivamente (c) e IgHGC\_v16 (d). KC, LC y GC en los nombres de los cebadores indican que el cebador es específico para la cadena kappa, la cadena lambda y la cadena pesada gamma, respectivamente. Las bandas blancas en las fotos de gel indican dónde se habían recortado los carriles no relevantes.

Figura 11. Secuencias de la región adaptadora. El ARN se sometió a transcripción reversa con oligonucleótidos que comprenden una región cebadora universal y una región adaptadora en el extremo 3' terminal. El ADNc se amplificó luego utilizando la secuencia de la región del cebador universal como cebador directo y las secuencias específicas del gen como cebadores reversos. Los productos amplificados se visualizaron en un gel de agarosa. La región del adaptador consta de G (a), GGGGG y rGrGrG en los carriles 1 y 2 respectivamente (b). rG indica nucleótidos de ARN en lugar de nucleótidos de ADN.

Figura 12. Secuencias de cebadores universales. El ARN se sometió a transcripción reversa con oligonucleótidos que comprendían una secuencia de cebador universal y una región adaptadora en el extremo 3'. El ADNc se amplificó luego por PCR utilizando un cebador directo complementario a la región del cebador universal y un cebador reverso complementario a la secuencia específica del gen. Univ\_seq\_4 (a), univ\_seq\_5 (b) y univ\_seq\_f (c). Las bandas blancas verticales en las fotos de gel indican dónde se han recortado los carriles no relevantes. De lo contrario, los carriles pertenecen a la misma fotografía de gel.

Figura 13. Secuencias de cebadores específicas del gen para la primera reacción de PCR. Cebadores reversos específicos del gen utilizados en la amplificación de secuencias en la primera reacción de PCR. Tanto la primera reacción de PCR como los posteriores productos de PCR anidados se ejecutaron y se visualizaron en un gel de agarosa. Los cebadores reversos utilizados son IgKC\_v4, IgLC\_v5, IgHGC\_v13 en los carriles 1-3 respectivamente (a), K\_GSP1, L\_GSP1, G\_GSP1 en los carriles 1-3 respectivamente (b), K\_GSP1c, L\_GSP1c en los carriles 1-2 respectivamente (c), G\_GSP1 (d), L\_GSP1d, G\_GSP1g en los carriles 1-2 respectivamente (e), G\_GSP1h, G\_GSP1k, L\_GSP1f, L\_GSP1g en los carriles 1-4 respectivamente (f), G\_GSP1d (g) L\_GSP1h-o en los carriles 1-8 respectivamente (h), G\_GSP1m-q y G\_GSP1t en los carriles 1-6 respectivamente (i). K, L y G en los nombres de los cebadores indican que los cebadores son específicos para las regiones constantes de inmunoglobulina kappa, lambda y gamma, respectivamente. Cada gel comienza con un marcador de carril a la izquierda seguido de los carriles de muestra. Las barras blancas entre carriles en la misma fotografía de gel indican dónde se han recortado los carriles no relevantes intermedios.

Figura 14. Secuencias genéticas específicas para la 2ª reacción de PCR. Cebadores reversos específicos del gen utilizados en la amplificación de secuencias en la 2ª reacción de PCR. Los productos de PCR se corrieron y se visualizaron en un gel de agarosa. Los cebadores reversos utilizados son K\_GSP2, L\_GSP2, G\_GSP2 en los carriles 1-3 respectivamente (a), K\_GSP2v2a, K\_GSP2v2b, L\_GSP2v2 en los carriles 1-3 respectivamente (b), K\_GSP2v2c, K\_GSP2v2c, G\_GSP2v2c1, G\_GSP2v2c2 en los carriles 1-4 respectivamente (c), K\_GSP2v2d-f en los carriles 1-3 respectivamente (d), K\_GSP2v2g, L\_GSP2v2d y G\_GSP2b en los carriles 1-3 respectivamente (e). K, L, G en los nombres de los cebadores indican que son específicos para las regiones constantes de inmunoglobulina kappa, lambda y gamma, respectivamente. Cada gel comienza con un marcador de carril a la izquierda seguido de los carriles de muestra. Las barras blancas entre carriles en la misma fotografía de gel indican que los carriles intermedios no relevantes se han recortado.

Figura 15. Ubicaciones potenciales de las secuencias de códigos de barras para identificar un par unido de secuencias de polinucleótidos. El esquema ilustra el enlace físico de dos segmentos de ácido nucleico, A y B (por ejemplo, dos ADNc). Un código de barras (BC) se adjunta a cualquiera de los extremos, o ambos extremos, o en cualquier lugar de la secuencia que une A y B. En una realización, A y B representan secuencias de cadena pesada y liviana de inmunoglobulina.

Figura 16. Diferentes tipos de colas de superposición-extensión. La línea en negrita corresponde a una secuencia específica del gen y la línea delgada corresponde a la cola solapada. Como se indicó, el solapamiento puede ser completamente debido al solapamiento de la secuencia del cebador o también a un solapamiento parcial o total con una secuencia específica del gen. Como se indica, la superposición también puede contener una secuencia de código de barras. Las estructuras I, II y III indican ubicaciones potenciales de las superposiciones.

Figura 17. Descripción esquemática de la adición de códigos de barras externos a un par de cadenas livianas y pesadas de anticuerpos. Se muestran los productos de una reacción de transcripción reversa. El cebador de PCR específico del gen LC contiene un código de barras, un sitio de cebador de secuenciación y un sitio de restricción (RE1) para permitir que estos elementos se agreguen al extremo 3' del producto de PCR resultante. Se indican los cebadores específicos para LC y HC con extensiones solapadas y que codifican un sitio de restricción (RE3). También se indica un cebador reverso específico para HC que contiene un sitio de cebador de secuenciación y un sitio de restricción (RE2). La amplificación produce un ácido nucleico con la estructura enlazada que se muestra con un código de barras en un extremo.

Figura 18. Descripción esquemática de la adición de códigos de barras internos a un par de cadenas livianas y pesadas de anticuerpos. Se muestra un método para usar adaptadores que contienen secuencias de superposición de extensión y de código de barras para unir los ADNc resultantes de la transcripción reversa de los ARNm utilizando cebadores oligo (dT). El método mostrado aprovecha las actividades en 3' de cola y cambio de plantilla de la transcriptasa reversa para agregar secuencias de superposición y extensión a los ADNc que se unirán. En este ejemplo, uno de los adaptadores agrega un código de barras y una secuencia de superposición a uno de los ADNc que se unirán, mientras que solo la secuencia de superposición y extensión se agrega al otro ADNc que se unirá.



Después de la amplificación, se genera una estructura unida que transporta una secuencia de código de barras entre los ADNc unidos.

5      Figura 19. Descripción esquemática de la adición de dos códigos de barras internos a un par de cadenas livianas y pesadas de anticuerpos que utilizan cebadores de superposición-extensión de secuencia universal. Se muestra un método para utilizar adaptadores que contienen una secuencia universal y un código de barras para unir los ADNc resultantes de la transcripción reversa de los ARNm utilizando cebadores oligo (dT). En este ejemplo, los cebadores de PCR a la secuencia universal agregan una secuencia de superposición y extensión a cada uno de los ADNc que se unirán. Después del esquema de amplificación mostrado, se genera una estructura unida que lleva dos códigos de barras entre los ADNc unidos.

10     Figura 20. Descripción esquemática de la adición de dos códigos de barras internos a un par de cadenas livianas y pesadas de anticuerpos que utilizan adaptadores de superposición y extensión. Se muestra un método para usar adaptadores que contienen un código de barras y una secuencia de superposición para unir los ADNc resultantes de la transcripción reversa de los ARNm utilizando cebadores específicos de genes (GSP). En este ejemplo, las secuencias de extensión de solapamiento en los adaptadores agregados a cada uno de los ADNc permiten la unión de los ADNc por fusión. Después del esquema de amplificación mostrado, se genera una estructura unida que lleva dos códigos de barras entre los ADNc unidos.

20     Figura 21. Uso de GSP con código de barras durante la transcripción reversa en combinación con adaptadores agregados de cambio de plantilla. RT se realizó con ARN de PBMC total y univ\_seq\_2 oligo de cambio de plantilla e IgKC\_v3 GSP (carriles 1-2) e IgLC\_v5 GSP (carriles 3-4) con una secuencia adicional de 5' flanqueante, de la cual la primera parte es la secuencia Fixed\_PCR3, y los últimos 8 pb es un código de barras. Se usaron partes alícuotas de la reacción de RT en las siguientes reacciones de PCR, ya sea con un cebador en 5' V<sub>K</sub> (carril 1) o V<sub>L</sub> (carril 3) o Univ\_seq\_2 (carriles 2 y 4) como cebador en 5', y Fixed\_PCR3 como el 3' cebador. Los productos de PCR en los carriles 2 y 4 se ejecutaron como un borrón, lo que demuestra que los GSP con código de barras no son específicos en la reacción de RT, y no son adecuados para su uso con adaptadores agregados de cambio de plantilla.

25     Figura 22. Esquema de activación para la clasificación por citometría de flujo de células individuales en placas de 96 pozos. Los plasmoblastos se definen como CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>CD38<sup>++</sup>. Las PBMC individuales se activaron por primera vez según su perfil FSC y SSC (no se muestra). Las células CD19<sup>+</sup> B vivas se clasificaron (panel izquierdo), y se redujeron aún más a células B CD20<sup>-</sup> (2do panel desde la izquierda), y se refinaron a células CD27<sup>+</sup>CD38<sup>++</sup>. A partir de esto, se determinaron los plasmoblastos IgG<sup>+</sup> como IgA<sup>-</sup> e IgM<sup>+</sup>, ya que los plasmoblastos IgG<sup>+</sup> no expresan IgG en la superficie celular. Esta población se clasificó en células individuales en placas de 96 pozos.

35     Figura 23. Los plasmoblastos están presentes en personas sometidas a desafío inmunológico. Los plasmoblastos constituyeron 0.15% de las células B de sangre periférica en un donante sano representativo, y oscilan entre 0.5% y 16.4% en personas con una variedad de desafíos inmunológicos, incluidas infecciones (*Staphylococcus aureus* y *Clostridium difficile*), cáncer (un paciente con melanoma metastático que no fue progresivo durante más de 4 años debido al tratamiento con ipilimumab y un paciente con adenocarcinoma de pulmón metastático que no fue progresivo durante más de 3 años después de recibir quimioterapia) y vacunación (recibo de la vacuna contra el virus de la influenza). Esto muestra que los plasmoblastos se elevan y pueden obtenerse de un rango de sujetos que desarrollan respuestas inmunes de interés para el aislamiento de plasmoblastos individuales para la secuenciación de alto-rendimiento del repertorio de anticuerpos para caracterizar la respuesta humoral activa.

40     Figura 24. Los anticuerpos recombinantes expresados se secretaron durante 2-3 semanas en transfecciones transitorias. Como se describe en la Figura 2, el ADNc de inmunoglobulina de cadena liviana y pesada emparejados se clonó mediante PCR y se cotransfectó en células 293T a una escala de 48 pozos. Los sobrenadantes se recogieron en días alternos durante 18 días. Se realizó ELISA de IgG antihumana para determinar la cantidad de anticuerpos secretados en los sobrenadantes recogidos, y se graficó la concentración de los anticuerpos en los sobrenadantes de un panel de cotransfectantes individuales. La secreción tendió a alcanzar su punto máximo en el día 9 y disminuyó sustancialmente en el día 18.

50     Figura 25. Cadena pesada (HC) y cadena liviana (LC) de anticuerpos emparejados de un humano vacunado contra la influenza exhiben una variación a través de las regiones determinantes del complemento (CDR). Figura 25A: Dendrograma parcial de anticuerpos contra la gripe. Después del emparejamiento de cadenas pesadas y livianas, se generó una alineación de secuencias múltiples para las cadenas pesadas, y se generó otra alineación de secuencias múltiples para las cadenas livianas. Ambas alineaciones de secuencias múltiples se generaron utilizando Clustalw2 2.1 con parámetros predeterminados. Las dos alineaciones se concatenaron juntas y se usaron para construir un árbol en CLC Sequence Viewer v. 6.5.2 utilizando el método de unión de vecinos con 100 repeticiones de arranque. Figura 25B: CDR de cadena pesada para una familia clonal de un paciente vacunado contra la gripe. Los identificadores en la figura corresponden a los nombres de secuencia en el Listado de Secuencias de la siguiente manera: 51.A11.1 = NA.51.11.A11.1.454.heavy.3.nb-aa, 49.A08.1 = NA.49.8.A08.1.454.heavy.3.nb-aa, 51.D07.1 es la secuencia de aminoácidos obtenida al traducir NA.51.40.D07.1.454.heavy.3.nb en el cuadro 1. Figura 25C: CDR de cadena liviana para una familia clonal de gripe paciente vacunado. Los identificadores en la figura corresponden a los nombres de

secuencia en el Listado de Secuencias de la siguiente manera: 51.A11.1 = NA.51.11.A11.1.454.light.4.nb-aa, 49.A08.1 = NA.49.8.A08.1.454.1ight. 4.nb-aa, 51.D07.1 = NA.51.40.D07.1.454.light.4.zerom50-aa.

5 Figura 26. Anticuerpos anti-influenza recombinantes unidos a la vacuna contra el virus de la gripe Fluzone. El análisis del árbol evolutivo (Figura 8) del conjunto de datos del repertorio de anticuerpos de cadena pesada y liviana generados para el humano vacunado contra la influenza descrito en la Figura 7 se realizó para seleccionar los anticuerpos representativos de las familias clonales identificadas. Las cadenas pesadas y livianas para los anticuerpos seleccionados que representan familias clonales, así como varias ramas singletes, se clonaron mediante PCR y se cotransfectaron en células 293T (como se describe en la Figura 2), y los sobrenadantes se recolectaron de los transfectantes como se describe en la Figura 24. Los anticuerpos recombinantes se probaron para determinar la reactividad contra la vacuna contra el virus de la gripe Fluzone (Sanofi) mediante ELISA, con la vacuna de Fluzone recubierta en la placa ELISA. Los anticuerpos del virus de la influenza recombinante se incubaron en la placa ELISA a 100 ng/ml, y se utilizó un anticuerpo anti-IgG humano conjugado con peroxidasa de rábano picante (HRP)- para detectar la unión del anticuerpo. La reacción del sustrato TMB se dejó ir durante 30 minutos antes de inactivar con parada ácida. La lectura se muestra como una absorbancia de 450 nm, ya que no había estándares disponibles. Los anticuerpos recombinantes múltiples que representan las familias clonales identificadas se unieron a la vacuna contra el virus de la influenza, mientras que los anticuerpos recombinantes representativos de otras familias clonales y los "callejones sin salida" no se unieron a la vacuna contra la influenza.

20 Figura 27. Anticuerpos recombinantes contra la influenza representativos de familias clonales se unen a las hemaglutininas del virus de la influenza con afinidades picomolares. Los anticuerpos recombinantes contra el virus de la influenza representativos de las familias clonales de humanos vacunados con Fluzone (Figura 7) que se unieron a la vacuna contra la influenza en un ensayo ELISA (Figura 26) se probaron utilizando un instrumento de resonancia de plasmón de superficie (SPR) (ProteOn System, Bio- Rad Laboratories) para determinar sus afinidades de unión a la hemaglutinina de la influenza (tanto las cepas H3N2 A/Perth/16/2009 como H1N1 A/California/07/2009 presentes en la vacuna). Los anticuerpos del virus anti-influenza recombinante se unieron a la superficie utilizando la química de EDAC-NHS, y las hemaglutininas de California H3N2 Perth y H1N1 se analizaron independientemente como ligandos, con hemaglutinina como analito. La columna Ka denota las tasas de activación, la columna Kd las velocidades de desconexión y la constante de disociación KD. Múltiples anticuerpos recombinantes se unieron a la hemaglutinina H3N2 Perth o H1N1 California con afinidades picomolares.

30 Figura 28. Los anticuerpos anti-influenza recombinantes neutralizan la infectividad del virus de la influenza en los ensayos de microneutralización. Se enviaron seis anticuerpos que muestran reactividad en el ELISA de Fluzone (Figura 26) a la organización de investigación contratada (CRO) Virapur, LLC, (San Diego, CA) para su análisis en un ensayo de microneutralización utilizando la cepa del virus de la influenza H1N1 California/07/2009 y La cepa del virus de la influenza H3N2 A/Perth/16/2009, 2 de las tres manchas del virus de la influenza en la vacuna Fluzone. 5 de los 6 anticuerpos recombinantes neutralizaron el virus de la influenza en el ensayo de microneutralización, lo que evita la infectividad a niveles de microgramos por mililitro y posiblemente a concentraciones de microgramos por mililitro. El anticuerpo recombinante F21 neutralizó el H3N2 Perth, y aunque se unió a Fluzone en el ensayo ELISA, no mostró unión en el análisis SPR (Figura 27) probablemente debido a que las concentraciones del analito de hemaglutinina que se usaron eran demasiado bajas para ser detectables.

40 Figura 29. Anticuerpos anti-*Staph. aureus* recombinantes unidos a *S. aureus* fijo por citometría de flujo. El análisis del árbol evolutivo (Figura 8) del conjunto de datos del repertorio de anticuerpos de cadena liviana y pesada generados por el ser humano que controló (sin antibióticos) una osteomielitis crónica por *Staph. aureus* (como se describe en la Figura 7) permitió la selección de anticuerpos representativos de las familias clonales identificadas. Las cadenas pesadas y livianas para los anticuerpos seleccionados que representan familias clonales, así como varias ramas singletes, se clonaron mediante PCR y se cotransfectaron en células 293T (como se describe en la Figura 2), y los sobrenadantes se recolectaron de los transfectantes como se describe en la Figura 24. Los anticuerpos anti-*Staph. aureus* recombinantes se probaron para determinar la reactividad contra *S. aureus* fijo. El anticuerpo secundario utilizado fue una IgG antihumana de ratón conjugada con FITC, y las muestras se analizaron en un BD LSR II o LSR Fortessa. Se muestra el porcentaje de tinción positiva, con 2 anticuerpos anti-influenza utilizados como controles negativos. El anticuerpo de 2° solo no dio como resultado una unión sobre el fondo, ya que la proteína A se une débilmente a la IgG1 de ratón, que es el isotipo del anticuerpo de 2°. La tinción observada en el fondo se debe a la unión de los anticuerpos recombinantes anti-*Staph. aureus* al pequeño porcentaje de la cepa de Wood de *S. aureus* que expresa la proteína A. (a). Se muestran gráficas de citometría de flujo para los 2 anticuerpos anti-*Staph. aureus* de unión positiva, S6 y S11, junto con los anticuerpos anti-influenza de control negativo con el mismo isotipo (b). El nivel de unión de los anticuerpos anti-*Staph. aureus* a *S. aureus* fue proporcional a la cantidad de anticuerpo utilizado. 55 La línea sólida oscura representa una concentración de anticuerpo de 10 µg/ml, la línea de puntos oscura 5 µg/ml y la línea de puntos grises 1 µg/ml (c).

60 Figura 30. Los anticuerpos anti-*Staph. aureus* redujeron el número de unidades formadoras de colonias de *S. aureus*. Se incubaron anticuerpos anti-*Staph. aureus* recombinantes con *S. aureus* en combinación con suero de conejo de bebé como fuente de complemento, antes de diluirse en serie y cultivar durante la noche en placas de agar sangre al 5% de agar de soja con tripticasa (TSA). Las unidades formadoras de colonias (UFC) se contaron y se representaron

gráficamente. Dos anticuerpos recombinantes anti-*Staph. aureus* (Ab-a y Ab-b) dieron como resultado la muerte del *Staph. aureus* y, por lo tanto, un número reducido de UFC/ml.

5 Figura 31. Identificación de los objetivos del antígeno de *S. aureus* de los anticuerpos recombinantes anti-*Staph. aureus* generados por un ser humano que desarrolla una respuesta inmune eficaz contra una infección crónica por *Staph. aureus*. Se generó un lisado de proteína a partir de un aislado clínico de *Staph. aureus*. Se usaron anticuerpos anti-*Staph. aureus* recombinantes representativos de familias clonales identificadas en el repertorio de anticuerpos de un ser humano que acumula una respuesta inmune que estaba impidiendo la progresión de una infección crónica por osteomielitis por *Staph. aureus* para inmunoprecipitar proteínas de un aislado clínico de *Staph. aureus* Spa<sup>r</sup>. Los inmunoprecipitados se separaron mediante SDS-PAGE, se eliminaron las bandas identificadas y se usó la espectrometría de masas (un espectrómetro de masas con trampa de iones Agilent XCT-Plus) para identificar las proteínas inmunoprecipitadas que se presentan en la Figura.

15 Figura 32. Generación de anticuerpos contra el adenocarcinoma de pulmón de un humano con adenocarcinoma de pulmón metastásico que no fue un progresor a largo plazo. Un humano con adenocarcinoma de pulmón metastásico que se convirtió en un no progresor a largo plazo después de la quimioterapia mostró plasmoblastos sanguíneos persistentemente elevados, lo que indica una respuesta inmune activada persistentemente (Figura 7). Se clasificaron los plasmoblastos de sangre periférica del paciente, se secuenció el repertorio de anticuerpos y se clonaron y expresaron de manera recombinante los anticuerpos representativos de las familias clonales (Figura 8). Uno de los anticuerpos recombinantes expresados, que es representativo de una de las familias clonales identificadas, se une a un adenocarcinoma de pulmón independiente en tinciones inmunohistoquímicas. Luego se utilizaron matrices de tejidos para caracterizar mejor la reactividad de este anticuerpo. Las matrices de tejidos que contienen múltiples adenocarcinomas de pulmón independientes, carcinoma de células escamosas y tejido de pulmón sano se bloquearon durante la noche con 100µg/ml de anticuerpo anti-humano de cabra F (ab). Los portaobjetos se tiñeron con anticuerpos anti-adenocarcinoma de pulmón o un anticuerpo anti-influenza como control negativo, y se visualizaron con Vector Red. Los portaobjetos se tiñeron con hematoxilina. El color azul de hematoxilina se eliminó con Photoshop, de modo que solo los núcleos y la tinción Vector Red (rojo) aparecen en gris más oscuro en la imagen. Este anticuerpo recombinante se unió a 4 de 5 muestras de tejido de adenocarcinoma de pulmón independientes analizadas (contenidas en matrices de tejido), pero no se unió al carcinoma de células escamosas de pulmón ni al tejido pulmonar sano.

30 Figura 33. El anticuerpo adenocarcinoma anti-pulmón identificado (Figura 32) se une a la superficie de una línea celular de adenocarcinoma de pulmón. El anticuerpo de adenocarcinoma anti-pulmón recombinante (Figura 32) tiñó fuertemente la superficie de la línea celular H1650 de adenocarcinoma de pulmón, y exhibió solo una tinción de bajo nivel de un epitelio de riñón y una línea de células de tumor escamoso de pulmón. Este anticuerpo anti-adenocarcinoma de pulmón no se unió a una segunda línea celular de adenocarcinoma de pulmón (H2009), lo que concuerda con nuestra observación de que este anticuerpo se unió a 4 de las 5 muestras de tejido de adenocarcinoma de pulmón independientes analizadas por inmunohistoquímica (Figura 32).

40 Figura 34. Generación de anticuerpos contra el factor reumatoide de pacientes con artritis reumatoide (AR). Los anticuerpos recombinantes representativos de las familias clonales identificadas en los árboles evolutivos de los repertorios de anticuerpos generados a partir de humanos con AR (Figura 8) se seleccionaron, clonaron y expresaron de forma recombinante. Se usaron anticuerpos recombinantes derivados de pacientes con AR como el anticuerpo primario en un ELISA directo y se usó anti-IgG-HRP humana como el anticuerpo secundario, y se visualizó la unión con el sustrato TMB. Los anticuerpos recombinantes RA2 y RA3 mostraron reactividad y, por lo tanto, representan anticuerpos contra el factor reumatoide.

45 Figura 35. Generación de anticuerpos anti-CCP y anti-histona 2A de pacientes con AR. Se caracterizaron anticuerpos recombinantes adicionales generados a partir de pacientes con AR con enfermedad activa utilizando un ELISA de histona 2A y un ELISA de péptido cíclico citrulinado (CCP) (utilizando el kit ELISA de CCP2 [Axis Shield]). Los anticuerpos recombinantes se utilizaron a 125µg/ml. El panel (a) presenta los resultados de un ELISA de histona 2A y múltiples anticuerpos recombinantes unidos a la histona 2A. El panel (b) presenta los resultados del ELISA de CCP2 y varios anticuerpos recombinantes mostraron una reactividad positiva. El ELISA anti-CCP2 incluyó un control seronegativo y 2 controles seropositivos. Para ambos ensayos, la absorbancia se registró como la lectura. Los valores de absorbancia por encima del fondo (línea de puntos) se consideraron positivos.

55 Figura 36. Experimento confirmatorio independiente que demuestra la generación de anticuerpos anti-histona 2A de pacientes con AR activa. Los anticuerpos recombinantes derivados de árboles evolutivos de pacientes con AR (Figura 8 y Figura 35) se probaron en un ensayo ELISA de histona 2A. Los anticuerpos se utilizaron a 30µg/ml, una concentración 4 veces menor que la utilizada en la Figura 35. La absorbancia se registró como la lectura. Los valores de absorbancia por encima del fondo se consideraron positivos.

Figura 37. Identificación de anticuerpos anti-histona y proteína anticitrulina utilizando matrices de antígenos RA. Los anticuerpos derivados de pacientes con AR se utilizaron para sondear una matriz de antígenos de AR que contiene un espectro de proteínas y péptidos nativos y citrulinados. Después de la incubación con un anticuerpo secundario anti-IgG humana marcado con Cy-3, la unión del anticuerpo recombinante se cuantificó mediante escaneo con un

escáner de micromatriz GenePix de Axon Instruments. Las reactividades se muestran como un mapa de calor. Los anticuerpos recombinantes derivados de la AR se unen a varios antígenos nativos o citrulinados.

5 Figura 38. La secuenciación por Pacific Biosciences proporciona lecturas de secuenciación de longitud completa del amplicón de cadena pesada de IgG. Los amplicones de la cadena pesada de IgG de la placa 44 se proporcionaron a Pacific Biosciences para la secuenciación SMRT. Se muestra el número de lecturas de secuencia de consenso circular (CCS) con códigos de barras correspondientes a los pozos seleccionados.

10 Figura 39. Uso de marcadores de superficie celular alternativos y otras características celulares para identificar plasmoblastos sanguíneos. Los plasmoblastos se pueden identificar y clasificar mediante el uso de una variedad de marcadores de la superficie celular y/o características celulares. El panel (a) demuestra que los plasmoblastos exhiben una mayor dispersión de avance (FSC) que las células B en reposo. Los plasmoblastos se identificaron según la tinción de CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>CD38<sup>hi</sup>, y estos resultados demuestran que las células B (grises) son más pequeñas que los plasmoblastos (negro). El panel (b) demuestra que el uso de tinción anti-CD 19 combinada con FSC identifica una población de células B que contiene 72% de plasmoblastos. El panel (c) demuestra que, para una población de células B CD19<sup>+</sup> (las células se clasificaron como CD19 positivas), la dispersión lateral (SSC) y FSC se pueden usar para identificar una población de células B que contiene un 37% de plasmoblastos. Los paneles (d-f) presentan varios enfoques para identificar plasmoblastos dentro de la población de células B CD19<sup>+</sup>. La activación de las células FSC<sup>hi</sup> proporciona una pureza del 37% de los plasmoblastos (c). La activación de las células FSC<sup>hi</sup>CD20<sup>-</sup> dio una pureza del 71% en los plasmoblastos (d). La activación de las células FSC<sup>hi</sup>CD38<sup>+</sup> dio una pureza del 80% en plasmoblastos (e). La activación de las células FSC<sup>hi</sup>CD27<sup>+</sup> dio una pureza del 44% en los plasmoblastos (f).

20 Figura 40. Las células B blásticas humanas (plasmoblastos) son más grandes que las células B en reposo, pero en promedio son más pequeñas que los monocitos. Monocitos singletes, células B y plasmoblastos fueron bloqueados y comparados para los parámetros de dispersión lateral y delantera. Los monocitos se definieron por su perfil característico de FSC y SSC, y como CD19<sup>-</sup>CD3<sup>-</sup>. Células mostradas en los ejes FSC-A (área de dispersión de avance) y SSC-A (área de dispersión lateral) (a). Las células se muestran en los ejes (b) de FSC-W (ancho de dispersión hacia delante) y SSC-W (ancho de dispersión lateral). La mediana de FSC-A, SSC-A, FSC-W, SSC-W de plasmoblastos se dividió por la de células B en reposo o monocitos para obtener una proporción que representa la relación de tamaño entre los tipos de células (c). La mediana de FSC-A, SSC-A, FSC-W, SSC-W del vigésimo percentil de plasmoblastos se dividió por la de la mediana de células B o monocitos en reposo para obtener una relación que represente la relación de tamaño entre la célula tipos en los que al menos el 80% de los plasmoblastos son más grandes que la relación (d). Las barras de error indican un intervalo de confianza del 95%.

35 Figura 41. Tamaño de los plasmoblastos humanos en comparación con las células B en reposo por microscopía. Plasmoblastos y células B en reposo fueron clasificados como CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>CD38<sup>++</sup> y CD19<sup>+</sup>CD20<sup>+</sup> respectivamente. Luego se generaron imágenes de las células utilizando un microscopio Olympus a 200x. El área celular se midió usando ImageJ y el diámetro se determinó usando el  $\text{área} = \pi \times r^2$  donde el diámetro = 2 x radio, y el volumen se determinó por  $\frac{4}{3} \times \pi r^3$ . Las barras de error indican el rango intercuartil. Un límite de  $\geq 8 \mu\text{M}$  o  $\geq 50 \mu\text{M}^2$  o  $\geq 268 \mu\text{M}^3$  incluirá el 96% de los plasmoblastos y excluirá el 92% de las células B en reposo.

40 Figura 42. El superíndice III tiene actividad de cambio de plantilla a temperaturas en y por debajo de 50°C. La transcripción reversa (RT) se realizó durante 90 minutos utilizando las temperaturas indicadas sobre los carriles utilizando un adaptador que termina con rGrGrG, y se realizó una ronda de PCR utilizando GAPDH como cebador 3' (secuencia ATGGTTCACCCCATGACG). Como se puede ver, no se pudo ver ninguna actividad de cambio de plantilla para agregar al adaptador a 55°C, y la actividad de cambio de plantilla aumenta desde el mínimo a 50°C hasta el máximo a 42°C, la temperatura más baja probada, según lo indica el brillo de la banda en ~450 pb. El marcador es un marcador de 100 pb. Superscript III es una transcriptasa reversa MMLV que tiene mutaciones específicas que dan como resultado una pérdida de la actividad de la ARNasa H, y también tiene mutaciones en el dominio de la polimerasa para aumentar la estabilidad térmica y tiene una vida-media de 220 minutos a una temperatura ambiente de 50°C. Se espera que otras enzimas MMLV H<sup>-</sup> que han sido diseñadas para una mayor estabilidad térmica exhiban una actividad similar.

50 Figura 43. Cebadores adicionales para las regiones constantes kappa, lambda y gamma humanas. Estos cebadores se utilizaron para la 1.<sup>a</sup> PCR, y luego se realizó la 2.<sup>a</sup> PCR utilizando los cebadores de la Tabla 1 y los productos de la PCR separados en un gel de agarosa al 2% y se tomó la imagen. Los cebadores utilizados para la 1.<sup>a</sup> PCR son Kappa GSP1, kappa GSP1e, kappa GSP1f, lambda GSP1, lambda GSP1x y lambda GSP1 y respectivamente. Las secuencias se encuentran en la Tabla 10. Las barras blancas entre carriles en la misma fotografía de gel indican que los carriles intermedios no relevantes se han recortado.

55 Figura 44. Cebadores adicionales para otras regiones constantes humanas y genes. La 1.<sup>a</sup> y la 2.<sup>a</sup> PCR se realizaron y los productos se corrieron en un gel de agarosa al 2% y se generaron imágenes. Los carriles son de izquierda: marcador, mu, regiones constantes alfa, TCR alfa (a) y marcador, TCR beta (b). Los cebadores utilizados y las secuencias están en la Tabla 10. Las barras blancas entre carriles en la misma fotografía de gel indican que los carriles intermedios no relevantes se han recortado.

Figura 45. Cebadores adicionales para genes de ratón. Se realizaron 1ª y 2ª PCR y los productos se corrieron en un gel de agarosa al 2% y se generaron imágenes. Los carriles son desde la izquierda: cadenas livianas de marcador, kappa, lambda, lambda, lambda, lambda y cadena pesada mu. Los 4 carriles de lambda tuvieron esta combinación de cebadores: ratón\_lambda\_GSP1a con ratón\_lambda\_GSP2a, ratón\_lambda\_GSP1a con ratón\_lambda\_GSP2b en la mesa de luz de la mesa por la que se puede encontrar un búfalo. Marcador y cadena pesada alfa (b). Cadenas pesadas de Gamma1, 2a, 2c con 2da PCR que utilizan mo\_g12\_GSP2d y mo\_g12\_GSP2e respectivamente, marcador (c). Marcador, cadena pesada gamma 3 con 2.º PCR usando mo\_g3\_GSP2d, mo\_g3\_GSP2e respectivamente seguido de cadena pesada gamma 2b con 2.º PCR usando mo\_g2b\_GSP2d, mo\_g2b\_GSP2e respectivamente (d). Marcador, TCR alfa (e). Marcador, TCR beta (f). Las barras blancas entre carriles en la misma fotografía de gel indican que los carriles intermedios no relevantes se han recortado.

Figura 46. Matanza mediada por anticuerpos anti-*S. aureus* de *S. aureus* por la línea celular de neutrófilos HL-60. Varios anticuerpos anti-*S. aureus* recombinantes (estafilococos 1, estafilococos 4, estafilococos 6, estafilococos 7, estafilococos 9, estafilococos 12) se incubaron a 4°C con *S. aureus* durante 30 minutos, después de lo cual se eliminó el anticuerpo no unido, y el *S. aureus* se incubó con células HL-60 activadas y complemento de conejo bebé durante 45 minutos a 37°C. Luego, las células se lavaron dos veces y las bacterias extracelulares se colocaron en serie en placas de agar-sangre al 5% TSA, se incubaron durante la noche y se contaron las unidades formadoras de colonias (UFC). Los anticuerpos recombinantes estafilococos 6, estafilococos 9 y estafilococos 12 indujeron más del 20% de muertes de *S. aureus*.

Descripción detallada

20 Composiciones

*Polinucleótidos*

En algunos aspectos, una composición puede incluir un polinucleótido. El término "polinucleótido(s)" se refiere a ácidos nucleicos tales como moléculas de ADN y moléculas de ARN y análogos de los mismos (por ejemplo, ADN o ARN generado usando análogos de nucleótidos o usando química del ácido nucleico). Como se desee, los polinucleótidos pueden prepararse sintéticamente, por ejemplo, utilizando una química de los ácidos nucleicos reconocida en la técnica o utilizando enzimáticamente, por ejemplo, una polimerasa, y, si se desea, pueden modificarse. Las modificaciones típicas incluyen metilación, biotilación y otras modificaciones conocidas en la técnica. Además, un polinucleótido puede ser de cadena sencilla o de doble cadena y, cuando se desee, unido a un resto detectable. En algunos aspectos, un polinucleótido puede incluir moléculas híbridas, por ejemplo, que comprenden ADN y ARN.

"G", "C", "A", "T" y "U" generalmente representan un nucleótido que contiene guanina, citosina, adenina, timidina y uracilo como base, respectivamente. Sin embargo, se entenderá que el término "ribonucleótido" o "nucleótido" también puede referirse a un nucleótido modificado o un resto sustituto. El experto en la materia sabe bien que la guanina, citosina, adenina y uracilo pueden reemplazarse por otros restos sin alterar sustancialmente las propiedades de apareamiento de bases de un oligonucleótido que comprende un nucleótido que lleva dicho resto de reemplazo. Por ejemplo, sin limitación, un nucleótido que comprende inosina como su base puede emparejarse con nucleótidos que contienen adenina, citosina o uracilo. Por lo tanto, los nucleótidos que contienen uracilo, guanina o adenina pueden reemplazarse en secuencias de nucleótidos por un nucleótido que contiene, por ejemplo, inosina. En otro ejemplo, la adenina y la citosina en cualquier parte del oligonucleótido pueden reemplazarse con guanina y uracilo, respectivamente, para formar un emparejamiento de bases W-Tob G-U con el ARNm diana. Las secuencias que contienen tales restos de reemplazo son adecuadas para las composiciones y métodos descritos en el presente documento.

Como se usa en este documento, y a menos que se indique lo contrario, el término "complementario", cuando se usa para describir una primera secuencia de nucleótidos en relación con una segunda secuencia de nucleótidos, se refiere a la capacidad de un polinucleótido que comprende la primera secuencia de nucleótidos para hibridar y formar una estructura dúplex bajo ciertas condiciones con un polinucleótido que comprende la segunda secuencia de nucleótidos, como entenderá el experto en la materia. Tales condiciones pueden ser, por ejemplo, condiciones rigurosas, donde las condiciones rigurosas pueden incluir: NaCl 400 mM, PIPES 40 mM pH 6.4, EDTA 1 mM, 50°C o 70°C durante 12-16 horas seguidas de lavado. Se pueden aplicar otras condiciones, como las condiciones fisiológicamente relevantes que pueden encontrarse dentro de un organismo. El experto en la materia podrá determinar el conjunto de condiciones más adecuadas para una prueba de complementariedad de dos secuencias de acuerdo con la aplicación final de los nucleótidos hibridados.

Las secuencias complementarias incluyen el apareamiento de bases de una región de un polinucleótido que comprende una primera secuencia de nucleótidos a una región de un polinucleótido que comprende una segunda secuencia de nucleótidos en la longitud o una porción de la longitud de una o ambas secuencias de nucleótidos. Dichas secuencias pueden denominarse "complementarias" entre sí en este documento. Sin embargo, cuando una primera secuencia se denomina "sustancialmente complementaria" con respecto a una segunda secuencia en el presente documento, las dos secuencias pueden ser complementarias, o pueden incluir una o más, pero en general no más de aproximadamente 5, 4, 3, o 2 pares de bases no coincidentes dentro de las regiones que están emparejadas en base.

Para dos secuencias con pares de bases no coincidentes, las secuencias se considerarán "sustancialmente complementarias" siempre que las dos secuencias de nucleótidos se unan entre sí mediante el emparejamiento de bases.

5 Las secuencias "complementarias", como se usan en este documento, también pueden incluir, o formarse completamente a partir de pares de bases que no sean de Watson-Crick y/o pares de bases formados a partir de nucleótidos no naturales y modificados, en la medida en que las realizaciones anteriores con respecto a su capacidad para hibridar se cumplen. Dichos pares de bases que no son de Watson-Crick incluyen, pero no están limitados a, el par de bases G:U Wobble o Hoogsteen.

10 El término porcentaje de "identidad", en el contexto de dos o más secuencias de ácido nucleico o polipéptido, se refiere a dos o más secuencias o subsecuencias que tienen un porcentaje específico de nucleótidos o residuos de aminoácidos que son iguales, cuando se comparan y se alinean para el máximo correspondencia, medida según uno de los algoritmos de comparación de secuencias que se describen a continuación (por ejemplo, BLASTP y BLASTN u otros algoritmos disponibles para personas con habilidades) o mediante inspección visual. Dependiendo de la aplicación, el porcentaje de "identidad" puede existir sobre una región de la secuencia que se está comparando, por ejemplo, sobre un dominio funcional, o, alternativamente, existe sobre la longitud completa de las dos secuencias a comparar.

20 Para la comparación de secuencias, normalmente una secuencia actúa como una secuencia de referencia con la que se comparan las secuencias de prueba. Cuando se usa un algoritmo de comparación de secuencias, las secuencias de prueba y referencia se ingresan en un ordenador, las coordenadas de la subsecuencia se designan, si es necesario, y se designan los parámetros del programa de algoritmos de secuencia. El algoritmo de comparación de secuencia calcula luego el porcentaje de identidad de secuencia para la secuencia de prueba en relación con la secuencia de referencia, en función de los parámetros de programa designados.

25 La alineación óptima de las secuencias para la comparación se puede realizar, por ejemplo, mediante el algoritmo de homología local de Smith & Waterman, Adv. Appl. Math. 2:482 (1981), por el algoritmo de alineamiento por homología de Needleman & Wunsch, J. Mol. Biol. 48:443 (1970), por la búsqueda del método de similitud de Pearson & Lipman, Proc. Nat'l. Acad. Sci. USA 85:2444 (1988), por implementación por ordenador de estos algoritmos (GAP, BESTFIT, FASTA, and TFASTA in the Wisconsin Genetics Software Package, Genetics Computer Group, 575 Science Dr., Madison, Wis.), o por inspección visual (véase en general Ausubel et al., infra).

30 Un ejemplo de un algoritmo que es adecuado para determinar el porcentaje de identidad de secuencia y la similitud de secuencia es el algoritmo BLAST, que se describe en Altschul et al., J. Mol. Biol. 215:403-410 (1990). El software para realizar los análisis de BLAST está disponible públicamente a través de la página del National Center for Biotechnology Information.

35 Las secuencias idénticas incluyen el 100% de identidad de un polinucleótido que comprende una primera secuencia de nucleótidos a un polinucleótido que comprende una segunda secuencia de nucleótidos en toda la longitud de una o ambas secuencias de nucleótidos. Dichas secuencias pueden denominarse "totalmente idénticas" con respecto a las demás en el presente documento. Sin embargo, en algunos aspectos, cuando una primera secuencia se denomina "sustancialmente idéntica" con respecto a una segunda secuencia en el presente documento, las dos secuencias pueden ser completamente complementarias, o pueden tener una o más, pero en general no más de aproximadamente 5, 4, 3 o 2 nucleótidos no coincidentes en la alineación. En algunos aspectos, cuando se hace referencia a una primera secuencia como "sustancialmente idéntica" con respecto a una segunda secuencia de este documento, las dos secuencias pueden ser completamente complementarias, o pueden ser aproximadamente 50, 60, 70, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98 o 99% idénticos entre sí.

45 Cuando una primera secuencia se denomina "distinta" con respecto a la identidad de una segunda secuencia en el presente documento, las dos secuencias tienen al menos uno o más nucleótidos no coincidentes en el alineamiento. En algunos aspectos, distintas secuencias pueden tener 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o más nucleótidos no coincidentes en la alineación. En algunos aspectos, las secuencias distintas pueden ser aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, o menos del 100% idénticos entre sí. En algunos aspectos, cuando una primera secuencia se denomina "distinta" con respecto a una segunda secuencia en el presente documento, las dos secuencias pueden tener secuencias sustancial o totalmente idénticas, pero en cambio difieren unas de otras en función de los diferentes patrones de modificación dentro de las secuencias. Tales modificaciones son generalmente conocidas en la técnica, por ejemplo, la metilación.

55 En algunos aspectos, un polinucleótido puede estar presente en una biblioteca de polinucleótidos. En algunos aspectos, una biblioteca de polinucleótidos puede incluir una pluralidad de polinucleótidos. En algunos aspectos, cada polinucleótido en la pluralidad de polinucleótidos puede derivarse de una sola muestra. En algunos aspectos, una sola muestra puede incluir una sola célula como una célula B.

- La notación convencional se usa en el presente documento para describir secuencias de nucleótidos: el extremo izquierdo de una secuencia de nucleótidos de cadena sencilla es el extremo 5'; La dirección de la izquierda de una secuencia de nucleótidos de doble cadena se conoce como la dirección 5'. La dirección de la adición de 5' a 3' de nucleótidos a los transcritos de ARN nacientes se denomina dirección de transcripción. La cadena de ADN que tiene la misma secuencia que un ARNm se conoce como la "cadena codificante"; las secuencias en la cadena de ADN que tienen la misma secuencia que un ARNm transcrito a partir de ese ADN y que están ubicadas en el extremo 5' al extremo 5' del transcrito de ARN se denominan "secuencias corriente arriba"; las secuencias en la cadena de ADN que tienen la misma secuencia que el ARN y que son 3' al extremo 3' de la transcripción del ARN codificante se denominan "secuencias corriente abajo".
- 5 El término "ARN mensajero" o "ARNm" se refiere a un ARN que no tiene intrones y que se puede traducir en un polipéptido.
- 10 El término "ADNc" se refiere a un ADN que es complementario o idéntico a un ARNm, ya sea de forma de cadena sencilla o de cadena doble.
- 15 El término "amplicón" se refiere al producto amplificado de una reacción de amplificación de ácido nucleico, por ejemplo, RT-PCR.
- 20 El término "hibridar" se refiere a una secuencia de interacción no covalente con un ácido nucleico complementario. La hibridación puede ocurrir en todo o en una parte de una secuencia de ácido nucleico. Los expertos en la materia reconocerán que la estabilidad de un dúplex de ácido nucleico, o híbridos, puede determinarse por la Tm. Puede encontrar orientación adicional sobre las condiciones de hibridación en: Current Protocols in Molecular Biology, John Wiley & Sons, N.Y., 1989, 6.3.1-6.3.6 y en: Sambrook et al., Molecular Cloning, a Laboratory Manual, Cold Spring Harbor Laboratory Press, 1989, Vol. 3.
- 25 Como se usa en este documento, "región" se refiere a una porción contigua de la secuencia de nucleótidos de un polinucleótido. Los ejemplos de regiones se describen en el presente documento e incluyen regiones de identificación, regiones de identificación de muestras, regiones de identificación de placas, regiones adaptadoras y similares. En algunos aspectos, un polinucleótido puede incluir una o más regiones. En algunos aspectos, un polinucleótido puede incluir menos de 2, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o más regiones. En algunos aspectos, las regiones se pueden acoplar. En algunos aspectos, las regiones pueden estar acopladas operativamente. En algunos aspectos, las regiones pueden estar físicamente acopladas.
- 30 Como se usa en este documento, "región variable" se refiere a una secuencia de nucleótidos variable que surge de un evento de recombinación, por ejemplo, puede incluir una región V, J y/o D de una secuencia del receptor de inmunoglobulina o célula T aislada de una célula T o célula B de interés, como una célula T activada o una célula B activada.
- 35 Como se usa en este documento, "región de inmunoglobulina variable de células B" se refiere a una secuencia de nucleótidos de inmunoglobulina variable aislada de una célula B. Por ejemplo, una secuencia de inmunoglobulina variable puede incluir una región V, J y/o D de una secuencia de inmunoglobulina aislada de una célula B de interés, como una célula B de memoria, una célula B activada o un plasmoblasto.
- 40 Como se usa en este documento, "región de identificación" se refiere a un marcador de secuencia de nucleótidos (por ejemplo, una secuencia de código de barras única) que se puede acoplar a al menos una secuencia de nucleótidos para, por ejemplo, una identificación posterior de la al menos una secuencia de nucleótidos.
- 45 Como se usa en este documento, "región de inmunoglobulina" se refiere a una porción contigua de la secuencia de nucleótidos de una o ambas cadenas (pesada y liviana) de un anticuerpo.
- 50 Como se usa en este documento, "región adaptadora" se refiere a un enlazador que acopla una primera secuencia de nucleótidos a una segunda secuencia de nucleótidos. En algunos aspectos, una región adaptadora puede incluir una parte contigua de la secuencia de nucleótidos que actúa como un enlazador. Por ejemplo, una región adaptadora puede tener la secuencia GGG y acoplar una primera secuencia a una segunda secuencia a través de la unión entre GGG y CCC.
- 55 En algunos aspectos, un polinucleótido puede incluir una región de ADNc. En algunos aspectos, un polinucleótido puede incluir una región de adaptador de identificación de muestra. En algunos aspectos, un polinucleótido puede incluir una región adaptadora. En algunos aspectos, un polinucleótido puede incluir una región cebadora universal. En algunos aspectos, un polinucleótido puede incluir una región de identificación de placa. En algunos aspectos, un polinucleótido puede incluir una primera región de identificación de placa. En algunos aspectos, un polinucleótido puede incluir una segunda región de identificación de placa. En algunos aspectos, un polinucleótido puede incluir una región de sitio de restricción. En algunos aspectos, un

polinucleótido puede incluir una primera región de sitio de restricción. En algunos aspectos, un polinucleótido puede incluir una segunda región de sitio de restricción. En algunos aspectos, un polinucleótido puede incluir una región de secuenciación. En algunos aspectos, un polinucleótido puede incluir una primera región de secuenciación. En algunos aspectos, un polinucleótido puede incluir una segunda región de secuenciación.

- 5 En algunos aspectos, un polinucleótido puede incluir una pluralidad de cualquier región descrita en este documento. Por ejemplo, un polinucleótido puede incluir una primera región de identificación de la muestra y una segunda región de identificación de la muestra. En algunos aspectos, la primera región de identificación de la muestra y la segunda región de identificación de la muestra son idénticas o sustancialmente idénticas. En algunos aspectos, la primera región de identificación de la muestra y la segunda región de identificación de la muestra son distintas. En algunos aspectos, una región de identificación está acoplada a una región de inmunoglobulina variable.

10 En algunos aspectos, la secuencia de una región será al menos lo suficientemente larga para servir como una secuencia diana para un cebador o una sonda en una reacción de PCR. En algunos aspectos, una región puede tener una longitud de 1 a más de 5000 pares de bases. Por ejemplo, una región puede tener una longitud de 1 a 10 000 nucleótidos, por ejemplo, de 2 a 30 nucleótidos, incluyendo todos los sub-rangos entre ellos. Como ejemplos no limitativos, una región puede ser de 1-30 nucleótidos, 1-26 nucleótidos, 1-23 nucleótidos, 1-22 nucleótidos, 1-21 nucleótidos, 1-20 nucleótidos, 1-19 nucleótidos, 1-18 nucleótidos, 1-17 nucleótidos, 18-30 nucleótidos, 18-26 nucleótidos, 18-23 nucleótidos, 18-22 nucleótidos, 18-21 nucleótidos, 18-20 nucleótidos, 19-30 nucleótidos, 19-26 nucleótidos, 19-23 nucleótidos, 19-22 nucleótidos, 19-21 nucleótidos, 19-20 nucleótidos, 20-30 nucleótidos, 20-26 nucleótidos, 20-25 nucleótidos, 20-24 nucleótidos, 20-23 nucleótidos, 20-22 nucleótidos, 20-21 nucleótidos, 21-30 nucleótidos, 21-26 nucleótidos, 21-25 nucleótidos, 21-24 nucleótidos, 21-23 nucleótidos, o 21-22 nucleótidos. En algunos aspectos, una región puede ser aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, o más nucleótidos en longitud. En algunos aspectos, una región puede ser menor que 50, 50-100, 100-200, 200-300, 300-400, 400-500, 500-600, 600-700, 700-800, 800-900, 900-1000, o más de 1000 nucleótidos de longitud. En algunos aspectos, una región puede ser menor que 1000, 1000-2000, 2000-3000, 3000-4000, 4000-5000, 5000-6000, 6000-7000, 7000-8000, 8000-9000, 9000-10000 o superior de 10000 nucleótidos de longitud. En algunos aspectos, una región puede incluir al menos dos nucleótidos, al menos 3, al menos 4, al menos 5, al menos 6, al menos 7, al menos 8, al menos 9, al menos 10, al menos 15, en al menos 20 o más nucleótidos de un polinucleótido descrito en el presente documento.

30 El término "muestra" puede incluir ARN, ADN, una célula única o múltiples células o fragmentos de células o una alícuota de líquido corporal, tomada de un sujeto (por ejemplo, un sujeto mamífero, un sujeto animal, un sujeto humano, o un sujeto animal no humano). Las muestras pueden seleccionarse por un experto en la técnica usando cualquier medio conocido o descubierto más tarde, incluyendo centrifugación, punción venosa, extracción de sangre, excreción, frotis, eyaculación, masaje, biopsia, aspirado con aguja, muestra de lavado, raspado, incisión quirúrgica, microdissección por captura con láser, gradiente de separación, o intervención u otros medios conocidos en la técnica. Las muestras también pueden seleccionarse por un experto en la técnica usando uno o más marcadores que se sabe que están asociados con una muestra de interés. Las muestras también pueden seleccionarse utilizando métodos conocidos en la técnica, tales como la selección de células y FACS. Otros ejemplos de métodos de selección de muestras se describen en la sección de Ejemplos a continuación.

40 En algunos aspectos, un polinucleótido puede derivarse de una muestra única o asociarse con ella. En algunos aspectos, una región puede derivarse o asociarse con una sola muestra. En algunos aspectos, una región de ADNc puede derivarse de una muestra única o asociarse con ella. En algunos aspectos, una región de amplicón puede derivarse o asociarse con una sola muestra. Una "muestra única" incluye una muestra que comprende polinucleótidos que se toma de una sola fuente. En algunos aspectos, una única fuente incluye una muestra tomada en un punto de tiempo particular o en una ubicación particular, por ejemplo, en un sujeto o matraz de células o placa de células. En algunos aspectos, una primera muestra única se toma de un primer sujeto en un primer punto de tiempo y una segunda muestra única se toma de un primer sujeto en un segundo punto de tiempo que es distinto del primer punto de tiempo. En algunos aspectos, una primera muestra única se toma de un primer sujeto en una primera ubicación y una segunda muestra se toma de un primer sujeto en una segunda ubicación que es distinta de la primera ubicación. En algunos aspectos, se toma una primera muestra única de un primer sujeto en un punto de tiempo y una segunda muestra única se toma de un segundo sujeto en un punto de tiempo. En algunos aspectos, una primera muestra única se toma de un primer sujeto en una ubicación y una segunda muestra se toma de un segundo sujeto en una ubicación. En una realización, una muestra comprende polinucleótidos que incluyen ARNm derivado de una o más células B. En otra realización, una muestra comprende polinucleótidos que incluyen ADNc derivado de una o más células B. En otra realización, una muestra única comprende ARNm derivado de una o más células B clasificadas en un solo pozo de una placa de 96 pozos o 384 pozos. Las muestras generalmente se derivan de una(s) célula(s) procariótica (por ejemplo, una(s) célula(s) bacteriana(s)), una(s) célula(s) eucariota (por ejemplo, una(s) célula(s) de mamífero y levadura) u otras fuentes de material genético como un virus o fagos. El término "mamífero" o "de mamífero", como se usa en el presente documento, incluye tanto a humanos como a no humanos e incluye, entre otros, humanos, primates no humanos, caninos, felinos, murinos, bovinos, equinos y porcinos. En algunos aspectos, los métodos de la invención se aplican a muestras individuales en una placa con al menos 96 pozos, al menos 384 pozos, al menos 1536 pozos o más pozos. En aspectos adicionales, los métodos de la invención se aplican a muestras individuales en



al menos una, dos, tres, cuatro, cinco, seis, siete, ocho, diez, quince, veinte, treinta o más placas con al menos 96 pozos cada una.

- 5 En algunos aspectos, una secuencia de la región del adaptador 5' y/o una región de identificación de la muestra se agregan a todos los ADNc de una sola muestra, por ejemplo, durante la RT y no solo a los genes Ig. En algunos aspectos, los cebadores específicos del gen 3'(GSP) se pueden usar para amplificar cualquier gen expresado en la muestra única. En algunos aspectos, los genes se amplifican y tienen una región variable 5', por ejemplo, receptores de células T y receptores de células B sin necesidad de múltiples cebadores en 5' degenerados para amplificar los genes de interés. Las GSP pueden incluir cebadores específicos para las cadenas de IgG, IgM, IgD, IgA, IgE, TCR y otros genes de interés.
- 10 En algunos aspectos, también se pueden realizar múltiples rondas de PCR, por ejemplo, utilizando GSP anidados. Para tales GSP anidados, el GSP para la segunda ronda de PCR se hibrida a su secuencia del gen objetivo en una posición 5' a lo largo de esa secuencia en relación con la posición hibridada por el GSP utilizado en la primera ronda de PCR.
- 15 En algunos aspectos, la región de ADNc o una región de amplicón puede incluir un polinucleótido de ADN. En algunos aspectos, la región de ADNc o una región de amplicón puede incluir un polinucleótido de ADNc. En algunos aspectos, la región de ADNc o una región de amplicón puede incluir un polinucleótido de ARN hibridado a un polinucleótido de ADN. En algunos aspectos, la región de ADNc o una región de amplicón puede incluir un polinucleótido de ARNm hibridado a un polinucleótido de ADNc.
- 20 En algunos aspectos, una región de cebador universal no es totalmente complementaria a ningún exón humano. En algunos aspectos, una región de cebador universal no es totalmente complementaria a ningún gen humano expresado. En algunos aspectos, una región de cebador universal tiene una estructura secundaria mínima.
- 25 En algunos aspectos, una región de amplicón comprende una secuencia de amplicón de cadena pesada de inmunoglobulina. En algunos aspectos, una región de amplicón comprende una secuencia de amplicón de cadena liviana de inmunoglobulina. En algunos aspectos, una región de amplicón comprende una secuencia de amplicón alfa del receptor de células T. En algunos aspectos, una región de amplicón comprende una secuencia de amplicón beta del receptor de células T.
- En algunos aspectos, un polinucleótido está presente en una biblioteca de polinucleótidos y puede diferenciarse de otros polinucleótidos presentes en la biblioteca basándose en una región del polinucleótido.
- 30 En algunos aspectos, la secuencia de la región de identificación de la muestra de cada polinucleótido en una biblioteca derivada de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca derivada de una o más muestras distintas de la primera muestra única. En algunos aspectos, la secuencia de la región de identificación de la muestra de cada polinucleótido en una biblioteca derivada de una primera muestra única difiere en al menos 1 nucleótido de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca derivada de una o más muestras distintas de la primera muestra única.
- 35 En algunos aspectos, la secuencia de la región de identificación de la muestra de cada polinucleótido en una biblioteca derivada de una primera muestra única difiere en al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o más nucleótidos de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca derivada de una o más muestras distintas de la primera muestra única.
- 40 En algunos aspectos, la secuencia de la región de identificación de la muestra de cada polinucleótido en una biblioteca derivada de una primera muestra única puede ser aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, o menos del 100% idéntico a la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca derivada de una o más muestras distintas de las Primera muestra única.
- 45 En algunos aspectos, la secuencia de la región de identificación de la muestra de cada polinucleótido en una biblioteca derivada de una primera muestra única es menos del 100% idéntica a la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca derivada de una o más muestras distintas de la primera muestra única. En algunos aspectos, una región de identificación de la muestra actúa como un código de barras digital en toda la primera hebra de ADNc transcrita a la reversa de una sola muestra.
- 50 En algunos aspectos, la región de identificación de la muestra tiene al menos 1 nucleótido de longitud. En algunos aspectos, una región de identificación de muestra puede comprender al menos 3 nucleótidos, y las regiones de identificación de muestra pueden diferir entre sí en al menos 1 nucleótido. En una realización, las regiones de identificación de muestra tienen una longitud de 3 a 15 nucleótidos y difieren entre sí en al menos 1 nucleótido. En algunos aspectos, las regiones de identificación de la muestra pueden comprender al menos 64 variantes (utilizando regiones de identificación de la muestra de 3 nucleótidos de longitud con cada identificación de la muestra que difieren entre sí por al menos 1 nucleótido), o en algunos aspectos, mayores números de variantes. En algunos aspectos, la secuencia unida 3' a la región de identificación de muestra puede ser una región adaptadora que comprende al menos 1 G. En una realización preferida, la secuencia unida 3' a la región de identificación de muestra puede ser una región adaptadora que comprende al menos 2 G's. En una realización, una secuencia unida al extremo 5' de una región de
- 55

identificación de muestra es una secuencia de cebador universal que se puede usar durante la amplificación por PCR para evitar la necesidad de la posterior adición de una secuencia de cebador universal 5' (por ligación u otro método) o el uso de múltiples cebadores en 5' degenerados para amplificar genes con regiones 5' variables. En las Tablas 2 y 8 se muestran ejemplos de regiones de identificación de muestras.

5 En algunos aspectos, la secuencia de la primera región de identificación de placa de cada polinucleótido en una biblioteca derivada de un primer conjunto de muestras únicas es distinta de la secuencia de la primera región de  
 10 identificación de placa de los otros polinucleótidos en la biblioteca derivada de uno o más conjuntos de muestra distintos del primer conjunto de muestras individuales. En algunos aspectos, la secuencia de la primera región de  
 15 identificación de placa de cada polinucleótido en una biblioteca derivada del primer conjunto de muestras únicas difiere en al menos 1 nucleótido de la secuencia de la primera región de identificación de placa de los otros polinucleótidos  
 en la biblioteca derivada de uno o más conjuntos de muestras individuales distintos del primer conjunto de muestras individuales. En algunos aspectos, la secuencia de la primera región de identificación de placa de cada polinucleótido  
 20 en una biblioteca derivada del primer conjunto de muestras únicas puede ser aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34,  
 25 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o más nucleótidos de la secuencia de la primera región de identificación de la placa de los otros polinucleótidos en la biblioteca derivada de uno o más conjuntos de muestras individuales distintos del primer conjunto  
 de muestras individuales. En algunos aspectos, la secuencia de la primera región de identificación de la placa de cada polinucleótido en una biblioteca derivada del primer conjunto de muestras únicas es menos del 100% idéntica a la  
 secuencia de la primera región de identificación de placa de los otros polinucleótidos en la biblioteca derivada de uno o más conjuntos de muestras individuales distintos del primer conjunto de muestras individuales. Los ejemplos de las  
 regiones de identificación de la primera placa se muestran en las Tablas 3 y 7.

En algunos aspectos, la secuencia de la segunda región de identificación de placa de cada polinucleótido en una  
 30 biblioteca derivada de un primer conjunto de muestras únicas es distinta de la secuencia de la segunda región de  
 identificación de placa de los otros polinucleótidos en la biblioteca derivada de uno o más conjuntos de muestras  
 distintas del primer conjunto de muestras individuales. En algunos aspectos, la secuencia de la segunda región de  
 35 identificación de placa de cada polinucleótido en una biblioteca derivada del primer conjunto de muestras individuales difiere en al menos 1 nucleótido de la secuencia de la segunda región de identificación de placa de los otros  
 polinucleótidos en la biblioteca derivada de uno o más conjuntos de muestras individuales distintos del primer conjunto  
 de muestras individuales. En algunos aspectos, la secuencia de la segunda región de identificación de placa de cada  
 polinucleótido en una biblioteca derivada del primer conjunto de muestras individuales difiere en al menos 2, 3, 4, 5,  
 40 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38,  
 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50 o más nucleótidos de la secuencia de la segunda región de identificación  
 de la placa de los otros polinucleótidos en la biblioteca derivada de uno o más conjuntos de muestras individuales  
 distintos del primer conjunto de muestras individuales. En algunos aspectos, la secuencia de la segunda región de  
 45 identificación de placa es idéntica a la secuencia de la primera región de identificación de placa en un polinucleótido.  
 En algunos aspectos, la secuencia de la segunda región de identificación de placa de cada polinucleótido en una  
 biblioteca derivada del primer conjunto de muestras únicas puede ser aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11  
 , 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42,  
 50 43, 44, 45, 46, 47, 48, 49, 50, 60, 70, 80, 81, 82, 83, 84, 85, 86, 87, 88, 89, 90, 91, 92, 93, 94, 95, 96, 97, 98, 99, o  
 menos del 100% idéntico a la secuencia de la segunda región de identificación de la placa de los otros polinucleótidos  
 en la biblioteca derivada de uno o más conjuntos de muestras distintas del primer conjunto de muestras individuales.  
 En algunos aspectos, la secuencia de la segunda región de identificación de placa de cada polinucleótido en una  
 biblioteca derivada del primer conjunto de muestras únicas es menos del 100% idéntica a la secuencia de la segunda  
 55 región de identificación de placa de los otros polinucleótidos en la biblioteca derivada de uno o más conjuntos de  
 muestras individuales distintos del primer conjunto de muestras individuales. En las Tablas 3 y 7 se muestran ejemplos  
 de regiones de identificación de la segunda placa.

En algunos aspectos, una región de identificación de placa (por ejemplo, una primera región de identificación de placa  
 o una segunda región de -identificación de placa) puede comprender al menos 2 nucleótidos, y las regiones de  
 55 identificación de placa difieren entre sí por al menos 1 nucleótido. En una realización, las regiones de identificación de  
 placa tienen una longitud de 2 a 10 nucleótidos y difieren entre sí en al menos 1 nucleótido. En algunos aspectos, el  
 uso de regiones de identificación de placa se encuentra en solo algunas realizaciones, ya que el uso de un número  
 mayor de regiones de identificación de muestra diferentes (una por muestra individual para analizar) puede eliminar la  
 necesidad de regiones de identificación de placa. En algunos aspectos, las regiones de identificación de la placa se  
 60 utilizan para reducir el número de oligonucleótidos únicos que contienen una región de identificación de la muestra  
 que deben sintetizarse.

En algunos aspectos, un polinucleótido incluye una o más regiones adaptadoras. En algunos aspectos, una región de adaptador incluye una o más G's. En algunos aspectos, una región de adaptador incluye 2, 3, 4, 5, 6, 7, 8, 9, 10 o más G's. En algunos aspectos, las regiones adaptadoras se adjuntan a los extremos 3' de los ADNc utilizando la propiedad de cambio de plantilla de MMLV H-transcriptasas reversas. Existen diferentes métodos para unir regiones adaptadoras, que incluyen, pero no se limitan a, hacer PCR con cebadores con secuencias de regiones adaptadoras 5' flanqueantes, ligación de extremos rígidos y pegajosos, adición de nucleótidos mediada por cambio de plantilla, u otros métodos para unir covalentemente los nucleótidos al extremo 5', al extremo 3' o a los extremos 5' y 3' de los polinucleótidos. Estos métodos pueden emplear propiedades de enzimas comúnmente usadas en biología molecular. La PCR puede usar, por ejemplo, ADN polimerasa termófila. Los extremos pegajosos que son complementarios o sustancialmente complementarios se crean a través del corte de dsADN con enzimas de restricción que dejan extremos sobresalientes o mediante actividades de cola de 3' de enzimas como TdT (transferasa terminal). Los extremos pegajosos y romos se pueden ligar con una región adaptadora complementaria utilizando ligasas como la ligasa T4. El cambio de plantilla utiliza la actividad de cola 3' de MMLV H transcriptasa reversa para agregar una o más citosinas (C) al extremo 3' de los ADNc y su capacidad para cambiar la plantilla de ARNm a una región adaptadora con G's complementarios. En algunos aspectos, un ADNc incluye 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más C en su extremo 3'.

En algunos aspectos, un polinucleótido incluye una o más regiones de sitios de restricción. Las regiones de sitios de restricción incluyen uno o más sitios de restricción. Los sitios de restricción pueden incluir: NheI, XhoI, BstBI, EcoRI, SacII, BbvCI, PspXI, AgeI, ApaI, KpnI, Acc65I, XmaI, BstEII, DraIII, PaeI, FseI, AsiSI y AscI. En algunos aspectos, se puede usar cualquier sitio de restricción de enzimas de 8 cortadores raros.

En algunos aspectos, una o más regiones de un polinucleótido descrito en el presente documento pueden estar acopladas operativamente a una o más regiones del polinucleótido. En algunos aspectos, dos o más regiones distintas de un único polinucleótido pueden estar acopladas operativamente. Por ejemplo, una región de cebador universal puede acoplarse operativamente a una región de adaptador. En algunos aspectos, se pueden acoplar operativamente dos o más regiones que son sustancialmente idénticas entre sí en secuencia o idénticas en descripción. Por ejemplo, una primera región de identificación de muestra puede estar acoplada operativamente a una segunda región de identificación de muestra. En algunos aspectos, las secuencias de la primera región de identificación de la muestra y la segunda región de identificación de la muestra son idénticas o sustancialmente idénticas. En algunos aspectos, las secuencias de la primera región de identificación de la muestra y la segunda región de identificación de la muestra son diferentes o distintas.

En algunos aspectos, una o más regiones de un polinucleótido descrito en el presente documento se pueden acoplar a una o más regiones del polinucleótido. En algunos aspectos, se pueden acoplar dos o más regiones distintas de un solo polinucleótido. Por ejemplo, una región de cebador universal se puede acoplar a una región de adaptador. En algunos aspectos, se pueden acoplar dos o más regiones que son sustancialmente idénticas entre sí en secuencia o idénticas en descripción. Por ejemplo, una primera región de identificación de muestra se puede acoplar a una segunda región de identificación de muestra. En algunos aspectos, las secuencias de la primera región de identificación de la muestra y la segunda región de identificación de la muestra son idénticas o sustancialmente idénticas. En algunos aspectos, las secuencias de la primera región de identificación de la muestra y la segunda región de identificación de la muestra son diferentes o distintas.

En algunos aspectos, un polinucleótido incluye la secuencia 5'-A-B-3', en donde A es una región de identificación de la muestra, y en donde B es una región adaptadora. En algunos aspectos, un polinucleótido incluye la secuencia 5'-A-B-C-3', en donde A es una región cebadora universal, en donde B es una región de identificación de la muestra y en donde C es una región adaptadora. En algunos aspectos, un polinucleótido incluye la secuencia 5'-A-B-C-3', en donde A es una región de identificación de la muestra, en donde B es una región adaptadora, y en donde C es una región de amplicón derivada de una sola muestra. En algunos aspectos, un polinucleótido incluye la secuencia 5'-ABCD-3', en donde A es una región cebadora universal, en donde B es una región de identificación de muestra, en donde C es una región adaptadora, y en donde D es una región de amplicón derivada de una muestra única. En algunos aspectos, un polinucleótido incluye la secuencia 5'-ABCDE-3', en donde A es una región de identificación de placa, en donde B es una región de cebador universal, en donde C es una región de identificación de muestra, en donde D es una región de adaptador, y en donde E es una región de amplicón derivada de una sola muestra. En algunos aspectos, un polinucleótido incluye la secuencia 5'-ABCDEF-3', en donde A es una primera región del sitio de restricción, en donde B es una región cebadora universal, en donde C es una región de identificación de muestra, en donde D es una región adaptadora, en donde E es una región de amplicón derivada de una sola muestra, y en donde F es una segunda región de sitio de restricción.

En algunos aspectos, las regiones de cada una de las secuencias anteriores se pueden reorganizar en un orden diferente, por ejemplo, 5'-C-A-D-B-3' o 5'-E-A-C-B-D-F-3' o 5'-B-A-3'. En algunos aspectos, se pueden eliminar una o más regiones de las secuencias anteriores, por ejemplo, 5'-A-D-3' o 5'-B-C-3'. En algunos aspectos, se pueden agregar una o más regiones adicionales a las secuencias anteriores, por ejemplo, 5'-A-A2-B-3' o 5'-A-B-C-D-E-F-G-3'. En tales ejemplos, la una o más regiones adicionales pueden ser cualquier región descrita en el presente documento o sus equivalentes. En algunos aspectos, una o más regiones de las secuencias anteriores se pueden modificar, por ejemplo, metilar.

En algunos aspectos, un polinucleótido puede incluir una molécula adaptadora. En algunos aspectos, una molécula adaptadora de polinucleótido puede incluir una región de cebador universal, una región de identificación de muestra y una región de adaptador, en donde el extremo 3' de la región de cebador universal está acoplada al extremo 5' de la región de identificación de muestra, y en donde el extremo 3' de la región de identificación de la muestra está acoplado al extremo 5' de la región del adaptador. En algunos aspectos, una molécula adaptadora incluye un polinucleótido que comprende al menos 2 nucleótidos que se unen a los C añadidos por una transcriptasa reversa en el extremo 3' de un ADNc de 1.<sup>a</sup> hebra. En algunos aspectos, una molécula adaptadora incluye un polinucleótido de desoxirribosa que comprende 3-6 G (ADN G). En otra realización, una molécula adaptadora incluye un polinucleótido de ribosa que consiste en 3-6 G (ARN G). En otras realizaciones, la molécula adaptadora puede utilizar análogos de nucleótidos, tales como ácidos nucleicos bloqueados (LNA), por ejemplo, LNA G. En otras realizaciones, la base de nucleótido también puede ser una base universal o degenerada tal como 5-nitroindol y 3-nitropirrol que puede emparejarse con C, así como con otros nucleótidos, en cualquier combinación.

En algunos aspectos, un polinucleótido puede incluir un cebador o una sonda. En algunos aspectos, un cebador puede incluir una región de cebador universal y una región de identificación de placa, y en donde el extremo 3' de la región de identificación de placa está acoplado al extremo 5' de la región de cebador universal.

En algunos aspectos, una composición puede incluir una biblioteca de composición de polinucleótidos. En algunos aspectos, una biblioteca de composiciones de polinucleótidos incluye una pluralidad de composiciones de polinucleótidos. En algunos aspectos cada composición está presente en un recipiente separado. En algunos aspectos, un recipiente puede ser un tubo de ensayo. En algunos aspectos, un recipiente puede ser un pozo en una placa. En algunos aspectos, un recipiente puede ser un pozo en una placa de 96 pozos. En algunos aspectos, un recipiente puede ser un pozo en una placa de 384 pozos. En algunos aspectos, cada composición comprende una región de ADNc derivada de una sola muestra. En algunos aspectos, cada composición comprende una región de adaptador de identificación de muestra que comprende una región de identificación de muestra acoplada a una región de adaptador. En algunos aspectos, la secuencia de la región de identificación de muestra de cada región de adaptador de identificación de muestra en una biblioteca es distinta de la secuencia de nucleótidos de la región de identificación de muestra de las otras regiones de adaptador de identificación de muestra presentes en cada contenedor separado en la biblioteca. En algunos aspectos, la región del adaptador de identificación de muestra se adjunta a la región de ADNc. En algunos aspectos, la región del adaptador de identificación de la muestra está unida a la región de ADNc mediante la unión entre sus regiones 3'. En algunos aspectos, la región del adaptador de identificación de la muestra está unida a la región de ADNc por la unión de G:C. En algunos aspectos, la región de ADNc comprende un polinucleótido de ARN hibridado con un polinucleótido de ADN. En algunos aspectos, la región de ADNc comprende un polinucleótido de ARNm hibridado a un polinucleótido de ADNc.

En algunos aspectos, la pluralidad de composiciones polinucleotídicas en una biblioteca de polinucleótidos puede comprender al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 3000, al menos 10 000, al menos 30 000, al menos 100 000, al menos 300 000, al menos 1 000 000, al menos 3 000 000, al menos 10 000 000, al menos 30 000 000, o más miembros. En otros aspectos, la pluralidad de composiciones polinucleotídicas en una biblioteca de polinucleótidos puede comprender al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 3000, al menos 10 000, al menos 30 000 o más genes del transcriptoma completo de una muestra de células. En otros aspectos, la pluralidad de composiciones polinucleotídicas en una biblioteca de polinucleótidos comprende al menos 1, al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 10 000, al menos 100 000, al menos 1 000 000, al menos 10 000 000, al menos 1 000 000 000 o más de las diferentes especies de anticuerpos presentes en la sangre de un individuo. Estas especies de anticuerpos pueden expresarse mediante plasmoblastos, células plasmáticas, células B de memoria, células plasmáticas de larga vida, células B ingenuas, otras células de linaje B o combinaciones de las mismas.

#### Vectores

En algunos aspectos, una composición puede incluir un vector. Los vectores se pueden usar en la transformación de una célula huésped con una secuencia de ácido nucleico. En algunos aspectos, un vector puede incluir uno o más polinucleótidos descritos en el presente documento. En una realización, puede introducirse una biblioteca de secuencias de ácido nucleico que codifican polipéptidos diana en una población de células, permitiendo de ese modo la selección de una biblioteca. El término "vector" se utiliza para referirse a una molécula portadora de ácido nucleico en donde se puede insertar una secuencia de ácido nucleico para su introducción en una célula donde se puede replicar. Una secuencia de ácido nucleico puede ser "exógena" o "heteróloga", lo que significa que es extraña a la célula en donde se está introduciendo el vector o que la secuencia es homóloga a una secuencia en la célula pero en una posición dentro del núcleo de la célula huésped. Ácido en donde normalmente no se encuentra la secuencia. Los vectores incluyen plásmidos, cósmidos y virus (por ejemplo, bacteriófagos). Un experto en la materia puede construir un vector mediante técnicas recombinantes estándar, que se describen en Maniatis et al., 1988 and Ausubel et al., 1994, ambas referencias se incorporan aquí como referencia. En algunos aspectos, un vector puede ser un vector con las regiones constantes de un anticuerpo pre-diseñado. De esta manera, un experto puede clonar solo las regiones VDJ de un anticuerpo de interés y clonar esas regiones en el vector pre-diseñado.

El término "vector de expresión" se refiere a un vector que contiene una secuencia de ácido nucleico que codifica al menos parte de un producto génico capaz de transcribirse. En algunos casos, las moléculas de ARN se traducen luego en una proteína, polipéptido o péptido. Los vectores de expresión pueden contener una variedad de "secuencias de control", que se refieren a secuencias de ácido nucleico para la transcripción y posiblemente la traducción de una secuencia codificante unida operativamente en un organismo huésped particular. Además de controlar las secuencias que gobiernan la transcripción y la traducción, los vectores y los vectores de expresión pueden contener secuencias de ácido nucleico que sirven también para otras funciones.

En algunos aspectos, un vector puede incluir un promotor. En algunos aspectos, un vector puede incluir un potenciador. Un "promotor" es una secuencia de control que es una región de una secuencia de ácido nucleico en donde se controlan el inicio y la velocidad de transcripción. Puede contener elementos genéticos a los que las proteínas y moléculas reguladoras pueden unirse, como la ARN polimerasa y otros factores de transcripción. Las frases "posicionado operativamente", "enlazado operativamente", "bajo control" y "bajo control transcripcional" significan que un promotor se encuentra en una ubicación y/u orientación funcional correcta en relación con una secuencia de ácido nucleico para controlar el inicio de la transcripción y/o expresión de esa secuencia. Un promotor puede o no ser usado en conjunto con un "potenciador", que se refiere a una secuencia reguladora que actúa en cis involucrada en la activación transcripcional de una secuencia de ácido nucleico.

Un promotor puede estar asociado de forma natural con un gen o secuencia, tal como puede obtenerse aislando las secuencias no codificantes 5' situadas corriente arriba del segmento de codificación y/o el exón. Dicho promotor puede denominarse "endógeno". De manera similar, un potenciador puede estar asociado naturalmente con una secuencia de ácido nucleico, ubicada corriente abajo o corriente arriba de esa secuencia. Alternativamente, se obtendrán ciertas ventajas colocando el segmento de ácido nucleico codificante bajo el control de un promotor recombinante o heterólogo, que se refiere a un promotor que normalmente no está asociado con una secuencia de ácido nucleico en su entorno natural. Un potenciador recombinante o heterólogo se refiere también a un potenciador que normalmente no está asociado con una secuencia de ácido nucleico en su entorno natural. Dichos promotores o potenciadores pueden incluir promotores o potenciadores de otros genes, y promotores o potenciadores aislados de cualquier otra célula procariótica, y promotores o potenciadores no "naturales", es decir, que contienen diferentes elementos de diferentes regiones reguladoras de la transcripción, y/o mutaciones que alteran la expresión. Además de producir secuencias de ácido nucleico de promotores y potenciadores sintéticamente, las secuencias pueden producirse utilizando la clonación recombinante y/o la tecnología de amplificación de ácido nucleico, incluida la PCR, en relación con las composiciones descritas en el presente documento (véase la Patente de Estados Unidos N° 4,683,202, Patente de Estados Unidos N° 5,928,906, cada una incorporada aquí como referencia).

En algunos aspectos, un promotor y/o potenciador que dirige efectivamente la expresión del segmento de ADN en el tipo de célula elegido para la expresión. Un ejemplo de dicho promotor que puede usarse es la arabinosa de *E. coli* o el promotor T7. Los expertos en la técnica de la biología molecular generalmente están familiarizados con el uso de promotores, potenciadores y combinaciones de tipo celular para la expresión de proteínas, por ejemplo, véase Sambrook et al. (1989), incorporado aquí como referencia. Los promotores empleados pueden ser constitutivos, específicos de tejido, inducibles y/o útiles en las condiciones apropiadas para dirigir la expresión de alto nivel del segmento de ADN introducido, tal como es ventajoso en la producción a gran escala de proteínas y/o péptidos recombinantes. El promotor puede ser heterólogo o endógeno.

En algunos aspectos, los vectores pueden incluir señales de iniciación y/o sitios de unión a ribosomas internos. También se puede incluir una señal de inicio específica para la traducción eficiente de las secuencias de codificación. Estas señales incluyen el codón de iniciación ATG o secuencias adyacentes. Es posible que se deban proporcionar señales de control de traducción exógenas, incluido el codón de iniciación ATG. Un experto en la técnica sería fácilmente capaz de determinar esto y proporcionar las señales necesarias. Es bien sabido que el codón de iniciación debe estar "en el marco" con el marco de lectura de la secuencia de codificación deseada para garantizar la traducción de todo el inserto. Las señales de control de traducción exógenas y los codones de iniciación pueden ser naturales o sintéticos. La eficacia de la expresión puede mejorarse mediante la inclusión de elementos potenciadores de la transcripción apropiados.

En algunos aspectos, un vector puede incluir secuencias que aumentan u optimizan el nivel de expresión del segmento de ADN que codifica el gen de interés. Un ejemplo de tales secuencias incluye la adición de intrones en el ARNm expresado (Brinster, RL et al. (1988). Introns increase transcriptional efficiency in transgenic mice. Proc. Natl. Acad. Sci. USA 85, 836-40; Choi, T. et al. (1991) A generic intron increases gene expression in transgenic mice. Mol. Cell. Biol. 11, 3070-4). Otro ejemplo de un método para optimizar la expresión del segmento de ADN es la "optimización de codones". La optimización de codones implica la inserción de mutaciones silenciosas en el segmento de ADN para reducir el uso de codones raros para optimizar la traducción de proteínas (Codon engineering for improved antibody expression in mammalian cells. Carton JM, Sauerwald T, Hawley-Nelson P, Morse B, Peffer N, Beck H, Lu J, Cotty A, Amegadzie B, Sweet R. Protein Expr Purif. 2007 Oct;55(2):279-86. Epub 2007 Jun 16.).

En algunos aspectos, un vector puede incluir múltiples sitios de clonación. Los vectores pueden incluir un sitio de clonación múltiple (MCS), que es una región de ácido nucleico que contiene múltiples sitios de enzimas de restricción, cualquiera de los cuales puede usarse junto con la tecnología recombinante estándar para digerir el vector (véase

- Carbonelli et al., 1999, Levenson et al., 1998, y Cocea, 1997, incorporados aquí como referencia.) "Digestión de enzimas de restricción" se refiere a la escisión catalítica de una molécula de ácido nucleico con una enzima que funciona solo en lugares específicos en una molécula de ácido nucleico. Muchas de estas enzimas de restricción están disponibles comercialmente. El uso de tales enzimas es entendido por los expertos en la técnica. Con frecuencia, un
- 5 vector se linealiza o se fragmenta utilizando una enzima de restricción que corta dentro del MCS para permitir que las secuencias exógenas se unan al vector. "Ligación" se refiere al proceso de formación de enlaces fosfodiéster entre dos fragmentos de ácido nucleico, que pueden o no ser contiguos entre sí. Las técnicas que implican enzimas de restricción y reacciones de ligación son bien conocidas por los expertos en la técnica de la tecnología recombinante.
- 10 En algunos aspectos, un vector puede incluir una señal de terminación. Los vectores o construcciones generalmente comprenderán al menos una señal de terminación. Una "señal de terminación" o "terminador" comprende las secuencias de ADN involucradas en la terminación específica de un transcrito de ARN por una ARN polimerasa. Por lo tanto, en ciertas realizaciones, se contempla una señal de terminación que termina la producción de un transcrito de ARN. Un terminador puede ser necesario in vivo para lograr niveles de mensaje deseables.
- 15 Los terminadores contemplados para su uso incluyen cualquier terminador conocido de transcripción descrito en el presente documento o conocido por un experto en la técnica, que incluye, entre otros, terminadores dependientes de rho o independientes de rho. En ciertas realizaciones, la señal de terminación puede ser una falta de secuencia transcribible o traducible, tal como debido a un truncamiento de secuencia.
- En algunos aspectos, un vector puede incluir un origen de replicación.
- 20 Para propagar un vector en una célula huésped, puede contener uno o más orígenes de sitios de replicación (a menudo denominados "ori"), que es una secuencia de ácido nucleico específica en donde se inicia la replicación.
- En algunos aspectos, un vector puede incluir uno o más marcadores seleccionables y/o seleccionables. En ciertas realizaciones, las células que contienen una construcción de ácido nucleico pueden identificarse in vitro o in vivo incluyendo un marcador en el vector de expresión. Tales marcadores conferirían un cambio identificable a la célula permitiendo una fácil identificación de las células que contienen el vector de expresión. Generalmente, un marcador
- 25 seleccionable es aquel que confiere una propiedad que permite la selección. Un marcador seleccionable positivo es aquel en donde la presencia del marcador permite su selección, mientras que un marcador seleccionable negativo es aquel en donde su presencia impide su selección. Un ejemplo de un marcador seleccionable positivo es un marcador de resistencia a fármacos.
- 30 Por lo general, la inclusión de un marcador de selección de fármacos ayuda en la clonación e identificación de los transformantes, por ejemplo, los genes que confieren resistencia a la neomicina, puromicina, higromicina, DHFR, GPT, zeocina e histidinol son marcadores seleccionables útiles. Además de los marcadores que confieren un fenotipo que permite la discriminación de los transformantes en función de la implementación de las condiciones, también se contemplan otros tipos de marcadores que incluyen marcadores detectables, como el GFP, cuya base es el análisis
- 35 colorimétrico. Alternativamente, pueden utilizarse enzimas analizables tales como cloranfenicol acetiltransferasa (CAT). Un experto en la técnica también sabría cómo emplear marcadores inmunológicos, posiblemente junto con el análisis FACS. El marcador utilizado no se considera importante, siempre que sea capaz de expresarse simultáneamente con el ácido nucleico que codifica un producto genético. Un experto en la técnica conoce bien otros ejemplos de marcadores seleccionables y seleccionables.
- 40 En un aspecto, el vector puede expresar segmentos de ADN que codifican múltiples polipéptidos de interés. Por ejemplo, los segmentos de ADN que codifican tanto la cadena pesada de la inmunoglobulina como la cadena liviana se pueden codificar y expresar mediante un solo vector. En un aspecto, ambos segmentos de ADN pueden incluirse en las mismas secuencias de ARN expresado y sitio de unión al ribosoma interno (IRES) que se utilizan para permitir la expresión de los segmentos de ADN como polipéptidos separados (Pinkstaff JK, Chappell SA, Mauro VP, Edelman GM, Krushel LA., Internal initiation of translation of five dendritically localized neuronal mRNAs., Proc Natl Acad Sci U
- 45 S A. 2001 Feb 27;98(5):2770-5. Epub 2001 Feb 20.). En otro aspecto, cada segmento de ADN tiene su propia región promotora que resulta en la expresión de ARNm separados (Andersen CR, Nielsen LS, Baer A, Tolstrup AB, Weilguny D. Efficient Expression from One CMV Enhancer Controlling Two Core Promoters. Mol Biotechnol. 2010 Nov 27. [Epub ahead of print]).
- Células huésped y Sistemas de Expresión.
- 50 En algunos aspectos, una composición puede incluir una célula huésped. En algunos aspectos, una célula huésped puede incluir un polinucleótido o vector descrito en el presente documento. En algunos aspectos, una célula huésped puede incluir una célula eucariota (por ejemplo, insecto, levadura o mamífero) o una célula procariota (por ejemplo, bacterias). En el contexto de la expresión de una secuencia de ácido nucleico heteróloga, "célula huésped" puede referirse a una célula procariótica, e incluye cualquier organismo transformable que sea capaz de replicar un vector
- 55 y/o expresar un gen heterólogo codificado por un vector. Una célula huésped puede, y ha sido, utilizada como receptor de vectores. Una célula huésped puede ser "transfectada" o "transformada", que se refiere a un proceso por el cual el

ácido nucleico exógeno se transfiere o se introduce en la célula huésped. Una célula transformada incluye la célula sujeto primaria y su progenie.

5 En realizaciones particulares, una célula huésped es una célula bacteriana Gram negativa. Estas bacterias son adecuadas para su uso ya que poseen un espacio periplásmico entre la membrana interna y externa y, en particular, la membrana interna mencionada anteriormente entre el periplasma y el citoplasma, que también se conoce como membrana citoplásmica. Como tal, podría usarse cualquier otra célula con tal espacio periplásmico. Los ejemplos de bacterias Gram negativas incluyen, pero no se limitan a, *E. coli*, *Pseudomonas aeruginosa*, *Vibrio cholera*, *Salmonella typhimurium*, *Shigella flexneri*, *Haemophilus influenza*, *Bordetella pertussi*, *Erwinia amylovora*, *Rhizobium sp.* La célula bacteriana Gram negativa puede definirse aún más como célula bacteriana que se ha transformado con la secuencia codificante de un polipéptido de fusión que comprende un polipéptido de unión candidato capaz de unirse a un ligando seleccionado. El polipéptido está anclado a la cara externa de la membrana citoplásmica, mirando hacia el espacio periplásmico, y puede comprender una secuencia de codificación de anticuerpos u otra secuencia. Un medio para la expresión del polipéptido es unir una secuencia líder al polipéptido capaz de causar dicha dirección.

15 Numerosas líneas celulares y cultivos procarióticos están disponibles para su uso como una célula huésped, y pueden obtenerse a través de la American Type Culture Collection (ATCC), que es una organización que sirve como un archivo para cultivos vivos y materiales genéticos. Un huésped apropiado puede ser determinado por un experto en la técnica con base en la estructura del vector y el resultado deseado. Un plásmido o cósmido, por ejemplo, puede introducirse en una célula huésped procariota para la replicación de muchos vectores. Las células bacterianas utilizadas como células huésped para la replicación y/o expresión del vector incluyen DH5-alfa, JM109 y KC8, así como una serie de hospedadores bacterianos disponibles comercialmente, tales como Células Competentes SURE™ y Células Doradas SOLOPACK™ (STRATAGENE™, La Jolla). En algunos aspectos, se contemplan otras células bacterianas tales como *E. coli* LE392 para uso como células huésped.

25 Muchas células huésped de diversos tipos de células y organismos están disponibles y serían conocidas por los expertos en la técnica. De manera similar, un vector viral se puede usar junto con una célula huésped procariota, particularmente una que es permisiva para la replicación o expresión del vector. Algunos vectores pueden emplear secuencias de control que permiten su replicación y/o expresión en células procariotas y eucariotas. Un experto en la técnica entenderá además las condiciones en las cuales incubar todas las células huésped descritas anteriormente para mantenerlas y permitir la replicación de un vector. También se entienden y se conocen técnicas y condiciones que permitirían la producción a gran escala de vectores, así como la producción de los ácidos nucleicos codificados por vectores y sus polipéptidos, proteínas o péptidos afines.

30 En algunos aspectos, una célula huésped es de mamífero. Los ejemplos incluyen células CHO, células CHO-K1 o células CHO-S. Otras células huésped de mamíferos incluyen células NS0 y células CHO que son dhfr-, por ejemplo, CHO-dhfr-, células DUKX-B11 CHO y células DG44 CHO.

35 Existen numerosos sistemas de expresión que pueden comprender al menos una parte o la totalidad de las composiciones descritas en este documento. Los sistemas de expresión pueden incluir sistemas de expresión eucarióticos y sistemas de expresión procarióticos. Tales sistemas podrían usarse, por ejemplo, para la producción de un producto polipeptídico identificado como capaz de unirse a un ligando particular. Los sistemas basados en procariotas pueden emplearse para producir secuencias de ácido nucleico, o sus polipéptidos, proteínas y péptidos afines. Muchos de estos sistemas están disponibles comercialmente y ampliamente. Otros ejemplos de sistemas de expresión comprenden vectores que contienen un promotor procariótico fuerte tal como T7, Tac, Trc, BAD, lambda pL, tetraciclina o Lac promotores, el sistema de expresión pET y un sistema de expresión de *E. coli*.

#### Polipéptidos

45 En algunos aspectos, una composición puede incluir un polipéptido. En algunos aspectos, un polipéptido codificado por un polinucleótido descrito en el presente documento puede expresarse, por ejemplo, desde una célula huésped. Los términos "polipéptido" o "proteína" incluyen una macromolécula que tiene la secuencia de aminoácidos de una proteína nativa, es decir, una proteína producida por una célula natural y no recombinante; o se produce por una célula genéticamente modificada o recombinante, y comprende moléculas que tienen la secuencia de aminoácidos de la proteína nativa, o moléculas que tienen eliminaciones, adiciones y/o sustituciones de uno o más aminoácidos de la secuencia nativa. El término también incluye polímeros de aminoácidos en los que uno o más aminoácidos son análogos químicos de un aminoácido y polímeros de origen natural correspondientes. Los términos "polipéptido" y "proteína" abarcan proteínas de unión a antígeno, anticuerpos o secuencias que tienen eliminaciones, adiciones y/o sustituciones de uno o más aminoácidos de proteína de unión a antígeno. El término "fragmento de polipéptido" se refiere a un polipéptido que tiene una eliminación amino-terminal, una eliminación carboxilo-terminal y/o una eliminación interna en comparación con la proteína nativa de longitud completa. Dichos fragmentos también pueden contener aminoácidos modificados en comparación con la proteína nativa. En ciertas realizaciones, los fragmentos tienen una longitud de aproximadamente cinco a 500 aminoácidos. Por ejemplo, los fragmentos pueden tener al menos 5, 6, 8, 10, 14, 20, 50, 70, 100, 110, 150, 200, 250, 300, 350, 400 o 450 aminoácidos de longitud. Los fragmentos de polipéptidos útiles incluyen fragmentos inmunológicamente funcionales de anticuerpos, incluyendo dominios de unión. En el caso de un anticuerpo de unión, los fragmentos útiles incluyen, pero no se limitan a, una región CDR, un dominio

variable de una cadena pesada y/o liviana, una parte de una cadena de anticuerpo o solo su región variable que incluye dos CDR, y similares.

5 El término "proteína aislada" significa que una proteína sujeto (1) está libre de al menos algunas otras proteínas con las que normalmente se encontraría, (2) está esencialmente libre de otras proteínas de la misma fuente, por ejemplo, de la misma especie, (3) se expresa por una célula de una especie diferente, (4) se ha separado de al menos aproximadamente el 50 por ciento de los polinucleótidos, lípidos, carbohidratos u otros materiales con los cuales está asociado en la naturaleza, (5) está asociado de manera operativa (por interacción covalente o no covalente) con un polipéptido con el que no está asociado en la naturaleza, o (6) no ocurre en la naturaleza. Típicamente, una "proteína aislada" constituye al menos aproximadamente el 5%, al menos aproximadamente el 10%, al menos aproximadamente el 25%, o al menos aproximadamente el 50% de una muestra dada. El ADN genómico, el ADNc, el ARNm u otro ARN, los ácidos nucleicos de origen sintético o cualquier combinación de los mismos pueden codificar dicha proteína aislada. Preferiblemente, la proteína aislada está sustancialmente libre de proteínas o polipéptidos u otros contaminantes que se encuentran en su entorno natural que podrían interferir con su uso terapéutico, diagnóstico, profiláctico, investigación u otro.

15 En algunos aspectos, un polipéptido puede incluir una proteína de unión a antígeno (ABP). Una "proteína de unión a antígeno" ("ABP") como se usa en este documento significa cualquier proteína que se una a un antígeno diana específico. La "proteína de unión a antígeno" incluye, pero no se limita a, anticuerpos y partes de unión de los mismos, tales como fragmentos inmunológicamente funcionales. Los peptidocuerpos son otro ejemplo de proteínas de unión a antígeno. El término "fragmento inmunológicamente funcional" (o simplemente "fragmento") de un anticuerpo o proteína de unión al antígeno de la cadena de la inmunoglobulina (cadena pesada o liviana), como se usa en este documento, es una especie de proteína de unión al antígeno que comprende una porción (independientemente de cómo esa porción se obtiene o se sintetiza) de un anticuerpo que carece de al menos algunos de los aminoácidos presentes en una cadena de longitud completa pero que aún es capaz de unirse específicamente a un antígeno. Dichos fragmentos son biológicamente activos porque se unen al antígeno diana y pueden competir con otras proteínas de unión a antígeno, incluidos los anticuerpos intactos, por la unión a un epítipo dado. En algunas realizaciones, los fragmentos son fragmentos neutralizantes. Estos fragmentos biológicamente activos pueden producirse mediante técnicas de ADN recombinante, o pueden producirse mediante escisión enzimática o química de proteínas de unión a antígeno, incluyendo anticuerpos intactos. Los fragmentos de inmunoglobulina inmunológicamente funcionales incluyen, pero no se limitan a, Fab, un diacuerpo (dominio variable de cadena pesada en el mismo polipéptido que un dominio variable de cadena liviana, conectado a través de un enlace péptido corto que es demasiado corto para permitir el emparejamiento entre los dos dominios en la misma cadena), Fab', F (ab')<sub>2</sub>, Fv, anticuerpos de dominio y anticuerpos de cadena única, y pueden derivarse de cualquier fuente de mamíferos, incluidos, entre otros, humanos, ratones, ratas, camélidos o conejos. Se contempla además que una parte funcional de las proteínas de unión a antígeno descritas en el presente documento, por ejemplo, una o más CDR, podrían unirse covalentemente a una segunda proteína o a una pequeña molécula para crear un agente terapéutico dirigido a un objetivo particular en el cuerpo, poseyendo propiedades terapéuticas bifuncionales, o teniendo una vida media sérica prolongada. Como apreciará un experto en la técnica, una proteína de unión a antígeno puede incluir componentes no proteicos. Detalles adicionales sobre las proteínas de unión a antígeno y los anticuerpos, tales como modificaciones, variantes, métodos de preparación y métodos de selección se pueden encontrar en la publicación de patente de Estados Unidos 20110027287, incorporada aquí como referencia en su totalidad para todos los fines.

45 En algunos aspectos, un polipéptido puede incluir un anticuerpo. El término "anticuerpo" se refiere a una inmunoglobulina intacta de cualquier isotipo, o un fragmento del mismo que puede competir con el anticuerpo intacto por la unión específica al antígeno diana, e incluye, por ejemplo, anticuerpos quiméricos, humanizados, completamente humanos y biespecíficos. Un "anticuerpo" es una especie de una proteína de unión a antígeno. Un anticuerpo intacto generalmente comprenderá al menos dos cadenas pesadas de longitud completa y dos cadenas livianas de longitud completa, pero en algunos casos puede incluir menos cadenas, como los anticuerpos que ocurren naturalmente en camélidos, que pueden comprender solo cadenas pesadas. Los anticuerpos pueden derivar únicamente de una fuente única, o pueden ser "quiméricos", es decir, diferentes porciones del anticuerpo pueden derivarse de dos anticuerpos diferentes. Las proteínas de unión a antígeno, anticuerpos o fragmentos de unión pueden producirse en hibridomas, por técnicas de ADN recombinante, o por escisión enzimática o química de anticuerpos intactos. A menos que se indique lo contrario, el término "anticuerpo" incluye, además de los anticuerpos que comprenden dos cadenas pesadas de longitud completa y dos cadenas livianas de longitud completa, derivados, variantes, fragmentos y muteínas de las mismas. Además, a menos que se excluyan explícitamente, los anticuerpos incluyen anticuerpos monoclonales, anticuerpos biespecíficos, minicuerpos, anticuerpos de dominio, anticuerpos sintéticos (a veces referidos en el presente documento como "miméticos de anticuerpos"), anticuerpos quiméricos, anticuerpos humanizados, anticuerpos humanos, fusiones de anticuerpos (a veces referidos en el presente documento como "conjugados de anticuerpos"), y fragmentos de los mismos, respectivamente. En algunas realizaciones, el término también abarca peptidocuerpos.

60 Se puede administrar una cantidad terapéuticamente eficaz de un ABP a un sujeto que lo necesite. Las ABP pueden formularse en composiciones farmacéuticas. Estas composiciones pueden comprender, además de una o más de las ABP, un excipiente, portador, regulador, estabilizador u otros materiales farmacéuticamente aceptables bien conocidos por los expertos en la técnica. Dichos materiales no deben ser tóxicos y no deben interferir con la eficacia del



ingrediente activo. La naturaleza precisa del portador u otro material puede depender de la vía de administración, por ejemplo oral, intravenoso, cutáneo o subcutáneo, nasal, intramuscular, intraperitoneal.

5 Las composiciones farmacéuticas para administración oral pueden estar en forma de tabletas, cápsulas, polvo o líquido. Una tableta puede incluir un portador sólido tal como gelatina o un adyuvante. Las composiciones farmacéuticas líquidas generalmente incluyen un vehículo líquido tal como agua, petróleo, aceites animales o vegetales, aceite mineral o aceite sintético. Se pueden incluir soluciones salinas fisiológicas, dextrosa u otra solución de sacáridos o glicoles como etilenglicol, propilenglicol o polietilenglicol.

10 Para inyección intravenosa, cutánea o subcutánea, o inyección en el lugar de la afección, el ingrediente activo estará en forma de una solución acuosa parenteralmente aceptable que está libre de pirógenos y tiene un pH, isotonicidad y estabilidad adecuados. Los expertos en la técnica pueden preparar soluciones adecuadas usando, por ejemplo, vehículos isotónicos tales como inyección de cloruro de sodio, inyección de Ringer, inyección de Ringer lactato. Se pueden incluir conservantes, estabilizantes, reguladores, antioxidantes y/u otros aditivos, según se requiera.

15 La administración de ABP es preferiblemente en una "cantidad terapéuticamente eficaz" o "cantidad profilácticamente eficaz" (según sea el caso, aunque la profilaxis puede considerarse terapia), siendo esto suficiente para mostrar beneficios para el individuo. La cantidad real administrada, y la velocidad y el tiempo de administración dependerán de la naturaleza y la gravedad de la enfermedad que se esté tratando. Prescripción de tratamiento, por ejemplo las decisiones sobre la dosificación, etc., están a cargo de los médicos generales y otros médicos, y generalmente tienen en cuenta el trastorno a tratar, la condición del paciente individual, el sitio de entrega, el método de administración y otros factores conocidos por practicantes. Ejemplos de las técnicas y protocolos mencionados anteriormente se pueden encontrar en Remington's Pharmaceutical Sciences, 16th edition, Osol, A. (ed), 1980.

Una composición puede administrarse sola o en combinación con otros tratamientos, ya sea simultánea o secuencialmente dependiendo de la condición a tratar.

#### Células inmunes

25 Una muestra puede incluir células inmunes. Las células inmunes pueden incluir células T y células B. Las células T (linfocitos T) incluyen, por ejemplo, células que expresan receptores de células T. Las células B incluyen, por ejemplo, células B activadas, células B de memoria, células B1, células B2, células B de la zona marginal y células B foliculares. Las células T incluyen células T activadas, células T blásticas, células T cooperadoras (células T efectoras o células Th), células T citotóxicas (CTL), células T de memoria, células T de memoria central, células T de memoria efectora y células T reguladoras. Una muestra puede incluir una sola célula en algunas aplicaciones (por ejemplo, una prueba de calibración para definir células T o B relevantes) o más generalmente al menos 1000, al menos 10 000, al menos 100 000, al menos 250 000, al menos 500 000, al menos 750 000, o al menos 1 000 000 células.

#### Células B

35 Como se usa en este documento, "célula B" se refiere a cualquier célula que tenga al menos un locus genético de inmunoglobulina reorganizado. Una célula B puede incluir al menos un locus de cadena pesada de inmunoglobulina reorganizada o al menos un locus de cadena liviana de inmunoglobulina reorganizada. Una célula B puede incluir al menos un locus de cadena pesada de inmunoglobulina reorganizada y al menos un locus de cadena liviana de inmunoglobulina reorganizada. Las células B son linfocitos que forman parte del sistema inmunitario adaptativo. Las células B pueden incluir cualquier célula que exprese anticuerpos en forma de membrana unida como el receptor de células B (BCR) en la superficie celular o como anticuerpos secretados. Las células B pueden expresar inmunoglobulinas (anticuerpos, receptor de células B). Los anticuerpos pueden incluir heterodímeros formados a partir de las cadenas de inmunoglobulina pesadas y livianas. La cadena pesada se forma a partir de reordenamientos genéticos de los genes de la variable, la diversidad y la unión (VDJ) para formar la región variable, que está unida a la región constante. La cadena liviana se forma a partir de reordenamientos genéticos de los genes variable y de unión (VJ) para formar la región variable, que luego se une a la región constante. Debido a un gran número posible de combinaciones de unión, las regiones variables del gen del anticuerpo (que también es el BCR) tienen una gran diversidad, lo que permite a las células B reconocer cualquier antígeno extraño y montar una respuesta contra él.

#### Activación y diferenciación de células B

50 Las células B se activan y diferencian cuando reconocen un antígeno en el contexto de una respuesta inmune inflamatoria. Por lo general, incluyen 2 señales para activarse, una señal administrada a través de BCR (una forma unida a la membrana de la inmunoglobulina reorganizada) y otra administrada a través de CD40 u otra molécula coestimuladora. Esta segunda señal puede proporcionarse a través de la interacción con células T auxiliares, que expresan el ligando para CD40 (CD40L) en su superficie. Luego, las células B proliferan y pueden sufrir una hipermutación somática, donde se realizan cambios aleatorios en las secuencias de nucleótidos de los genes de anticuerpos, y se seleccionan las células B cuyos anticuerpos tienen una mayor afinidad de células B. También pueden sufrir un "cambio de clase", en donde la región constante de la cadena pesada que codifica el isotipo IgM se cambia

a la región constante que codifica el isotipo IgG, IgA o IgE. La diferenciación de las células B puede terminar como células B de memoria, que suelen ser de mayor afinidad y con cambio de clase, aunque algunas células B de memoria aún son del isotipo IgM. Las células B de memoria también pueden activarse y diferenciarse en plasmoblastos y, en última instancia, en células plasmáticas. Las células B diferenciadoras también pueden convertirse primero en plasmoblastos, que luego se diferencian para convertirse en células plasmáticas.

Maduración por afinidad y familias clonales.

Una familia clonal se define generalmente por el uso de secuencias relacionadas con la cadena pesada de inmunoglobulina y/o la cadena liviana V(D)J por 2 o más muestras. Las secuencias V(D)J de la cadena pesada de inmunoglobulina relacionadas pueden identificarse por su uso compartido de los segmentos del gen V(D)J codificados en el genoma. Dentro de una familia clonal hay generalmente subfamilias que varían según las mutaciones compartidas dentro de sus segmentos V(D)J, que pueden surgir durante la recombinación de genes de células B y la hipermutación somática.

Las células B activadas migran y forman centros germinales dentro de los tejidos linfoides u otros tejidos, donde experimentan una maduración de afinidad. Las células B también pueden experimentar una maduración de afinidad fuera de los centros germinales. Durante la maduración de la afinidad, las células B experimentan mutaciones aleatorias en sus genes de anticuerpos, concentradas en las regiones determinantes complementarias (CDR) de los genes, que codifican las partes del anticuerpo que se unen directamente y reconocen el antígeno objetivo contra el cual se activó la célula B. Esto crea subclones a partir de las células B en proliferación originales que expresan inmunoglobulinas que son ligeramente diferentes del clon original y entre sí. Los clones compiten por el antígeno y los clones de mayor afinidad se seleccionan, mientras que los clones de menor afinidad mueren por apoptosis. Este proceso da como resultado la "maduración de la afinidad" de las células B y, en consecuencia, la generación de células B que expresan inmunoglobulinas que se unen al antígeno con mayor afinidad. Todas las células B que se originan a partir de las mismas células B 'progenitoras' forman familias clonales, y estas familias clonales incluyen células B que reconocen epítomos antigénicos iguales o similares. En algunos aspectos, se espera que los clones presentes en frecuencias más altas representen clones que se unen al antígeno con mayor afinidad, porque los clones de mayor afinidad se seleccionan durante la maduración de la afinidad. En algunos aspectos, los clones con diferente uso del segmento V(D)J exhiben diferentes características de enlace. En algunos aspectos, los clones con el mismo uso del segmento V(D)J pero diferentes mutaciones exhiben diferentes características de unión.

Células B de memoria

Las células B de memoria suelen ser células B maduras por afinidad y pueden ser de clase cambiada. Estas son células que pueden responder más rápidamente a un desafío antigénico posterior, reduciendo significativamente el tiempo incluido para la secreción de anticuerpos madurados por afinidad contra el antígeno de aproximadamente ~14 días en un organismo sin tratamiento previo a aproximadamente ~7 días.

Plasmoblastos y células plasmáticas.

Las células plasmáticas pueden ser de larga vida o de corta vida. Las células plasmáticas de larga vida pueden sobrevivir durante toda la vida del organismo, mientras que las células plasmáticas de vida corta pueden durar de 3 a 4 días. Las células plasmáticas de larga vida residen en áreas de inflamación, en las áreas de la mucosa (en el caso de células plasmáticas secretoras de IgA), en tejidos linfoides secundarios (como el bazo o los ganglios linfáticos) o en la médula ósea. Para alcanzar estas áreas divergentes, los plasmoblastos destinados a convertirse en células plasmáticas de larga vida pueden viajar primero a través del torrente sanguíneo antes de utilizar varios gradientes de quimioquinas para transitar a las áreas apropiadas. Los plasmoblastos son células que maduran por afinidad, generalmente se cambian de clasificación y generalmente secretan anticuerpos, aunque generalmente en cantidades más bajas que la cantidad de anticuerpo producido por las células plasmáticas. Las células plasmáticas son secretoras de anticuerpos dedicados.

Características de las series TCR y BCR.

Dado que las recombinaciones de identificación están presentes en el ADN de cada célula inmune adaptativa individual, así como sus transcripciones de ARN asociadas, se puede secuenciar el ARN o el ADN. Una secuencia recombinada de una célula T o una célula B también puede denominarse un clonotipo. El ADN o el ARN pueden corresponder a secuencias de genes del receptor de células T (TCR) o genes de inmunoglobulina (Ig) que codifican anticuerpos. Por ejemplo, el ADN y el ARN pueden corresponder a secuencias que codifican cadenas alfa, beta, gamma o delta de un TCR. En la mayoría de las células T, el TCR es un heterodímero que consiste en una cadena alfa y una cadena beta. La cadena TCR-alfa se genera por recombinación VJ, y el receptor de la cadena beta se genera por recombinación V(D)J. Para la cadena TCR-beta, en humanos hay 48 segmentos V, 2 segmentos D y 13 segmentos J. Se pueden eliminar varias bases y agregar otras (llamadas nucleótidos N y P) en cada una de las dos uniones. En una minoría de células T, los TCR consisten en cadenas gamma y delta. La cadena gamma TCR se genera por recombinación VJ, y la cadena delta TCR se genera por recombinación V(D)J (Kenneth Murphy, Paul

Travers, and Mark Walport, Janeway's Immunology 7th edition, Garland Science, 2007, que se incorpora aquí como referencia en su totalidad).

5 El ADN y el ARN analizados en los métodos pueden corresponder a secuencias que codifican inmunoglobulinas de cadena pesada (IgH) con regiones constantes (alfa, delta, gamma, épsilon o mu) o inmunoglobulinas de cadena liviana (IgK o IgL) con regiones constantes lambda o kappa. Cada anticuerpo puede tener dos cadenas livianas idénticas y dos cadenas pesadas idénticas. Cada cadena está compuesta por una constante (C) y una región variable. Para la cadena pesada, la región variable se compone de una variable (V), diversidad (D) y segmentos de unión (J). Varias secuencias distintas que codifican para cada tipo de estos segmentos están presentes en el genoma. Un evento específico de recombinación VDJ ocurre durante el desarrollo de una célula B, marcando esa célula para generar una  
10 cadena pesada específica. La diversidad en la cadena liviana se genera de manera similar, excepto que no hay una región D, por lo que solo hay una recombinación VJ. La mutación somática a menudo ocurre cerca del sitio de la recombinación, causando la adición o eliminación de varios nucleótidos, lo que aumenta aún más la diversidad de las cadenas pesadas y livianas generadas por las células B. La posible diversidad de los anticuerpos generados por una célula B es, entonces, el producto de las diferentes cadenas pesadas y livianas. Las regiones variables de las cadenas pesada y liviana contribuyen a formar la región o sitio de reconocimiento (o unión) del antígeno. A esta diversidad se agrega un proceso de hipermutación somática que puede ocurrir después de que una respuesta específica se monta contra algún epítipo. En este proceso, se producen mutaciones en aquellas células B que son capaces de reconocer el epítipo específico, lo que conduce a una mayor diversidad de anticuerpos que pueden unirse más fuertemente al epítipo específico. Todos estos factores contribuyen a una gran diversidad de anticuerpos generados por las células B. Se pueden generar muchos miles de millones y quizás más de un billón de anticuerpos distintos. La premisa básica para generar diversidad de células-T es similar a la de generar anticuerpos por células B. Un elemento de la activación de las células-T y células B es su unión a los epítopos. La activación de una célula específica conduce a la producción de más del mismo tipo de células que conducen a una expansión clonal.

25 Las regiones determinantes de complementariedad (CDR), o regiones hipervariables, son secuencias en los dominios variables de los receptores de antígeno (por ejemplo, receptor de células T e inmunoglobulina) que pueden unirse a un antígeno. La cadena de cada receptor de antígeno contiene tres CDR (CDR1, CDR2 y CDR3). Los dos polipéptidos que producen las células T (alfa y beta) y la inmunoglobulina (IgH e IgK o IgL) contribuyen a la formación de las tres CDR.

30 La parte de CDR1 y CDR2 que está codificada por TCR-beta se encuentra dentro de uno de los 47 segmentos V funcionales. La mayor parte de la diversidad de las CDR se encuentra en la CDR3, y la diversidad se genera por los eventos de recombinación somática durante el desarrollo de los linfocitos T.

35 Una gran diversidad de BCR está presente inter e intraindividuos. El BCR está compuesto por dos genes IgH e IgK (o IgL) que codifican cadenas pesadas y livianas de anticuerpos. Las tres secuencias de la Región de Determinación de Complementariedad (CDR) que se unen a antígenos y moléculas MHC tienen la mayor diversidad en IgH e IgK (o IgL). La parte de CDR1 y CDR2 codificada por IgH se encuentra dentro de uno de los 44 segmentos V funcionales. La mayor parte de la diversidad en células B ingenua emerge en la generación de CDR3 a través de eventos de recombinación somática durante el desarrollo de linfocitos B. La recombinación puede generar una molécula con uno de cada uno de los segmentos V, D y J. En los seres humanos, hay 44 segmentos V, 27 D y 6 J; Por lo tanto, existe una posibilidad teórica de más de 7,000 combinaciones. En una pequeña fracción de BCR (alrededor del 5%) se encuentran dos segmentos D. Además, se pueden eliminar varias bases y agregar otras (llamadas nucleótidos N y P) en cada una de las dos uniones que generan un gran grado de diversidad. Después de la activación de las células B se produce un proceso de maduración por afinidad a través de la hipermutación somática. En este proceso, las células de la progenie de las células B activadas acumulan distintas mutaciones somáticas en todo el gen con una mayor concentración de mutación en las regiones CDR, lo que lleva a generar anticuerpos con mayor afinidad por los antígenos. Además de la hipermutación somática, las células B se someten al proceso de cambio de isotipo. Los anticuerpos con los mismos segmentos variables pueden tener diferentes formas (isotipos) dependiendo del segmento constante. Mientras que todas las células B ingenuas expresan IgM (o IgD), las células B activadas expresan principalmente IgG, pero también IgM, IgA e IgE. Esta expresión que cambia de IgM (y/o IgD) a IgG, IgA o IgE se produce a través de un evento de recombinación que hace que una célula se especialice en la producción de un isotipo específico. Hay un segmento para cada IgM, IgD e IgE, dos segmentos para IgA y cuatro segmentos para IgG.

## Métodos

Aplicación a la atención sanitaria y usos biotecnológicos.

55 El uso de las composiciones y métodos descritos en el presente documento para identificar anticuerpos y TCR y para agrupar secuencias de anticuerpos y TCR en familias clonales tiene muchas aplicaciones nuevas y útiles para la investigación en atención de la salud y biotecnología. Las familias clonales de anticuerpos pueden comprender clones no idénticos madurados por afinidad y las familias clonales de TCR pueden comprender clones idénticos. Estas aplicaciones incluyen, pero no se limitan a: 1) el descubrimiento y desarrollo de anticuerpos o terapias derivadas de anticuerpos; 2) el descubrimiento y desarrollo de diagnósticos; 3) el descubrimiento y desarrollo de herramientas de

investigación útiles en investigación en salud y biotecnología; y 4) el desarrollo y la evaluación de vacunas candidatas y la identificación de antígenos útiles como componentes de vacunas.

5 Como la invención se puede utilizar con cualquier tipo de célula B o T, la fuente de la célula y el subtipo(s) específico(s) de célula B o T se eligen en función del perfil del producto final deseado. Los ejemplos de subclases específicas de células B o T y su uso se describen en la subsección "Aislamiento y enriquecimiento de células y subpoblaciones de células" en la sección de Materiales y Métodos Generales. En general, las células pueden ser de un sujeto humano o animal en particular que tenga un estado clínico o curso de enfermedad particular, o que haya recibido un régimen de tratamiento particular, o que haya estado expuesta a un desafío particular, inmunización o conjunto de condiciones que induzcan una respuesta inmune.

10 Aplicación al descubrimiento y desarrollo de herramientas terapéuticas, diagnósticas y de investigación.

15 Para desarrollar un anticuerpo o molécula derivada de un anticuerpo para su uso como herramienta terapéutica, de diagnóstico o de investigación, el anticuerpo y/o los derivados de las regiones de unión a antígeno del anticuerpo pueden identificarse o descubrirse primero como que se unen a/el antígeno(s) deseado(s) o epítipo(s) y/o que tienen una consecuencia funcional deseada en un sistema in vivo o in vitro. A continuación, estos anticuerpos candidatos se analizan en busca de otras propiedades deseadas específicas para el producto deseado. Estas propiedades del producto objetivo serán diferentes para los diferentes tipos de anticuerpos terapéuticos, diagnósticos y de investigación, y la invención proporciona un medio útil para identificar candidatos para un mayor desarrollo hacia cualquiera de estas rutas de productos.

20 Según el perfil deseado de las propiedades del producto final, la fuente de células B relevantes puede ser, entre otras, un paciente con una enfermedad, como una enfermedad infecciosa, un cáncer o una afección autoinmune; un paciente que recibe un tratamiento, como terapia contra el cáncer o una vacuna; o un animal con enfermedad o tratado de manera que induzca una respuesta inmune, como la inmunización o la inducción/establecimiento de un modelo de enfermedad.

25 En general, los anticuerpos candidatos, o las macromoléculas candidatas derivadas de las regiones de unión a antígeno, que están destinadas al desarrollo como agentes terapéuticos, diagnósticos o de investigación se descubren a través de múltiples tecnologías que forman parte de uno de los dos enfoques generales: 1) aislamiento de anticuerpos de interés de las células B de la respuesta inmune de un humano o un animal; y 2) aislamiento de anticuerpos derivados de bibliotecas de expresión de moléculas de inmunoglobulina, o derivados de las mismas, expresados de forma heteróloga y cribados utilizando una o más tecnologías de visualización (revisado en  
30 Hoogenboom HR, Trends Biotechnol., 1997,15:62-70; Hammond PW, MAbs, 2010, 2:157-64; Nissim A, Chemajovsky Y, Handb Exp Pharmacol., 2008, (181):3-18; Steinitz M, Hum Antibodies, 2009;18:1-10; Bradbury AR, Sidhu S, Dübel S, and McCafferty, Nat Biotechnol., 2011, 29:245-54; Antibody Engineering (Kontermann RE and Dübel S eds., Springer, 2nd edition)).

35 Para la metodología anterior (#1), los anticuerpos candidatos se seleccionan de familias clonales particulares identificadas de donantes relevantes como se describe en, por ejemplo, la sección de Materiales y métodos generales. La invención se puede aplicar como se describe a las células B apropiadas (por ejemplo, células B blásticas) del donante humano o animal apropiado para descubrir o identificar anticuerpos candidatos. Por ejemplo, para un candidato a anticuerpo terapéutico contra el cáncer, el donante humano apropiado puede ser un paciente que haya suprimido con éxito la progresión del cáncer a través de una respuesta inmune; o para un candidato a anticuerpo de diagnóstico particular, el donante apropiado puede ser un paciente que tenga autoanticuerpos contra el marcador de diagnóstico o un ratón inmunizado contra el marcador; o para una herramienta de reactivo de anticuerpos candidato, el donante apropiado puede ser un ratón, un conejo, una cabra, una rata, un caballo, un pollo, un perro u otro animal  
40 inmunizado con la molécula y/o epítipo objetivo que el anticuerpo reactivo está destinado a reconocer. Las secuencias y la selección de anticuerpos para la expresión y la prueba se pueden realizar como se describe en la sección de Materiales y Métodos Generales. Dichas aplicaciones de la tecnología pueden proporcionar anticuerpos candidatos que a menudo se obtienen a través de métodos más laboriosos y que requieren más tiempo (por ejemplo, tecnología de hibridomas, immortalización inducida por virus de células B, etc.).

45 Para la última metodología (# 2), un subconjunto de, o el conjunto completo de secuencias de cadenas pesadas y livianas emparejadas de uno o más repertorios de anticuerpos humanos o animales, obtenidos como en # 1, se utilizan para sembrar bibliotecas de expresión que contienen regiones de identificación para rastrear el origen de la muestra y los pares afines originales de la muestra cuando se secuencian una biblioteca y/o un subconjunto seleccionado/  
50 enriquecido de una biblioteca utilizando una plataforma de secuenciación de la próxima generación. Las regiones variables y la información de la región marco pueden incorporarse en uno o más formatos de biblioteca de visualización de anticuerpos para descubrir anticuerpos candidatos. Las regiones variables de los genes de Ig pueden clonarse e incorporarse en vectores de expresión utilizando los métodos descritos en la subsección "Clonación y expresión de pares de inmunoglobulina de cadena pesada y cadena liviana" en la sección de Materiales y Métodos generales. Por ejemplo, los fragmentos y/o dominios de las cadenas pesada y liviana de pares afines obtenidos como en el # 1 se pueden usar para sembrar levadura Fab (Weaver-Feldhaus JM, Lou J, Coleman JR, y otros, FEBS Lett, 2004, 564:24-34) o bibliotecas de fagémidos (Kashyap AK, Steel J, Oner AF, y otros, Proc Natl Acad Sci USA, 2008, 105:5986-91)

con seguimiento de la región de identificación de cada cadena a la célula B original apropiada de origen independientemente de la combinación combinatoria de diferentes cadenas pesadas y livianas en emparejamientos no endógenos (no afines). Las cadenas pesada y liviana de pares afines obtenidas como en # 1 también se pueden usar con otras plataformas de visualización, más allá del fagémido o la levadura, y se pueden usar con otros constructos de expresión de derivados de anticuerpos más allá de los constructos de expresión del fragmento Fab [Antibody Engineering (Kontermann, RE and Dübel S eds., Springer, 2nd edition)]. En una aplicación alternativa de las regiones de identificación, las regiones de identificación se pueden agregar a las bibliotecas de visualización ya existentes para proporcionar los beneficios del seguimiento de la región de identificación y la corrección de errores de los datos de secuenciación de la próxima generación. Dependiendo del tipo de biblioteca, el formato y el sistema de expresión/visualización, las regiones de identificación pueden incorporarse utilizando reacciones de PCR o transcriptasa reversa seguidas de reacciones de PCR (Consúltese, por ejemplo, la subsección "Secuenciación de genes de inmunoglobulina pareados de cadena liviana y cadena pesada de células B únicas" en la sección de Materiales y Métodos Generales).

Anticuerpos candidatos, ya sea de repertorios de células B (véase, por ejemplo, la sección de Materiales Generales y Métodos) o mostrar "repertorios" de biblioteca (Kashyap AK, Steel J, Oner AF, et al., Proc Natl Acad Sci U S A, 2008, 105:5986-91; Weaver-Feldhaus JM, Lou J, Coleman JR, et al., FEBS Lett, 2004, 564:24-34; Ravn U, Gueneau F, Baerlocher L, et al., Nucleic Acids Res, 2010, 38:e193; Antibody Engineering (Kontermann, RE and Dübel S eds., Springer, 2nd edition), se identifican expresando y probando el anticuerpo o las moléculas derivadas del anticuerpo, o bibliotecas de moléculas, en ensayos de unión contra el antígeno/objetivo deseado(s) y/o epitopo(s) o en ensayos para evaluar la consecuencia funcional en un entorno in vivo o in vitro (incluyendo muestras/preparaciones ex vivo). Informes publicados han descrito el uso de regiones de identificación para rastrear la fuente donante de anticuerpo secuencias obtenidas de un repertorio de células B para usar en una biblioteca de expresión (por ejemplo, Kashyap AK, Steel J, Oner AF, et al., Proc Natl Acad Sci U S A, 2008, 105:5986-91). La tecnología de la región de identificación descrita en el presente documento proporciona mejoras útiles en el uso de dicha región de identificación. La invención: 1) proporciona un medio para rastrear, no solo a cada donante, sino a las células B de cada donante para el emparejamiento afinado de cadenas pesadas y livianas; 2) proporciona un medio para indexar de nuevo a la muestra de células B original para recuperar más muestra para clonación y/o prueba; 3) proporciona un medio para rastrear el origen de cadenas pesadas y livianas a pesar de los pares combinatorios no afines dentro de la biblioteca de expresiones; 4) proporciona un medio para rastrear el origen de cadenas pesadas y livianas a través de rondas de selección-enriquecimiento (por ejemplo, al monitorizar la evolución de secuencias durante la selección de grupos in vitro, como en Ravn U, Gueneau F, Baerlocher L, et al., Nucleic Acids Res, 2010, 38:e193).

Se ha demostrado que la identificación de las secuencias de cadena pesada o liviana representadas con mayor frecuencia en un repertorio de respuesta inmune de células B, y la combinación de pares de cadenas pesada y liviana en función de la frecuencia de orden de las cadenas individuales, es una forma viable de identificar algunos anticuerpos candidatos. a pesar del hecho de que la información del par cognado no se conserva en los análisis de secuencia de la siguiente generación cuando se realiza de esta manera (Reddy ST, Ge X, Miklos AE, et al., Nat Biotechnol, 2010, 28:965-9). La invención también permite este tipo de metodología de análisis de frecuencia, pero puede proporcionar además un medio para usar la secuenciación de la próxima generación para evaluar la frecuencia de anticuerpos reales en el repertorio, no simplemente cadenas pesadas o livianas independientes aisladas.

Además, la invención proporciona al menos tres mejoras de utilidad significativa más allá del análisis de frecuencia: 1) porque se puede rastrear el emparejamiento análogo de cadenas pesadas y livianas, el descubrimiento de anticuerpos reales de la respuesta inmune y las familias clonales de anticuerpos reales producidas por el B las células en la respuesta inmune pueden identificarse (un clon involucra un par específico, afin de cadenas pesadas y livianas que evolucionaron conjuntamente desde los mismos progenitores celulares y la información sobre los emparejamientos naturales dentro del proceso de maduración de afinidad mejoraría los enfoques descritos en la literatura para analice las respuestas inmunitarias mediante la secuenciación de la próxima generación [por ejemplo Wu X, Yang ZY, Li Y, Hogerkorp CM, et al., Science, 2010, 329:856-61]); 2) las regiones de identificación proporcionan los medios para minimizar, o incluso eliminar, el efecto en los análisis de secuencia de los errores de secuenciación comunes a las plataformas de secuenciación de la próxima generación (Consúltese, por ejemplo, la subsección "Otras opciones de análisis de datos de secuenciación" en la sección de Materiales y Métodos Generales); y 3) las regiones de identificación proporcionan la capacidad de vincular y rastrear  $\geq 2$  secuencias coexpresadas a nivel de una sola célula.

Para aquellos anticuerpos candidatos que se han identificado que tienen propiedades de unión deseables a un antígeno, diana o epitopo, o que tienen un efecto funcional deseado, se pueden clonar y expresar más anticuerpos de la familia clonal respectiva (véase, por ejemplo, la sección de Materiales y Métodos Generales) para probar la presencia de anticuerpos o anticuerpos similares, pero potencialmente más óptimos, que tengan las mismas propiedades funcionales o de unión pero que contengan otra diferencia de importancia para el perfil del producto final.

Para los casos en los que los candidatos se identifican a partir de bibliotecas de expresión de visualización, las regiones de identificación pueden proporcionar un medio para identificar anticuerpos de secuencia potencialmente similar a los candidatos mediante la identificación de secuencias que no se seleccionaron en el enriquecimiento de selección pero que contienen regiones de identificación de los candidatos identificados y, por lo tanto, se derivan de

las mismas cadenas pesadas y/o livianas originales que sembraron la biblioteca. Los anticuerpos que se pierden en rondas de selección in vitro, pero son similares a los seleccionados, los anticuerpos candidatos pueden recuperarse o "rescatarse" para su posterior análisis como posibles candidatos. Dicho rescate puede obviar el efecto del sesgo en la expresión o los sistemas de análisis que pueden pasar por alto los anticuerpos útiles y funcionales (Ravn U, Gueneau F, Baerlocher L, et al., *Nucleic Acids Res*, 2010, 38:e193).

Una vez que se identifican los candidatos con propiedades de unión y/o funcionales deseadas, se pueden avanzar a ensayos relevantes y otras evaluaciones basadas en el perfil de producto deseado, en sentido descendente. Para los anticuerpos terapéuticos destinados a la inmunización pasiva, los candidatos son avanzados en ensayos y modelos de pruebas preclínicos para determinar los mejores candidatos para las pruebas clínicas en humanos o para el uso en salud animal, incluidas, entre otras, evaluaciones de propiedades como la estabilidad y la agregación. facilidad de formulación y dosificación, expresión y fabricación de proteínas, selectividad de especies, farmacología, farmacocinética, seguridad y toxicología, absorción, metabolismo y recambio de anticuerpos diana, así como inmunogenicidad [Ver, por ejemplo, Lynch CM, Hart BW, and Grewal IS, *MAbs*, 2009, 1: 2-11; Chapman K, Pullen N, Coney L, et al., *mAbs*, 2009, 1, 505-516; S. Dübel, *Handbook of Therapeutic Antibodies: Technologies, Emerging Developments and Approved Therapeutics* (John Wiley & Sons, 2010); *Therapeutic Monoclonal Antibodies: From Bench to Clinic*, (Z. An ed., John Wiley & Sons, 2009) *Antibody Engineering* (Kontermann RE and Dübel S eds., Springer, 2nd edition)]. Por lo tanto, se seleccionan muchos candidatos porque la mayoría será insuficiente para las pruebas terapéuticas en humanos con respecto a al menos una de las muchas propiedades que deben evaluarse antes de las pruebas en humanos (es decir, desgaste). Las familias clonales pueden ser minadas, por ejemplo, como se describió anteriormente, para los candidatos similares a los que ya están caracterizados, pero posiblemente albergan diferencias con respecto a una o más de las propiedades que se evalúan en el trabajo preclínico. Es posible que deban emplearse estrategias específicas de ingeniería de anticuerpos para optimizar ciertas propiedades [Antibody Engineering (Kontermann RE and Dübel S eds., Springer, 2nd edition)].

Para el diagnóstico, la invención se puede usar para identificar anticuerpos, TCR y familias clonales producidas por infección o vacunación para uso en la detección de agentes infecciosos (Selvarajah S, Chatterji U, Kuhn R, et al., 2012, 6: 29- 37; Berry JD, *Vet J*, 2005, 170: 193-211), así como para cualquier enfermedad no infecciosa, afección patológica, tratamiento médico o terapia. Dichos anticuerpos, TCR y/o familias clonales pueden proporcionar sondas de diagnóstico útiles para biomarcadores o proporcionar información del sistema inmunitario sobre el estado de la enfermedad o el efecto del tratamiento en un humano o animal. Como tales, los anticuerpos o TCR específicos, o las familias clonales específicas de cualquiera de las clases de receptores inmunes pueden proporcionar utilidad para las herramientas de diagnóstico y la medicina personalizada. En otra aplicación para diagnóstico, se pueden usar biomarcadores de enfermedad o respuesta al tratamiento como inmunógenos para inmunizar ratones u otros animales de los cuales se recolectan células B para identificar anticuerpos (véase, por ejemplo, la sección de Materiales Generales y Métodos) contra el biomarcador que posteriormente podría ser utilizado en pruebas de diagnóstico, como ELISA u otros inmunoensayos (Selvarajah S, Chatterji U, Kuhn R, et al., 2012, 6:29-37). Una vez que se identifique que tienen potencial utilidad de diagnóstico, los anticuerpos candidatos, TCR y/o familias clonales pueden avanzar a ensayos, modelos y posiblemente ensayos relevantes para el perfil deseado del producto de diagnóstico [Berry JD, *Vet J*, 2005, 170:193-211; *Diagnostic and Therapeutic Antibodies in Methods in Molecular Medicine*, Vol. 40 (George AJT and Urch CE eds., Humana Press); *Antibody Engineering* (Kontermann RE and Dübel S eds., Springer, 2nd edition) ; Colwill K, Renewable Protein Binder Working Group, y Gräslund S, *Nat Methods*, 2011, 8:551-8; Pershad K, Pavlovic JD, Gräslund S, et al., *Protein Eng Des Sel*, 2010, 23:279-88.]. Es posible que deban emplearse estrategias específicas de ingeniería de anticuerpos para optimizar ciertas propiedades [Antibody Engineering (Kontermann RE and Dübel S eds., Springer, 2nd edition)].

Para los anticuerpos de la herramienta de investigación, los candidatos identificados, por ejemplo, como se describió anteriormente, pueden ser avanzados para probar su desempeño en la aplicación de investigación para la cual está destinado el anticuerpo de la herramienta de investigación (por ejemplo, inmunoprecipitación, inmunotransferencia, inmunotinción e histología, purificación por inmutoaafinidad, captura, e inmunoensayos de detección y sándwich, por ejemplo, como se describe en *Antibodies: A Laboratory Manual*, E Harlow and D Lane (Cold Spring Harbor Laboratory Press, 1988). Los criterios de validación se basarán en el uso de la investigación final prevista (Colwill K, Renewable Protein Binder Working Group, and Gräslund S, *Nat Methods*, 2011, 8:551-8; Pershad K, Pavlovic JD, Gräslund S, et al., *Protein Eng Des Sel*, 2010, 23:279-88). Puede ser necesario emplear estrategias de ingeniería específica de anticuerpos para optimizar ciertas propiedades [Antibody Engineering (Kontermann RE and Dübel S eds., Springer, 2nd edition)].

Aplicación al descubrimiento y desarrollo de vacunas.

La invención se puede usar para identificar anticuerpos, TCR y familias clonales de cada una de estas clases de receptores inmunes para un desafío de vacuna en un humano o animal. Se pueden usar anticuerpos específicos como sondas para identificar los componentes de la vacuna reconocidos por el anticuerpo y la familia clonal a la que pertenece el anticuerpo que se usó como sonda. Esta información sobre el anticuerpo y las familias clonales se puede usar para hacer evaluaciones sobre las proporciones o la fuerza de la respuesta inmune que se dirige a antígenos y/o epítopos particulares de la vacuna. La evaluación de las respuestas inmunes de anticuerpos a diferentes componentes de la vacuna se puede complementar con la información recopilada sobre la respuesta del repertorio de TCR

- concomitante a la vacuna (véase, por ejemplo, las subsecciones, "Para otros tipos de células" y "PCR de otras cadenas pesadas de inmunoglobulina y T-cadenas de receptores celulares (TCR) "en la sección Materiales y Métodos Generales). Esta información puede usarse posteriormente para comprender qué componentes de la vacuna, o qué variantes de una vacuna, o qué adyuvantes producen respuestas efectivas o más óptimas de un sistema inmunológico de un ser humano o animal (Haynes BF, Gilbert PB, McElrath MJ, et al., N Engl J Med, 2012, 366:1275-86). LA metodología también se puede utilizar para comparar individuos o poblaciones en su respuesta a una vacuna.
- Se pueden realizar análisis similares para identificar y evaluar los anticuerpos, los TCR y las familias clonales producidas en respuesta a un patógeno real y que pueden correlacionarse con resultados clínicos de interés, como las respuestas protectoras a la infección (por ejemplo, Identification of antibodies from survivors of a severe influenza pandemic, Yu X, Tsibane T, McGraw PA, et al., Nature, 2008, 455:532-6; o Identification of specific antibodies from HIV-infected individuals with broadly HIV-neutralizing sera, Wu X, Yang ZY, Li Y, Hogerkorp CM, et al., Science, 2010, 329:856-61 and Walker LM, Phogat SK, Chan-Hui PY, et al., Science, 2009, 326:285-9). La identificación de dichos correlatos de protección se puede comparar con la respuesta producida por la vacuna y/o los componentes específicos de la vacuna como se describe anteriormente y los dos conjuntos de datos se pueden comparar para evaluar la capacidad de la vacuna para producir respuestas inmunitarias que se correlacionen con los resultados deseados observados en casos de infección real.
- Por lo tanto, la invención proporciona un medio útil para obtener una lectura sustituta de la protección de la enfermedad y la respuesta de la vacuna, mediante el análisis de la secuencia del anticuerpo y/o del repertorio del TCR, antes de que un humano o animal se enfrente a una infección real. Una vez que una familia clonal ha sido identificada como que se une a un antígeno o epítipo en particular, la identificación de antígenos o epítipos a los que apuntan otros repertorios de respuesta inmunitaria es posible sin realizar ensayos en los casos en que se encuentran familias clonales iguales o similares en todos los repertorios. Por lo tanto, en aquellos casos en que se conoce suficiente información sobre una familia clonal y su unión antígeno/epítipo (véase, por ejemplo, la subsección, "Detección de anticuerpos humanos expresados" en la sección de Materiales y Métodos Generales), análisis de secuencia solo de repertorios recientemente analizados puede proporcionar una lectura de los antígenos de ese repertorio para las familias clonales conocidas que contiene. Esta aplicación puede proporcionar un medio útil para monitorizar las respuestas en uno, unos pocos o muchos sujetos en ensayos clínicos de vacunas y para monitorizar la inmunidad y las relaciones de enfermedades infecciosas para una, unas pocas o muchas personas a nivel de la población.
- Además, las familias de anticuerpos, TCR y clonales que se correlacionan con la protección de un patógeno se pueden usar para identificar los antígenos específicos y los conjuntos de antígenos (incluidos los antígenos conocidos y nuevos) que median las respuestas inmunitarias protectoras y/o efectivas contra el patógeno. La identificación de los antígenos dirigidos a las respuestas inmunitarias efectivas se puede usar para guiar la selección de antígenos que se incluirán en las vacunas que se espera que produzcan respuestas protectoras mediadas por anticuerpos y TCR en humanos o animales inmunizados.
- Los anticuerpos, TCR y familias clonales que no se unen a antígenos conocidos en ensayos se convierten en candidatos para la identificación de antígenos y/o epítipos potencialmente nuevos del patógeno contra el cual los anticuerpos y/o TCR proporcionan protección. Los anticuerpos que se sabe que no se unen a antígenos ya conocidos pueden usarse como sondas en combinación con inmunoseparación y espectroscopia de masas para identificar el antígeno o epítipo no identificado previamente (Zhu YZ, Cai CS, Zhang W, et al., PLoS One, 2010, 5: e13915 y ver, por ejemplo, las subsecciones, "Inmunoprecipitación de antígenos de estafilococos con anticuerpos derivados de pacientes infectados con estafilococos" y "Identificación por espectrometría de masas de péptidos" en la sección Materiales y Métodos generales). Tales nuevos antígenos o epítipos se pueden usar como componentes de vacunas que se espera que produzcan o contribuyan a la producción de respuestas protectoras mediadas por anticuerpos- y TCR -en humanos o animales inmunizados.
- Además de facilitar el desarrollo de vacunas para patógenos microbianos, el anticuerpo, TCR y familias clonales también pueden usarse para desarrollar vacunas contra tumores. Los seres humanos o animales que desarrollan una respuesta inmunitaria contra un cáncer o células precancerosas pueden producir anticuerpos, TCR y familias clonales que pueden usarse para identificar individuos y combinaciones de antígenos que pueden incorporarse en vacunas preventivas o terapéuticas para el cáncer.
- Métodos para producir uno o más polinucleótidos de interés.
- En algunos aspectos, un método incluye obtener una biblioteca de ADNc que comprende una pluralidad de ADNc asociados con una pluralidad de muestras obtenidas de uno o más sujetos, en donde cada ADNc está asociado con una sola muestra en la pluralidad de muestras, y en donde cada ADNc asociado con cada muestra está presente en un recipiente separado; agregar una molécula adaptadora al ADNc asociado con cada muestra, en donde la molécula adaptadora comprende una región de identificación de muestra y una región adaptadora, en donde la región de identificación de muestra está acoplada a la región adaptadora, y en donde la secuencia de la región de identificación de muestra de cada adaptador la molécula es distinta de la secuencia de la región de identificación de la muestra de las otras moléculas adaptadoras agregadas a cada ADNc en la biblioteca; y permitir que la región del adaptador se una a cada ADNc en la biblioteca para producir uno o más polinucleótidos de interés.

En algunos aspectos, obtener la biblioteca de ADNc comprende obtener la pluralidad de muestras y procesar las muestras para preparar la biblioteca de ADNc. En algunos aspectos, obtener la biblioteca de ADNc comprende recibir la biblioteca de ADNc directa o indirectamente de un tercero que ha procesado la pluralidad de muestras para preparar la biblioteca de ADNc.

5 En algunos aspectos, la molécula adaptadora comprende además una región de cebador universal, en donde el extremo 3' de la región de cebador universal está acoplada al extremo 5' de la región de identificación de la muestra. En algunos aspectos, cada región de ADNc comprende un polinucleótido de ARNm hibridado con un polinucleótido de ADNc.

10 En algunos aspectos, cada muestra comprende una célula. En algunos aspectos, la célula es una célula B. En algunos aspectos, la célula B es un plasmoblasto, una célula B de memoria o una célula plasmática. En algunos aspectos, cada muestra comprende una pluralidad de células.

En algunos aspectos, cada región del adaptador se une a cada ADNc a través de la unión, por ejemplo, la unión de G:C.

15 En algunos aspectos, la molécula adaptadora es de cadena sencilla, y comprende además incorporar la molécula adaptadora en cada ADNc permitiendo que una enzima haga que la molécula adaptadora sea de cadena doble. En algunos aspectos, la molécula adaptadora se incorpora a cada ADNc para producir el polinucleótido de interés mediante una transcriptasa reversa MMLV H<sup>-</sup>. En algunos aspectos, los métodos pueden incluir pasos de amplificación tales como PCR y otras reacciones de amplificación generalmente conocidas en la técnica.

Métodos de enlace y polinucleótidos de código de barras de interés.

20 En algunos aspectos, el método incluye la unión de dos secuencias polinucleotídicas de interés, por ejemplo, una cadena liviana (LC) y una cadena pesada (HC) de anticuerpo de una sola muestra, y proporcionar una o más secuencias de identificación de secuencia o código de barras. En este aspecto, se proporciona un enlace físico entre las secuencias polinucleotídicas de interés, así como uno o más códigos de barras para proporcionar un identificador para permitir que se determinen las secuencias polinucleotídicas derivadas de una fuente o muestra particular, por ejemplo, célula única, pozo de muestra, muestra única, etc. Las muestras individuales pueden comprender una o más células de linaje B u otros tipos de células. Ejemplos de métodos para unir dos secuencias polinucleotídicas de interés son conocidos en la técnica, por ejemplo, los documentos WO 99/16904, WO 93/03151 y US 7,749,697, que se incorporan aquí como referencia. Entre otras ventajas asociadas con el uso de códigos de barras en secuencias de polinucleótidos enlazadas se incluye la facilitación de la secuenciación de alto rendimiento y el mapeo de una secuencia de nuevo a una muestra original, de modo que pueda volver a secuenciarse y clonarse por PCR para expresar las secuencias de polinucleótidos, por ejemplo, HC y polinucleótidos de inmunoglobulina LC. Algunas de las tecnologías de secuenciación de alto rendimiento muestran tasas de error de secuenciación de 1-10 +%, y el uso de códigos de barras permite la secuenciación repetida de plantillas para facilitar la corrección de errores bioinformáticos. Esto es particularmente importante para distinguir los errores de secuenciación de las variaciones genéticas, como las de los polinucleótidos de inmunoglobulina. Específicamente, puede ser difícil determinar si las secuencias estrechamente relacionadas son de hecho secuencias distintas o si en su lugar representan artefactos producidos por errores de secuenciación. Los códigos de barras, al permitir el análisis de la secuenciación repetida de plantillas individuales, permiten la corrección de errores de secuenciación, proporcionando así la determinación de si las secuencias son distintas frente a los artefactos de los errores de secuenciación. En una realización, las secuencias de polinucleótidos son inmunoglobulina HC y secuencias de LC que se han desviado debido a la hipermutación somática, y difieren en solo 1 nucleótido.

En este aspecto, las estructuras físicamente unidas y con código de barras, como se muestra en la Figura 15, generalmente se obtienen. La Figura 15 ilustra el enlace físico de dos segmentos de ácido nucleico, A y B (por ejemplo, dos ADNc). Un código de barras (BC) se adjunta a cualquiera de los extremos o en el enlazador que conecta A y B. Este enlace físico de A y B, así como también, la adición del código de barras se realiza a través de una serie de medios, incluso mediante ligación, recombinación, amplificación o superposición-extensión, o una combinación de estos métodos, como se describe con mayor detalle a continuación. Opcionalmente, se pueden agregar códigos de barras adicionales a la estructura que se muestra en la Figura 15, para proporcionar códigos de barras compuestos que permitan la secuenciación de un gran número de polinucleótidos enlazados utilizando un número menor de códigos de barras. Además, se apreciará que, dependiendo de la estrategia particular utilizada para unir los dos segmentos de ácido nucleico, se puede obtener cualquier orientación relativa de los segmentos, con respecto a las orientaciones sentido y antisentido, es decir, los segmentos, como los ADNc, pueden ser unidos de cabeza a cola, de cabeza a cabeza o de cola a cola.

55 Los códigos de barras se pueden agregar a las secuencias de polinucleótidos antes, durante o después del enlace físico utilizando métodos conocidos en la técnica. Estos métodos incluyen, pero no se limitan a, por ejemplo, métodos de ligación, como la ligación de los extremos romos de los adaptadores de códigos de barras, y mediante el fusiónamiento y la ligación de las terminaciones compatibles, como las generadas por la cola homopolimérica, la digestión con enzimas de restricción de un enlazador, o 3' cola de ADNc por transcriptasa reversa. Los códigos de



barras también se pueden agregar en reacciones de amplificación utilizando cebadores adecuados que llevan secuencias de códigos de barras. Los códigos de barras también se pueden agregar en una reacción de transcripción reversa utilizando cebadores adecuados que contienen la secuencia de códigos de barras. Los códigos de barras también se pueden agregar mediante la incorporación en oligonucleótidos utilizados para vincular los genes de interés entre sí, a través de colas de superposición-extensión u otros métodos, de manera que estén ubicados entre los dos genes de interés. Por consiguiente, utilizando estos métodos, los códigos de barras pueden incorporarse en los extremos de secuencias polinucleotídicas unidas físicamente o en las secuencias enlazadoras que unen las dos secuencias polinucleotídicas de interés.

En una realización, el enlace se logra mediante el uso de la superposición-extensión (véase Figura 16). En general, las colas de extensión solapadas son secuencias complementarias que se agregan a las secuencias de polinucleótidos de interés para unirse. El fusionamiento de las colas de superposición y extensión anexadas a las moléculas polinucleotídicas de interés permite que se unan (véase Figuras 17, 18, 19 y 20). Como se describe a continuación, las colas de superposición y extensión se pueden agregar a través de una serie de métodos bien conocidos que incluyen, entre otros, reacciones de síntesis de polinucleótidos, como la amplificación de ácidos nucleicos y la transcripción reversa, la ligación y la recombinación. Debido a la variedad de métodos disponibles para efectuar el enlace, se reconocerá que pueden obtenerse diferentes orientaciones relativas de los segmentos de polinucleótidos, con respecto a las orientaciones de sentido y antisentido, es decir, los segmentos, por ejemplo, cadenas pesadas y livianas de anticuerpos, pueden ser unidos de cabeza a cola, de cabeza a cabeza o de cola a cola.

En una realización, las colas de superposición-extensión permiten el enlace de secuencias de polinucleótidos generadas durante una reacción de síntesis de polinucleótidos. Por ejemplo, las colas de extensión solapadas se pueden introducir durante el curso de las reacciones de síntesis de polinucleótidos, como las reacciones de amplificación o transcripción reversa, mediante el uso de cebadores que llevan una cola de superposición-extensión. Alternativamente, pueden usarse reacciones de ligación. Como se muestra en las Figuras 17, 18, 19 y 20, después del fusionamiento de colas de extensión de solapamiento complementarias, el ADN se rellena en una dirección de 5' a 3' durante la fase de extensión de una reacción de síntesis de polinucleótidos, como la PCR, para generar un polinucleótido de doble cadena con los dos polinucleótidos de interés unidos físicamente.

En algunas realizaciones, un método de RT-PCR de superposición-extensión permite que las secuencias se unan simultáneamente a medida que la reacción avanza en un solo tubo, eliminando así la necesidad de una purificación intermedia. En algunas realizaciones, una cola de extensión solapada comprende una secuencia de código de barras.

La Figura 17 ilustra en general un ejemplo del uso de colas de extensión solapada para unir secuencias de polinucleótidos que codifican cadenas livianas y pesadas de anticuerpos y para proporcionar al menos un código de barras. Otros métodos útiles para unir dos secuencias de polinucleótidos de interés se discuten a continuación. En este ejemplo, después de que haya ocurrido la síntesis de polinucleótidos, por ejemplo, la transcripción reversa, el uso de un cebador de PCR específico del gen LC que contiene un código de barras, un sitio de cebador de secuenciación opcional y un sitio de restricción opcional (RE1) permite que estos elementos se agreguen al final del producto de PCR resultante. Se indican los cebadores específicos para LC (en una realización la región  $V_L$ ) y HC (en una realización la región  $V_H$ ) con superposiciones de extensión y que codifican un sitio de restricción opcional (RE3). En una realización, la LC comprende la VJ reorganizada con o sin un segmento corto de la región constante de la cadena pesada. En una realización, los cebadores de superposición y extensión también contienen una secuencia de código de barras. También se usa un cebador reverso específico para HC que contiene un RE2 opcional. A medida que avanza la amplificación con estos cebadores, se genera un ácido nucleico con la estructura enlazada mostrada con un código de barras en un extremo. Los productos de las reacciones realizadas en muestras individuales pueden integrarse fácilmente en los otros flujos de trabajo descritos en este documento. Por ejemplo, un segundo código de barras opcional se puede agregar y usar junto con el primer código de barras para permitir que la multiplexación identifique grandes cantidades de secuencias usando un número mínimo relativo de códigos de barras. Para la secuenciación, un solo código de barras puede ser suficiente.

Las variaciones del esquema general mostrado en la Figura 17, cuyos ejemplos se ilustran aquí, serán evidentes. Por ejemplo, el código de barras se puede colocar en el otro extremo del producto final, o en ambos extremos, o entre los polinucleótidos. Además, se puede incluir un código de barras como parte de la región de superposición de la extensión (por ejemplo, a cada lado de RE3 o el código de barras se puede dividir por la secuencia de RE3).

Las secuencias LC (incluidas las secuencias  $V_L$ ) y HC (incluidas las secuencias  $V_H$ ) pueden derivarse a través de una variedad de medios. Por ejemplo, pueden generarse a través de la transcripción reversa de ARNm con oligo dT o cebadores específicos de genes. En una realización, una transcripción reversa y las reacciones de amplificación subsiguientes se realizan simultáneamente, es decir, una reacción de RT-PCR, para llegar al producto final. Cuando se utiliza la transcripción reversa, la región de superposición de la extensión, así como otros elementos, como los sitios de restricción, los sitios de cebadores de secuenciación, o las secuencias universales, se pueden agregar a través del fusionamiento de un adaptador que comprende uno o más residuos G a uno, dos, tres o más residuos de C generados por la cola 3' del ADNc generado en la reacción de transcripción reversa como se muestra, por ejemplo, en las Figuras 18, 19 y 20. El cambio de plantilla por la transcriptasa reversa permite una región de superposición de

extensión (y otros elementos de secuencia) que se añadirán al ADNc. Por ejemplo, como se muestra en la Figura 18, cuando se aprovechan las actividades de cambio de cola y plantilla en 3' de la transcriptasa reversa, se puede usar un primer adaptador para agregar una secuencia de superposición de extensión y un código de barras a un primer polinucleótido de interés, mientras que se puede agregar un segundo adaptador con una secuencia complementaria a la superposición-extensión del primer adaptador a un segundo ADNc de interés. Las secuencias de solapamiento de la extensión complementaria se reasocian durante una reacción de síntesis de ácido nucleico posterior, como la PCR, para unir los dos polinucleótidos de interés. La extensión desde el punto de superposición da como resultado una molécula de ADN de doble cadena en donde los dos polinucleótidos de interés están unidos con un código de barras entre ellos. En las Figuras 19 y 20 se muestran las variaciones que permiten la generación de dos códigos de barras ubicados internamente entre dos secuencias de polinucleótidos enlazadas.

Otros métodos para unir o unir las secuencias polinucleotídicas de interés incluyen la ligación. En esta realización, la mezcla de cebadores utilizada para la amplificación se diseña de tal manera que las secuencias diana amplificadas se pueden escindir con enzimas de restricción apropiadas, y se puede realizar un enlace covalente mediante la ligación del ADN. Después de la amplificación con dicha mezcla de cebadores, las enzimas de restricción necesarias para formar extremos compatibles de las secuencias diana, se agregan a la mezcla. Las secuencias diana se ligan luego con una ligasa. No es necesaria la purificación de los productos de la PCR antes de la digestión con enzimas de restricción o las etapas de ligación, aunque se puede realizar la purificación.

En otra realización, las secuencias polinucleotídicas de interés pueden unirse por recombinación. En esta metodología, las secuencias polinucleotídicas de interés amplificadas se pueden unir utilizando sitios de recombinación idénticos. El enlace se realiza agregando la recombinasa apropiada para facilitar la recombinación. Los sistemas de recombinasa adecuados incluyen la recombinasa F1p con una variedad de sitios FRT, la recombinasa Cre con una variedad de sitios lox, la integrasa  $\Phi$ C31 que lleva a cabo la recombinación entre el sitio attP y el sitio attB, el sistema  $\beta$ -recombinase-six así como la Gin sistema gix. La unión por recombinación se ha ejemplificado para dos secuencias de nucleótidos ( $V_H$  unidas a  $V_L$ ) (Chapal, N. et al. 1997 *BioTechniques* 23, 518-524), incorporadas aquí como referencia.

Por consiguiente, en un aspecto, el método comprende amplificar por PCR o RT-PCR amplificación, secuencias de nucleótidos de interés usando una plantilla derivada de una célula aislada o una población de células isogénicas y (1) efectuar el enlace de las secuencias de nucleótidos amplificadas de interés y (2) agregar uno más de los códigos de barras a las secuencias de polinucleótidos enlazadas. El método comprende una etapa opcional de realizar una amplificación adicional de los productos enlazados para, por ejemplo, agregar códigos de barras adicionales, sitios de restricción, sitios de cebadores de secuenciación, y similares.

En otro aspecto, se proporciona un método para producir una biblioteca de pares enlazados con código de barras que comprenden cadenas pesadas y livianas de anticuerpos de células individuales de un donante. Este aspecto comprende proporcionar una fracción celular que contiene linfocitos de un donante, que se enriquece opcionalmente para una población de linfocitos particular de dicha fracción celular. Además, se obtiene una población de células individuales aisladas distribuyendo células de la fracción de células que contienen linfocitos, o la fracción de células enriquecidas, individualmente entre una pluralidad de vasos, recipientes o pozos. La amplificación molecular múltiple (por ejemplo, la amplificación por RT-PCR múltiple) de las secuencias codificantes de la región variable contenida en la población de células individuales aisladas se lleva a cabo y el enlace de los pares de cadenas pesadas y livianas y la adición de códigos de barras se ve afectada, en donde un par individual se deriva de una sola célula. Además, en diferentes realizaciones, el método puede comprender dos pasos opcionales: en el primer paso, la célula individual aislada en la población de células individuales puede expandirse a una población de células isogénicas antes de realizar la amplificación RT-PCR multiplex, proporcionando así una pluralidad de vasos, contenedores o pozos con una población de células isogénicas (una población de células isogénicas en un vaso, contenedor o pozo). Otro paso opcional incluye realizar una amplificación adicional de las secuencias de codificación ligadas y de las cadenas pesadas enlazadas. Esta etapa de amplificación adicional se puede usar para simplemente aumentar la cantidad del ácido nucleico enlazado, o para agregar una primera o segunda secuencia de código de barras u otros elementos de secuencia al ácido nucleico enlazado.

En algunos aspectos, la amplificación RT-PCR multiplex puede realizarse como un proceso de dos pasos, donde la transcripción reversa (RT) se realiza por separado de la amplificación PCR multiplex (o amplificación molecular multiplex alternativa), o como un proceso de un solo paso, donde los pasos de amplificación de PCR y RT multiplex se realizan con los mismos cebadores en un tubo.

La transcripción reversa (RT) se realiza con una enzima que contiene actividad de transcriptasa reversa que resulta en la generación de ADNc a partir de ARN total, ARNm o ARN específico diana de una célula aislada. Los cebadores que pueden utilizarse para la transcripción reversa incluyen cebadores oligo-dT, hexámeros aleatorios, descamadores aleatorios, otros cebadores aleatorios, o cebadores que son específicos para las secuencias de nucleótidos de interés. En algunas realizaciones, dichos cebadores pueden contener elementos tales como códigos de barras, sitios de cebado universales, sitios de restricción, sitios de cebadores de secuenciación, y similares.

El procedimiento de amplificación de RT-PCR multiplex de dos pasos permite que el ADNc generado en el paso de RT se distribuya a más de un recipiente, lo que permite el almacenamiento de una fracción de plantilla antes de

continuar con la amplificación, si se desea. Además, la distribución de ADNc a más de un tubo permite el rendimiento de más de una amplificación de PCR multiplex de ácido nucleico derivado de la misma plantilla. Esta metodología de dos pasos se puede utilizar, por ejemplo, para amplificar y vincular las secuencias codificantes de región variable de cadena pesada y de cadena variable kappa en un tubo y la región variable de cadena pesada y las secuencias codificantes de región variable de cadena liviana lambda en un tubo diferente utilizando la misma plantilla. Una sola célula generalmente solo expresa una de las cadenas livianas. Sin embargo, a menudo será más fácil realizar las reacciones simultáneamente en lugar de esperar el resultado de una de las reacciones antes de realizar la otra. Además, la amplificación de kappa y lambda sirve como un control interno negativo, ya que se esperaría que solo kappa o lambda se amplificaran desde una sola célula.

- 5
- 10 En el procedimiento de RT-PCR multiplex de un solo paso, la transcripción reversa y la amplificación por PCR múltiple se llevan a cabo en el mismo recipiente, recipiente o pozo. Todos los componentes necesarios para realizar tanto la transcripción reversa como la PCR multiplex en un solo paso se agregan inicialmente a los recipientes, contenedores o pozos y se realiza la reacción. En general, no es necesario agregar componentes adicionales una vez que se haya iniciado la reacción. La ventaja de la amplificación de RT-PCR multiplex de un solo paso es que reduce el número de pasos necesarios para generar las secuencias de nucleótidos ligadas a códigos de barras de la presente invención aún más. Esto es particularmente útil cuando se realiza RT-PCR multiplex en una matriz de células individuales, donde la misma reacción se lleva a cabo en una pluralidad de vasos. En general, la composición necesaria para la RT-PCR multiplex de un solo paso comprende una plantilla de ácido nucleico, una enzima con actividad de transcriptasa reversa, una enzima con actividad de ADN polimerasa, una mezcla de desoxinucleósido trifosfato (mezcla de dNTP que comprende dATP, dCTP, dGTP y dTTP) y una mezcla de cebador multiplex. La plantilla de ácido nucleico es preferiblemente ARN total o ARNm derivado de una célula aislada en forma purificada, como un lisado de la célula, o como está contenido en la célula intacta.
- 15
- 20

- En un aspecto, los métodos generan bibliotecas de polinucleótidos de interés unidos y con código de barras. En algunos aspectos, la pluralidad de composiciones polinucleotídicas en una biblioteca de polinucleótidos puede comprender al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 3000, al menos 10 000, al menos 30 000, al menos 100 000, al menos 300 000, al menos 1 000 000, al menos 3 000 000, al menos 10 000 000, al menos 30 000 000, o más miembros. En otros aspectos, la pluralidad de composiciones polinucleotídicas en una biblioteca de polinucleótidos puede comprender al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 3000, al menos 10 000, al menos 30 000 o más genes del transcriptoma completo de una muestra de células. En otros aspectos, la pluralidad de composiciones polinucleotídicas en una biblioteca de polinucleótidos comprende al menos 1, al menos 2, al menos 3, al menos 10, al menos 30, al menos 100, al menos 300, al menos 1000, al menos 10 000, al menos 100 000, al menos 1 000 000, al menos 10 000 000, al menos 1 000 000 000 o más de las diferentes especies de anticuerpos presentes en la sangre de un individuo. Estas especies de anticuerpos pueden expresarse mediante plasmoblastos, células plasmáticas, células B de memoria, células plasmáticas de larga vida, células B ingenuas, otras células de linaje B o combinaciones de las mismas.
- 25
- 30
- 35

Las composiciones de polinucleótidos enlazadas y con código de barras generadas por los métodos descritos anteriormente pueden someterse ventajosamente a una secuenciación alta y multiplexada, preferiblemente, utilizando plataformas de secuenciación NextGen como se describe en el presente documento.

- 40 Las composiciones de polinucleótidos enlazadas y con código de barras generadas por los métodos descritos anteriormente también se pueden usar para clonar, producir polipéptidos de interés y seleccionar como se describe en este documento.

Métodos para producir uno o más polinucleótidos de interés para la secuenciación.

- En algunos aspectos, el método incluye obtener una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una región de cebador universal, una región de identificación de muestra, una región de adaptador y una región de amplicón derivada de una sola muestra, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido derivada de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca derivado de una o más muestras distintas de la primera muestra única; y amplificar la biblioteca de polinucleótidos con un conjunto de cebadores para producir uno o más polinucleótidos de interés para la secuenciación, en donde el uno o más polinucleótidos de interés para la secuenciación comprende una primera región de secuenciación, una región de identificación de la primera placa, una región de cebador universal, una región de identificación de la muestra, una región del adaptador, una región del amplicón derivada de una sola muestra y una segunda región de secuenciación.
- 45
- 50
- 55

En algunos aspectos, un método incluye además la secuenciación de uno o más polinucleótidos de interés. En algunos aspectos, la secuenciación es de secuenciación 454.

En algunos aspectos, la secuenciación incluye lecturas de secuenciación más largas, de modo que las lecturas de secuencia directa e inversa se solapan lo suficiente como para permitir la reconstrucción de la secuencia completa, aproximadamente 600 pares de bases (pb) de, por ejemplo, cadenas livianas de anticuerpos (LC) (donde la longitud exacta de la secuencia puede depender de la longitud de la región 5' no traducida (UTR), y de aproximadamente 700 pb de la secuencia de las cadenas pesadas (HC). Por lo tanto, en algunos aspectos, se puede usar cualquier tecnología de secuenciación que pueda producir lecturas de secuenciación de al menos 350-400pb y, de ese modo, lograr la superposición incluida para el ensamblaje de secuencia, y las tecnologías de secuenciación que permiten lecturas de 600 a 700+ pb permitirían una secuencia utilizando sólo un cebador directo (secuenciación desde el extremo 5').

Se puede usar cualquier técnica para secuenciar el ácido nucleico conocido por los expertos en la técnica. Las técnicas de secuenciación de ADN incluyen reacciones de secuenciación didesoxi clásica (método de Sanger) que usan terminadores o cebadores marcados y separación de gel en electroforesis en placa o capilar. En una realización preferida, las plataformas de secuenciación de la próxima generación (NextGen) se usan ventajosamente en la práctica de la invención. La secuenciación NextGen se refiere a cualquiera de una serie de métodos de secuenciación tipo Sanger postclásicos que son capaces de un alto rendimiento, secuenciación múltiple de grandes cantidades de muestras simultáneamente. Las plataformas actuales de secuenciación de NextGen, como las que se describen con mayor detalle a continuación, son capaces de generar lecturas de múltiples ácidos nucleicos distintos en la misma secuencia de ejecución. El rendimiento es variado, con 100 millones de bases a 600 giga base por ejecución, y el rendimiento está aumentando rápidamente debido a las mejoras en la tecnología. El principio de funcionamiento de las diferentes plataformas de secuenciación de NextGen también varía y puede incluir: secuenciación por síntesis utilizando nucleótidos marcados terminados reversiblemente, pirosecuenciación, secuenciación 454, hibridación específica de alelo a una biblioteca de sondas de oligonucleótidos marcadas, secuenciación por síntesis utilizando hibridación específica de alelo a biblioteca de clones marcados que es seguida por ligación, monitorización en tiempo real de la incorporación de nucleótidos marcados durante un paso de polimerización, secuenciación polónica, secuenciación en tiempo real de una sola molécula y secuenciación SOLiD. La secuenciación se ha demostrado mediante reacciones de extensión secuencial o única utilizando polimerasas o ligasas, así como mediante hibridaciones diferenciales únicas o secuenciales con bibliotecas de sondas. Estas reacciones se han realizado en muchas secuencias clonales en paralelo, incluidas demostraciones en aplicaciones comerciales actuales de más de 100 millones de secuencias en paralelo. Estos enfoques de secuenciación pueden utilizarse para estudiar el repertorio del receptor de células T (TCR) y/o del receptor de células B (BCR) y otras secuencias de interés.

Las técnicas de secuenciación pueden generar al menos 1000 lecturas por ejecución, al menos 10 000 lecturas por ejecución, al menos 100 000 lecturas por ejecución, al menos 500 000 lecturas por ejecución o al menos 1 000 000 lecturas por ejecución.

Las técnicas de secuenciación pueden generar aproximadamente 30 bp, aproximadamente 40 bp, aproximadamente 50 bp, aproximadamente 60 bp, aproximadamente 70 bp, aproximadamente 80 bp, aproximadamente 90 bp, aproximadamente 100 bp, aproximadamente 110, aproximadamente 120 bp por lectura, aproximadamente 150 bp, aproximadamente 200 bp, aproximadamente 250 bp, aproximadamente 300 bp, aproximadamente 350 bp, aproximadamente 400 bp, aproximadamente 450 bp, aproximadamente 500 bp, aproximadamente 550 bp, aproximadamente 600 bp, aproximadamente 650 bp, o aproximadamente 700 bp o más bp por lectura.

Las técnicas de secuenciación pueden generar al menos 30, 40, 50, 60, 70, 80, 90, 100, 110, 120, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700 o más nucleótidos por lectura.

Una técnica de secuenciación que se puede utilizar, por ejemplo, Helicos True Single Molecule Sequencing (tSMS) (Harris T. D. et al. (2008) Science 320:106-109). En la técnica tSMS, una muestra de ADN se divide en cadenas de aproximadamente 100 a 200 nucleótidos, y se agrega una secuencia poliA al extremo 3' de cada cadena de ADN. Cada hebra se marca mediante la adición de un nucleótido de adenosina marcado de forma fluorescente. Las cadenas de ADN se hibridan luego con una célula de flujo, que contiene millones de sitios de captura de oligo-T que se inmovilizan en la superficie de la célula de flujo. Las plantillas pueden tener una densidad de alrededor de 100 millones de plantillas/cm<sup>2</sup>. La celda de flujo luego se carga en un instrumento, por ejemplo, el secuenciador HeliScope™, y un láser ilumina la superficie de la celda de flujo, revelando la posición de cada plantilla. Una cámara CCD puede asignar la posición de las plantillas en la superficie de la celda de flujo. La plantilla de la etiqueta fluorescente se escinde y se lava. La reacción de secuenciación comienza con la introducción de una ADN polimerasa y un nucleótido marcado con fluorescencia. El ácido nucleico oligo-T sirve como cebador. La polimerasa incorpora los nucleótidos marcados al cebador de una manera dirigida a la plantilla. La polimerasa y los nucleótidos no incorporados se eliminan. Las plantillas que han incorporado la incorporación del nucleótido marcado de forma fluorescente se detectan mediante imágenes de la superficie de la célula de flujo. Después de la obtención de imágenes, una etapa de escisión elimina la etiqueta fluorescente y el proceso se repite con otros nucleótidos marcados con fluorescencia hasta que se alcanza la longitud de lectura deseada. La información de la secuencia se recopila con cada paso de adición de nucleótidos.

Otro ejemplo de una técnica de secuenciación de ADN que puede usarse es la secuenciación 454 (Roche) (Margulies, M et al. 2005, Nature, 437, 376-380). La secuenciación 454 implica dos pasos. En el primer paso, el ADN se corta en fragmentos de aproximadamente 300-800 pares de bases, y los fragmentos tienen extremos romos. Los adaptadores

- de oligonucleótidos se ligan a los extremos de los fragmentos. Los adaptadores sirven como cebadores para la amplificación y secuenciación de los fragmentos. Los fragmentos se pueden unir a las perlas de captura de ADN, por ejemplo, perlas recubiertas con estreptavidina usando, por ejemplo, el Adaptador B, que contiene la etiqueta 5'-biotina. Los fragmentos unidos a las perlas se amplifican por PCR dentro de gotitas de una emulsión de aceite-agua. El resultado son múltiples copias de fragmentos de ADN amplificados clonalmente en cada cuenta. En el segundo paso, las cuentas se capturan en pozos (tamaño pico-litro). La pirosecuenciación se realiza en cada fragmento de ADN en paralelo. La adición de uno o más nucleótidos genera una señal de luz que es grabada por una cámara CCD en un instrumento de secuenciación. La intensidad de la señal es proporcional al número de nucleótidos incorporados.
- 5
- La pirosecuenciación hace uso del pirofosfato (PPi) que se libera tras la adición de nucleótidos. El PPi se convierte en ATP por la ATP sulfurilasa en presencia de adenosina 5' fosfosulfato. La luciferasa usa ATP para convertir la luciferina en oxiluciferina, y esta reacción genera luz que se detecta y analiza.
- 10
- Otro ejemplo de una técnica de secuenciación de ADN que se puede usar es la tecnología SOLiD (Applied Biosystems). En la secuenciación SOLiD, el ADN genómico se divide en fragmentos y los adaptadores se unen a los extremos 5' y 3' de los fragmentos para generar una biblioteca de fragmentos. Alternativamente, pueden introducirse adaptadores internos ligando los adaptadores a los extremos 5' y 3' de los fragmentos, circularizando los fragmentos, digiriendo el fragmento circularizado para generar un adaptador interno, y uniendo adaptadores a los extremos 5' y 3' de los extremos resultantes. Fragmentos para generar una librería emparejada. A continuación, se preparan poblaciones de perlas clonales en microrreactores que contienen perlas, cebadores, plantilla y componentes de PCR. Después de la PCR, las plantillas se desnaturalizan y las perlas se enriquecen para separar las perlas con plantillas extendidas. Las plantillas en las perlas seleccionadas se someten a una modificación 3' que permite la unión a un portaobjetos de vidrio.
- 15
- La secuencia puede determinarse por hibridación secuencial y ligación de oligonucleótidos parcialmente aleatorios con una base determinada central (o un par de bases) que se identifica por un fluoróforo específico. Después de registrar un color, el oligonucleótido ligado se escinde y se elimina, y luego se repite el proceso.
- 20
- Otro ejemplo de una tecnología de secuenciación que se puede utilizar es la secuenciación SOLEXA (Illumina). La secuenciación de SOLEXA se basa en la amplificación del ADN en una superficie sólida mediante PCR plegable y cebadores anclados. El ADN genómico está fragmentado y los adaptadores se agregan a los extremos 5' y 3' de los fragmentos. Los fragmentos de ADN que se unen a la superficie de los canales celulares de flujo se extienden y se amplifican en puente. Los fragmentos se vuelven de doble cadena, y las moléculas de doble cadena se desnaturalizan. Los ciclos múltiples de la amplificación en fase sólida seguidos de la desnaturalización pueden crear varios millones de agrupaciones de aproximadamente 1000 copias de moléculas de ADN de cadena sencilla de la misma plantilla en cada canal de la celda de flujo. Se utilizan cebadores, ADN polimerasa y cuatro nucleótidos terminados de manera reversible y marcados con fluoróforo para realizar la secuenciación secuencial. Después de la incorporación de nucleótidos, se usa un láser para excitar los fluoróforos, se captura una imagen y se registra la identidad de la primera base. Los terminadores 3' y los fluoróforos de cada base incorporada se eliminan y se repiten los pasos de incorporación, detección e identificación.
- 25
- Otro ejemplo de una tecnología de secuenciación que se puede utilizar incluye la tecnología de molécula única, en tiempo real (SMRT™) de Pacific Biosciences. En SMRT, cada una de las cuatro bases de ADN está unida a uno de cuatro tintes fluorescentes diferentes. Estos tintes están fosfoenlazados. Una sola ADN polimerasa se inmoviliza con una sola molécula de ADN de cadena sencilla de plantilla en la parte inferior de una guía de onda de modo cero (ZMW). Una ZMW es una estructura de confinamiento que permite la observación de la incorporación de un solo nucleótido por la ADN polimerasa contra el fondo de nucleótidos fluorescentes que se difunden rápidamente en una salida de la ZMW (en microsegundos). Se requieren varios milisegundos para incorporar un nucleótido en una hebra en crecimiento. Durante este tiempo, la etiqueta fluorescente se excita y produce una señal fluorescente, y la etiqueta fluorescente se separa. La detección de la fluorescencia correspondiente del tinte indica qué base se incorporó. El proceso se repite.
- 30
- Otro ejemplo de una técnica de secuenciación que se puede usar es la secuenciación de nanoporos (Soni G V y Meller A. (2007) Clin Chem 53: 1996-2001). Un nanoporo es un agujero pequeño, del orden de 1 nanómetro de diámetro. La inmersión de un nanoporo en un fluido conductor y la aplicación de un potencial a través de él da como resultado una liviana corriente eléctrica debido a la conducción de iones a través del nanoporo. La cantidad de corriente que fluye es sensible al tamaño del nanoporo. Cuando una molécula de ADN pasa a través de un nanoporo, cada nucleótido en la molécula de ADN obstruye el nanoporo en un grado diferente. Por lo tanto, el cambio en la corriente que pasa a través del nanoporo a medida que la molécula de ADN pasa a través del nanoporo representa una lectura de la secuencia de ADN.
- 35
- Otro ejemplo de una técnica de secuenciación que se puede usar implica el uso de una matriz de transistores de efecto de campo (ChemFET) sensibles a los agentes químicos para secuenciar el ADN (por ejemplo, como se describe en la Publicación de Solicitud de Patente de Estados Unidos No. 20090026082). En un ejemplo de la técnica, las moléculas de ADN se pueden colocar en cámaras de reacción, y las moléculas de plantilla se pueden hibridar con un cebador de secuenciación unido a una polimerasa. La incorporación de uno o más trifosfatos en una nueva cadena de ácido
- 40
- 45
- 50
- 55

nucleico en el extremo 3' del cebador de secuenciación puede detectarse mediante un cambio en la corriente por parte de un chemFET. Una matriz puede tener múltiples sensores chemFET. En otro ejemplo, se pueden unir ácidos nucleicos individuales a las perlas, y los ácidos nucleicos se pueden amplificar en la perla, y las perlas individuales se pueden transferir a cámaras de reacción individuales en una matriz chemFET, con cada cámara teniendo un sensor chemFET, y los ácidos nucleicos pueden ser secuenciados.

Otro ejemplo de una técnica de secuenciación que se puede usar es el uso de un microscopio electrónico (Moudrianakis E. N. y Beer M. Proc Natl Acad Sci USA. 1965 de marzo; 53:564-71). En un ejemplo de la técnica, las moléculas de ADN individuales se marcan con etiquetas metálicas que se pueden distinguir con un microscopio electrónico. Estas moléculas se estiran luego sobre una superficie plana y se toman imágenes con un microscopio electrónico para medir secuencias.

En algunos aspectos, obtener la biblioteca de polinucleótidos comprende preparar la biblioteca de polinucleótidos en un laboratorio. En algunos aspectos, obtener la biblioteca de polinucleótidos comprende recibir la biblioteca de polinucleótidos directa o indirectamente de un tercero que ha preparado la biblioteca de polinucleótidos.

#### Métodos para analizar los datos de secuenciación

En algunos aspectos, el método incluye obtener un conjunto de datos asociado con una pluralidad de polinucleótidos, en donde el conjunto de datos comprende datos de secuenciación para la pluralidad de polinucleótidos, en donde cada polinucleótido en la pluralidad de polinucleótidos comprende una región de identificación de muestra, y en donde cada región de identificación de muestra en cada polinucleótido es único para una sola muestra, en donde la secuencia de la región de identificación de la muestra de cada polinucleótido derivada de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la pluralidad de polinucleótidos derivados de uno o más muestras distintas de la primera muestra única; y analizar el conjunto de datos para hacer coincidir polinucleótidos con regiones idénticas de identificación de la muestra, en donde una coincidencia indica que los polinucleótidos se originaron en la misma muestra.

En algunos aspectos, cada polinucleótido en la pluralidad de polinucleótidos comprende además una primera región de identificación de la placa, en donde cada combinación de cada primera región de identificación de la placa y la región de identificación de la muestra en cada polinucleótido es única para una muestra única, en donde la secuencia de la primera región de identificación de la placa de cada polinucleótido derivado de un primer conjunto de muestras individuales es distinto de la secuencia de la primera región de identificación de la placa de los otros polinucleótidos en la pluralidad de polinucleótidos derivados de uno o más conjuntos de muestras individuales distintos del primer conjunto de muestras individuales, y además que comprende analizar el conjunto de datos para hacer coincidir polinucleótidos con regiones idénticas de identificación de la primera placa y regiones idénticas de identificación de la muestra, en donde una coincidencia entre ambas regiones indica que los polinucleótidos se originaron en la misma muestra.

En algunos aspectos, ambos polinucleótidos incluyen una región variable. En algunos aspectos, un polinucleótido incluye una región variable. En algunos aspectos, ninguno de los polinucleótidos incluye una región variable.

En algunos aspectos, obtener el conjunto de datos comprende obtener la pluralidad de polinucleótidos y secuenciar la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. En algunos aspectos, obtener el conjunto de datos comprende recibir el conjunto de datos directa o indirectamente de un tercero que ha secuenciado la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. En algunos aspectos, el conjunto de datos se almacena en un medio de almacenamiento electrónico. En algunos aspectos, el conjunto de datos se transfiere a través de Internet.

En algunos aspectos, el método se implementa en un ordenador, por ejemplo, es un método implementado por ordenador.

En algunos aspectos, la muestra única es una sola célula. En algunos aspectos, la muestra única comprende una sola célula. En algunos aspectos, la muestra única comprende una única célula B. En algunos aspectos, la muestra única comprende una pluralidad de células B. En algunos aspectos, la muestra única comprende una sola célula B y una o más células.

En algunos aspectos, los datos generados a partir de la secuenciación (por ejemplo, la secuenciación 454) se pueden analizar mediante el software de análisis de datos GS4X 454, y las secuencias con puntajes de calidad deficiente se pueden filtrar. Las secuencias de buena calidad se pueden subdividir de acuerdo con su región de identificación de la muestra (y en algunas realizaciones, la combinación de su región de identificación de la muestra y la región de identificación de la placa) mediante el uso de un guion en Pitón antes de ensamblar las secuencias utilizando enfoques bioinformáticos, por ejemplo, utilizando Newbler. Debido a que las lecturas reversas pueden tener solo una segunda región de identificación de la placa en algunos aspectos, es posible que el ensamblaje de secuencias pueda ocurrir entre lecturas directas e reversas de secuencias de diferentes células. Para evitar este problema potencial, el uso de V(D)J de la cadena pesada y liviana de las lecturas tanto directas como reversas se puede identificar primero utilizando

HighV-QUEST. Luego, las secuencias se pueden agrupar según su uso V(D)J antes de ser ensambladas. En algunos aspectos, el ensamblaje de secuencias puede ser intolerante con los desajustes de nucleótidos, lo que evita el ensamblaje de lecturas directas e reversas de diferentes células que comparten el mismo uso de V(D)J. En algunos aspectos, las secuencias se pueden agrupar según su uso de V(D)J mediante un programa informático.

- 5 En algunos aspectos, los métodos bioinformáticos se pueden usar para identificar grupos de secuencias que forman familias y subfamilias clonales, y por lo tanto secuencias de interés de inmunoglobulina. Dichos métodos de bioinformática implican medidas de similitud de secuencia. Dichos métodos bioinformáticos pueden usarse para identificar secuencias de interés derivadas de un humano individual, derivadas de uno o más humanos, derivadas de uno o más humanos con una condición, o derivadas de uno o más humanos con diferentes condiciones.
- 10 En algunos aspectos, las secuencias pesadas y/o livianas de inmunoglobulina relacionadas pueden identificarse a través del análisis filogenético computacional de la homología entre la cadena pesada de inmunoglobulina y/o las secuencias V(D)J de cadena liviana. En algunos aspectos, se pueden utilizar métodos de clasificación estándar (es decir, agrupación) de las secuencias que representan el individuo o combinaciones de los segmentos de los genes V, D y/o J y otras secuencias derivadas de la cadena pesada y/o la cadena liviana de inmunoglobulina. para identificar familias clonales o subfamilias (por ejemplo, utilizando ClustalX).
- 15

Como se usa en el presente documento, "familia clonal" se refiere a una pluralidad de secuencias de inmunoglobulina que tienen cada una las regiones V, D y/o J, en donde cada secuencia es una versión mutada de la misma secuencia de inmunoglobulina de la línea germinal que tiene una región V, D y/o J o la secuencia de inmunoglobulina de la línea germinal que tiene la región V, D y/o J. En algunos aspectos, la pluralidad es una pluralidad de secuencias de cadenas pesadas. En algunos aspectos, la pluralidad es una pluralidad de secuencias de cadenas livianas. En algunos aspectos, la pluralidad es una pluralidad de secuencias pareadas de cadenas pesadas y livianas. En algunos aspectos, cada secuencia tiene regiones V, D y J. En algunos aspectos, cada secuencia tiene regiones V y D. En algunos aspectos, cada secuencia tiene regiones D y J. En algunos aspectos, cada secuencia tiene regiones V y J. En algunos aspectos, cada secuencia tiene una V región. En algunos aspectos, cada secuencia tiene una región D. En algunos aspectos, cada secuencia tiene una región J. En algunos aspectos, la una o más mutaciones están ubicadas dentro de las regiones V, D y/o J. En algunos aspectos, la una o más mutaciones están ubicadas entre las regiones V, D y/o J.

20

25

En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de genes V y J son una familia clonal. En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de los genes V y J y cuya suma de la longitud de los nucleótidos P/N y los nucleótidos D son de la misma longitud son una familia clonal. En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de genes V, D y J son una familia clonal. En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de los genes V, D y J, y cuyos nucleótidos P/N entre los segmentos de los genes V y D tienen la misma longitud, y cuyos nucleótidos P/N entre los D y los segmentos del gen J son de la misma longitud, son una familia clonal. En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de los genes V y J y cuyas cadenas livianas usan todos los mismos segmentos de los genes V y J son una familia clonal. En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de los genes V y J y cuya suma de la longitud de los nucleótidos P/N y los nucleótidos D son de la misma longitud, y cuyas cadenas livianas usan la misma V y Los segmentos del gen J y cuyos nucleótidos P/N son de la misma longitud, son una familia clonal. En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de los genes V, D y J y cuyas cadenas livianas usan todos los mismos segmentos de los genes V y J son una familia clonal. En algunos aspectos, un conjunto de anticuerpos cuyas cadenas pesadas usan todos los mismos segmentos de los genes V, D y J, y cuyos nucleótidos P/N entre los segmentos de los genes V y D tienen la misma longitud, y cuyos nucleótidos P /N entre los D y los segmentos del gen J son de la misma longitud, y cuyas cadenas livianas usan todos los mismos segmentos de los genes V y J, y cuyos nucleótidos P/N entre los segmentos de los genes V y J tienen la misma longitud, son una familia clonal.

30

35

40

45

#### Métodos para construir familias clonales

El uso de V, D y J para un receptor de células T (TCR) o una secuencia de consulta del gen variable de inmunoglobulina se puede determinar mediante la identificación de la línea germinal V, D (si corresponde) y los segmentos del gen J que probablemente hayan dado lugar a la secuencia. Los segmentos D están presentes en algunas secuencias de TCR e inmunoglobulina (por ejemplo, TCR $\beta$ , TCR $\delta$  y secuencias de cadena pesada de anticuerpos) pero no en otras (por ejemplo, TCR $\alpha$ , TCR $\gamma$  y secuencias de cadenas livianas de anticuerpos). La siguiente descripción incluye segmentos D, pero los mismos enfoques se pueden aplicar a secuencias de regiones variables que carecen de segmentos D. En todos los casos, la determinación del uso de V (D) J utiliza una base de datos de referencia de las secuencias de los segmentos de los genes V, D y J de la línea germinal, como IMG/ GENE-DB (Giudicelli V, Chaume D, Lefranc MP. IMG/GENE-DB: a comprehensive database for human and mouse immunoglobulin and T cell receptor genes. Nucleic Acids Res. 2005 Jan 1;33(Database issue):D256-61.).

50

55

En una metodología para determinar el uso de V(D)J, la secuencia de consulta se compara en serie con cada segmento génico de la línea germinal V, D y J por separado y el segmento génico más similar de cada tipo (V, D o J) se selecciona como el más probable que haya dado lugar a la secuencia de consulta. V-QUEST y Alto V-QUEST son ejemplos de esta metodología (Giudicelli V, Chaume D, Lefranc MP. *IMGT/V-QUEST, an integrated software program for immunoglobulin and T cell receptor V-J and V-D-J rearrangement analysis*. *Nucleic Acids Res.* 2004 Jul 1;32(Web Server issue):W435-40.; Brochet X, Lefranc MP, Giudicelli V. *IMGT/V-QUEST: the highly customized and integrated system for IG and TR standardized V-J and V-D-J sequence analysis*. *Nucleic Acids Res.* 2008 Jul 1;36(Web Server issue):W503-8.). V-QUEST primero genera alineaciones de pares para la secuencia de consulta y cada secuencia de segmento del gen V. Luego genera alineaciones por pares para la región de secuencia de consulta corriente abajo del extremo 3' deducido del segmento V y cada secuencia de segmento de gen J. Si está presente un segmento D, V-QUEST genera alineaciones por pares para la región de la secuencia de consulta que se encuentra entre las regiones que coinciden con los segmentos V y J y cada secuencia de segmento del gen D. V-QUEST también puede inferir los límites de las regiones de unión V-D, V-J y/o D-J.

En otro enfoque para determinar el uso de V(D)J, la combinación de los segmentos V, D y J de la línea germinal que probablemente hayan dado lugar a la secuencia de consulta se identifica en un solo paso en lugar de en tres pasos separados para V, D y J respectivamente. Esta metodología tiene la ventaja de que la identificación de un tipo de segmento (V, D o J) puede tener en cuenta la información sobre las posibles coincidencias con los otros dos tipos de segmentos. Por ejemplo, el mejor segmento D coincidente podría depender de qué coincidencia del segmento V se está considerando. SoDA es un ejemplo de esta metodología (Volpe JM, Cowell LG, Kepler TB. *SoDA: implementation of a 3D alignment algorithm for inference of antigen receptor recombinations*. *Bioinformatics.* 2006 Feb 15;22(4):438-44.). SoDA primero selecciona los segmentos candidatos V, D y J. Genera alineaciones locales por pares para la secuencia de consulta y cada secuencia de segmento del gen V y, a continuación, mantiene solo los segmentos V con alineaciones que cumplen con un umbral de puntuación. Repite estos pasos para los segmentos J y los segmentos D. Luego se genera una alineación óptima para cada combinación posible de segmentos candidatos V, D y J. Las alineaciones se generan utilizando el mismo enfoque de programación dinámica general ampliamente utilizado en la alineación de secuencias (Needleman, S.B. and Wunsch, C.D. (1970) *A general method applicable to the search for similarities in the amino acid sequence of two proteins*. *J. Mol. Biol.*, 48,443-453.), Pero permitiendo la inserción de nucleótidos adicionales en las uniones VD, VJ y/o DJ. Dicha inserción comúnmente tiene lugar durante el proceso biológico de la recombinación V(D)J. En la alineación de secuencias mediante la programación dinámica, normalmente hay puntuaciones de penalización asociadas con inserciones, eliminaciones y desajustes. Sin embargo, en esta metodología para determinar el uso de V(D)J, no se aplican penalizaciones para la inserción de nucleótidos en las uniones entre segmentos. La combinación V(D)J que produce la alineación con la puntuación más alta se selecciona para indicar el uso de V(D)J para la secuencia de consulta. Esta metodología también puede identificar los límites de las regiones de secuencia de unión.

De las familias y subfamilias clonales, se puede usar una variedad de metodologías para seleccionar clones específicos para la expresión de sus genes de inmunoglobulina de cadena pesada y liviana emparejados codificados y la caracterización de sus propiedades de unión. En algunos aspectos, los clones de mayor frecuencia de las familias clonales y/o las subfamilias clonales, así como otros clones representativos de las familias clonales y las subfamilias, se expresan y se analizan para determinar sus propiedades de unión. Los clones también pueden seleccionarse al azar, de todos los clones, de todas las familias clonales seleccionadas o, y/o de todas las subfamilias clonales seleccionadas o para la expresión y caracterización de sus características de unión. Los clones también pueden seleccionarse basándose en que poseen un mayor número de variaciones en la región variable del anticuerpo. Se puede construir un árbol filogenético y los clones pueden seleccionarse en función de las características del árbol filogenético, por ejemplo, al descender el árbol, siempre seleccionando la rama con el mayor número de nodos de hojas debajo.

En algunos aspectos, un método incluye además la selección de uno o más polinucleótidos para la clonación.

Métodos para identificar un segundo polinucleótido de interés con base en la selección de un primer polinucleótido de interés

En algunos aspectos, el método incluye obtener un conjunto de datos asociado con una pluralidad de polinucleótidos, en donde el conjunto de datos comprende datos de secuenciación para la pluralidad de polinucleótidos, en donde cada polinucleótido en la pluralidad de polinucleótidos comprende una región de identificación de muestra, y en donde cada región de identificación de muestra en cada polinucleótido es único para una sola muestra, asociando así cada polinucleótido en la pluralidad de polinucleótidos con una muestra única distinta, en donde la secuencia de la región de identificación de la muestra de cada polinucleótido derivada de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la pluralidad de polinucleótidos derivados de una o más muestras distintas de la primera muestra única; y seleccionando un primer polinucleótido de interés asociado con una primera muestra única del conjunto de datos e identificando un segundo polinucleótido de interés en la primera muestra única en base a la región de identificación de la muestra del primer polinucleótido de interés.



En algunos aspectos, cada polinucleótido en la pluralidad de polinucleótidos comprende además una primera región de identificación de la placa, en donde cada combinación de cada primera región de identificación de la placa y la región de identificación de la muestra en cada polinucleótido es única para una sola muestra, en donde la secuencia de la primera identificación de la placa la región de cada polinucleótido derivada de un primer conjunto de muestras  
 5 únicas es distinta de la secuencia de la primera región de identificación de la placa de los otros polinucleótidos en la pluralidad de polinucleótidos derivados de uno o más conjuntos de muestras individuales distinta del primer conjunto de muestras individuales, y que comprende además la identificación de un segundo polinucleótido de interés en la primera muestra única en base a la región de identificación de la muestra y la primera región de identificación de la placa del primer polinucleótido de interés.

10 En algunos aspectos, ambos polinucleótidos incluyen una región variable. En algunos aspectos, un polinucleótido incluye una región variable. En algunos aspectos, ninguno de los polinucleótidos incluye una región variable.

En algunos aspectos, el método se implementa en un ordenador, por ejemplo, es un método implementado por ordenador.

15 En algunos aspectos, la primera muestra única comprende una célula B. En algunos aspectos, la primera muestra única comprende una única célula B y una o más células. En algunos aspectos, la primera muestra única comprende una pluralidad de células B. En algunos aspectos, la primera muestra única comprende una célula B, en donde el primer polinucleótido de interés comprende una secuencia de nucleótidos de cadena pesada de anticuerpo, y en donde el segundo polinucleótido de interés comprende una secuencia de nucleótidos de cadena liviana de anticuerpo. En algunos aspectos, la primera muestra única comprende una célula B, en donde el primer polinucleótido de interés  
 20 comprende una secuencia de nucleótidos de cadena liviana de anticuerpo, y en donde el segundo polinucleótido de interés comprende una secuencia de nucleótidos de cadena pesada de anticuerpo.

En algunos aspectos, obtener el conjunto de datos comprende obtener la pluralidad de polinucleótidos y secuenciar la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. En algunos aspectos, obtener el conjunto de datos comprende recibir el conjunto de datos directa o indirectamente de un tercero que ha secuenciado la pluralidad de polinucleótidos para determinar experimentalmente el conjunto de datos. En algunos aspectos, el conjunto de datos se almacena en un medio de almacenamiento electrónico.  
 25

Métodos para producir uno o más polinucleótidos de interés para la clonación.

En algunos aspectos, el método incluye obtener una biblioteca de polinucleótidos que comprende una pluralidad de polinucleótidos, en donde cada polinucleótido comprende una región de cebador universal, una región de identificación de muestra, una región de adaptador y una región de amplicón derivada de una sola muestra, en donde la secuencia de la región del cebador universal es sustancialmente idéntica en cada polinucleótido en la pluralidad de polinucleótidos, y en donde la secuencia de la región de identificación de la muestra de cada polinucleótido derivada de una primera muestra única es distinta de la secuencia de la región de identificación de la muestra de los otros polinucleótidos en la biblioteca derivado de una o más muestras distintas de la primera muestra única; y amplificar la biblioteca de polinucleótidos con un conjunto de cebadores para producir uno o más polinucleótidos de interés para la clonación, en donde el uno o más polinucleótidos de interés para la clonación comprende una primera región de sitio de restricción, una región de cebador universal, una región de identificación de muestra, una región del adaptador, una región de amplicón derivada de una sola muestra y una segunda región de sitio de restricción.  
 30  
 35

En algunos aspectos, obtener la biblioteca de polinucleótidos comprende preparar la biblioteca de polinucleótidos en un laboratorio. En algunos aspectos, obtener la biblioteca de polinucleótidos comprende recibir la biblioteca de polinucleótidos directa o indirectamente de un tercero que ha preparado la biblioteca de polinucleótidos.  
 40

En algunos aspectos, un método incluye además la clonación de uno o más polinucleótidos, por ejemplo, en un vector descrito en el presente documento.

Métodos para producir una molécula de interés.

45 En algunos aspectos, el método incluye obtener una célula huésped que comprende un polinucleótido de interés; y cultivar la célula huésped en condiciones suficientes para producir la molécula de interés.

En algunos aspectos, la obtención de la célula huésped comprende la preparación de la célula huésped que comprende el polinucleótido en un laboratorio. En algunos aspectos, obtener la célula huésped comprende recibir la célula huésped que comprende el polinucleótido directa o indirectamente de un tercero que ha preparado la célula huésped.  
 50

En algunos aspectos, la molécula de interés es un polipéptido. En algunos aspectos, la molécula de interés es un anticuerpo. En algunos aspectos, la molécula de interés es un anticuerpo monoclonal humano.

En algunos aspectos, el método incluye además la recolección de la molécula de interés.

En algunos aspectos, es deseable "replegar" ciertos polipéptidos, por ejemplo, polipéptidos que comprenden uno o más componentes de ABP o el propio ABP. En ciertas realizaciones, tales polipéptidos se producen usando sistemas de expresión descritos en el presente documento. En ciertas realizaciones, los polipéptidos se "repliegan" y /o se oxidan para formar la estructura terciaria deseada y/o para generar enlaces disulfuro. En ciertas realizaciones, dicha estructura y/o enlaces están relacionados con cierta actividad biológica de un polipéptido. En ciertas realizaciones, el replegamiento se realiza utilizando cualquiera de una serie de procedimientos conocidos en la técnica. Los métodos de ejemplo incluyen, pero no se limitan a, exponer el agente polipeptídico solubilizado a un pH típicamente superior a 7 en presencia de un agente caotrópico. Un agente caotrópico de ejemplo es la guanidina. En ciertas realizaciones, la solución de replegamiento/oxidación también contiene un agente reductor y la forma oxidada de ese agente reductor.

En ciertas realizaciones, el agente reductor y su forma oxidada están presentes en una relación que generará un potencial rédox particular que permite que ocurra la mezcla de disulfuro. En ciertas realizaciones, tal barajado permite la formación de puentes de cisteína. Las parejas rédox de ejemplo incluyen, pero no se limitan a, cisteína/cistamina, glutatión/ditiobisGSH, cloruro cúprico, ditiotreitól DTT/ditiane DTT y 2-mercaptoetanol (bME)/ditio-bME. En ciertas realizaciones, se usa un codisolvente para aumentar la eficiencia del replegamiento. Los codisolventes de ejemplo incluyen, pero no se limitan a, glicerol, polietilenglicol de diversos pesos moleculares y arginina.

En ciertas realizaciones, uno purifica sustancialmente un polipéptido, por ejemplo, un polipéptido que comprende uno o más componentes de ABP o el propio ABP. Los expertos en la técnica conocen ciertas técnicas de purificación de proteínas. En ciertas realizaciones, la purificación de proteínas implica el fraccionamiento en bruto de fracciones de polipéptidos de fracciones no polipeptídicas. En ciertas realizaciones, los polipéptidos se purifican usando técnicas cromatográficas y/o electroforéticas. Los métodos de purificación de ejemplo incluyen, pero no se limitan a, precipitación con sulfato de amonio; precipitación con PEG; inmunoprecipitación; desnaturalización por calor seguida de centrifugación; cromatografía, que incluye, pero no se limita a, cromatografía de afinidad (por ejemplo, Proteína-A-Sepharose), cromatografía de intercambio iónico, cromatografía de exclusión y cromatografía de fase reversa; filtración en gel; cromatografía de hidroxapatita; enfoque isoelectrico; electroforesis en gel de poli(acrilamida); y combinaciones de tales y otras técnicas. En ciertas realizaciones, un polipéptido se purifica por cromatografía líquida rápida de proteínas o por cromatografía líquida a alta presión (HPLC). En ciertas realizaciones, las etapas de purificación se pueden cambiar o ciertas etapas se pueden omitir y aún dar como resultado un método adecuado para la preparación de un polipéptido sustancialmente purificado.

En ciertas realizaciones, se cuantifica el grado de purificación de una preparación de polipéptido. Ciertos métodos para cuantificar el grado de purificación son conocidos por los expertos en la técnica. Ciertos métodos de ejemplo incluyen, pero no se limitan a, determinar la actividad de unión específica de la preparación y evaluar la cantidad de un polipéptido dentro de una preparación mediante análisis SDS/PAGE. Ciertos métodos de ejemplo para evaluar la cantidad de purificación de una preparación polipeptídica comprenden calcular la actividad de unión de una preparación y compararla con la actividad de unión de un extracto inicial. En ciertas realizaciones, los resultados de dicho cálculo se expresan como "purificación por plegado". Las unidades utilizadas para representar la cantidad de actividad de unión dependen del ensayo particular realizado.

En ciertas realizaciones, un polipéptido que comprende uno o más componentes de ABP o la propia ABP se purifica parcialmente. En ciertas realizaciones, la purificación parcial se puede lograr usando menos pasos de purificación o utilizando diferentes formas del mismo esquema de purificación general. Por ejemplo, en ciertas realizaciones, la cromatografía en columna de intercambio catiónico realizada utilizando un aparato de HPLC generalmente dará como resultado una mayor "purificación" que la misma técnica que utiliza un sistema de cromatografía de baja presión. En ciertas realizaciones, los métodos que dan como resultado un grado menor de purificación pueden tener ventajas en la recuperación total del polipéptido, o en el mantenimiento de la actividad de unión de un polipéptido.

En ciertos casos, la migración electroforética de un polipéptido puede variar, a veces significativamente, con diferentes condiciones de SDS/PAGE. Véase, por ejemplo, Capaldi et al, Biochem. Biophys. Res. Comm., 76: 425 (1977). Se apreciará que bajo diferentes condiciones de electroforesis, los pesos moleculares aparentes del polipéptido purificado o parcialmente purificado pueden ser diferentes.

#### Métodos de criba

En algunos aspectos, una molécula de interés se analiza para determinar su actividad. En algunos aspectos, la molécula de interés es un ABP. En algunos aspectos, la molécula de interés es un anticuerpo.

En algunos aspectos, los métodos de selección de las bibliotecas descritas en el presente documento se utilizan para identificar ABPs capaces de unirse a un objetivo deseado. Se contempla cualquier método de selección *in vitro* o *in vivo* que permita la selección de un ABP de una biblioteca, con base en la unión de ABP a una molécula diana.

En una realización, una biblioteca puede explorarse utilizando una criba unida a genotipo-genotipo libre de células *in vitro* reconocida en la técnica. Dichos métodos son bien conocidos en la técnica y se describen, por ejemplo, en las Patentes de Estados Unidos Nos. 7,195,880; 6,951,725; 7,078,197; 7,022,479; 6,518,018; 7,125,669; 6,846,655; 6,281,344; 6,207,446; 6,214,553; 6,258,558; 6,261,804; 6,429,300; 6,489,116; 6,436,665; 6,537,749; 6,602,685; 6,623,926; 6,416,950; 6,660,473; 6,312,927; 5,922,545; y 6,348,315. Estos métodos implican la transcripción de

proteínas in vitro a partir de un ácido nucleico de tal manera que la proteína está físicamente asociada o unida al ácido nucleico del cual se originó. Al seleccionar una proteína expresada con una molécula diana, también se puede seleccionar el ácido nucleico que codifica la proteína.

5 Para mejorar la expresión de las proteínas scFv, los ensayos de cribado in vitro mencionados anteriormente pueden incluir la adición o eliminación de ciertos reactivos. En una realización, pueden añadirse enzimas isomerasas de disulfuro de proteínas al sistema de expresión in vitro para mejorar la producción de moléculas scFv funcionales. En otra realización, se puede agregar un agente oxidante suave (por ejemplo, GSSG (glutatiión oxidado)/GSH (glutatiión reducido), por ejemplo, GSSG 100 mM/GSH 10 mM) a la mezcla de reacción de traducción in vitro de las proteínas scFv para permitir formación de enlaces disulfuro de cadena en las regiones VH y VL de la molécula scFv. En otra  
10 realización, los agentes reductores (por ejemplo, ditioneitol (DTT)) se pueden eliminar de la mezcla de reacción de traducción in vitro del scFv.

En otra realización, uno o más aminoácidos marcados, o derivados de los mismos, se pueden agregar al sistema de traducción in vitro de manera que el aminoácido(s) marcado(s) se incorpore en el anticuerpo resultante. Se contempla cualquier aminoácido marcado reconocido en la técnica, por ejemplo, un aminoácido radiomarcado, por ejemplo, metionina o cisteína marcada con <sup>35</sup>S.  
15

En una realización, los ensayos de selección in vitro pueden incluir que después de la selección in vitro de un anticuerpo o la pluralidad de anticuerpos, el ARNm que está físicamente asociado con el anticuerpo o la pluralidad de anticuerpos se puede transcribir de manera reversa para generar un ADNc que codifica dicho anticuerpo o la pluralidad de anticuerpos. Se contempla cualquier método adecuado para la transcripción reversa, por ejemplo, mediada por  
20 enzimas, por ejemplo, transcriptasa reversa del virus de la leucemia murina de Moloney.

Los métodos de selección pueden incluir la amplificación del ácido nucleico que codifica los anticuerpos que se unen específicamente a un objetivo deseado. En una realización, el ARNm que está asociado físicamente con un anticuerpo o una pluralidad de anticuerpos puede amplificarse para producir más ARNm. Se contempla cualquier método de replicación de ARN reconocido en la técnica, por ejemplo, usando una enzima ARN replicasa. En otra realización, el  
25 ARNm que está asociado físicamente con un anticuerpo o una pluralidad de anticuerpos primero se transcribe de forma reversa en ADNc antes de amplificarse por PCR. En una realización, la amplificación por PCR se realiza utilizando una polimerasa de lectura de alta fidelidad, por ejemplo, la ADN polimerasa termoestable KOD1 de Thermococcus kodakaraensis o Platinum Taq ADN Polymerase High Fidelity (Invitrogen, Carlsbad, California). En otra realización, la amplificación por PCR puede realizarse en condiciones que dan como resultado la introducción de mutaciones en el ADN amplificado, es decir, PCR propensa a errores.  
30

Los métodos de cribado también pueden incluir que la rigurosidad del ensayo de cribado de unión a la diana se incremente para seleccionar anticuerpos con afinidad mejorada por la diana. Se contemplan todos los métodos reconocidos en la técnica para aumentar la rigurosidad de un ensayo de interacción anticuerpo-diana. En una  
35 realización, una o más de las condiciones de ensayo pueden variarse (por ejemplo, la concentración de sal del regulador de ensayo) para reducir la afinidad de las moléculas de anticuerpo por el objetivo deseado. En otra realización, se puede reducir el tiempo permitido para que los anticuerpos se unan al objetivo deseado. En otra realización, puede añadirse una etapa de unión competitiva al ensayo de interacción anticuerpo-diana. Por ejemplo, primero se puede permitir que los anticuerpos se unan a un objetivo inmovilizado deseado. Luego se puede agregar una concentración específica de diana no inmovilizada, que sirve para competir por la unión con la diana inmovilizada,  
40 de modo que los anticuerpos con la afinidad más baja por el antígeno se eluyan de la diana inmovilizada, lo que resulta en el enriquecimiento de los anticuerpos con afinidad mejorada de unión al antígeno. En una realización, la rigurosidad de las condiciones de ensayo puede aumentarse adicionalmente aumentando la concentración de diana no inmovilizada que se añade al ensayo.

Los métodos de selección también pueden incluir múltiples rondas de selección para enriquecer uno o más anticuerpos con una unión mejorada al blanco. En una realización, en cada ronda de selección pueden introducirse otras mutaciones de aminoácidos en los anticuerpos usando métodos reconocidos en la técnica. En otra realización, en cada ronda de selección, la rigurosidad de la unión al objetivo deseado puede aumentarse para seleccionar anticuerpos con afinidad incrementada para un objetivo deseado.  
45

Los métodos de selección pueden incluir la purificación de proteínas de fusión ARN-anticuerpo de los componentes de un sistema de traducción in vitro. Esto se puede lograr utilizando cualquier método de separación reconocido en la técnica. En una realización, las proteínas de fusión ARN-anticuerpo pueden separarse por cromatografía utilizando una resina de polidesoxitimidina (polydT). En otra realización, las proteínas de fusión ARN-anticuerpo pueden separarse por cromatografía usando un anticuerpo específico para un epítipo presente en el componente de anticuerpo de la proteína de fusión ARN-anticuerpo. En una realización, el epítipo puede ser una etiqueta de  
50 secuencia de aminoácidos, por ejemplo, etiquetas FLAG o HA, incorporadas en la secuencia de aminoácidos del componente de anticuerpo de la proteína de fusión ARN-anticuerpo, por ejemplo, en el extremo N-terminal, C-terminal o en el enlazador de región intervariable.  
55

La selección de anticuerpos de una biblioteca puede incluir el uso de moléculas diana inmovilizadas. En una realización, la molécula diana puede unirse directamente a un sustrato sólido, por ejemplo, perlas de agarosa. En otra realización, la molécula diana puede modificarse primero, por ejemplo, la molécula diana biotinilada y modificada puede unirse mediante la modificación a un soporte sólido, por ejemplo, estreptavidina-M280, neutravidina-M280, SA-M270, NA-M270, SA-MyOne, NA-MyOne, SA-agarosa y NA-agarosa.

En algunos aspectos, los antígenos marcados con fluorescencia se usan para clasificar de una sola célula solo plasmoblastos u otras células del linaje B con reactividad contra antígenos marcados específicos. En otros aspectos, los antígenos marcados con fluorescencia se usan para enriquecer los plasmoblastos u otras células del linaje B con reactividad contra antígenos marcados específicos, antes de que ocurra la selección de células individuales. En algunos aspectos, se pueden usar moléculas fluorogénicas o cromogénicas para identificar y clasificar las células del linaje B. En algunos aspectos, los plasmoblastos deseados u otras células del linaje B pueden aislarse mediante clasificación celular activada magnéticamente (MACS) o incluso mediante desplazamiento. Los productos resultantes son generalmente anticuerpos monoclonales, contra una variedad de objetivos, entre los que se incluyen: antígenos del cáncer, citoquinas, quimioquinas, factores de crecimiento, proteínas secretadas, superficie celular y otros antígenos para agotar tipos de células de interés específico, microbios, bacterias, micobacterias, parásitos y virus. Otros métodos de selección se describen en la sección de Ejemplos a continuación.

#### Implementación por ordenador

En algunos aspectos, uno o más métodos descritos aquí pueden implementarse en un ordenador. En una realización, un ordenador comprende al menos un procesador acoplado a un conjunto de chips. También se acoplan al conjunto de chips una memoria, un dispositivo de almacenamiento, un teclado, un adaptador de gráficos, un dispositivo señalador y un adaptador de red. Una pantalla está acoplada al adaptador de gráficos. En una realización, la funcionalidad del conjunto de chips es proporcionada por un concentrador de controlador de memoria y un concentrador de controlador de E/S. En otra realización, la memoria está acoplada directamente al procesador en lugar del conjunto de chips.

El dispositivo de almacenamiento es cualquier dispositivo capaz de almacenar datos, como un disco duro, una memoria de solo lectura (CD-ROM), un disco duro o un dispositivo de memoria de estado sólido. La memoria contiene instrucciones y datos utilizados por el procesador. El dispositivo señalador puede ser un ratón, bola de desplazamiento u otro tipo de dispositivo señalador, y se usa en combinación con el teclado para ingresar datos en el sistema informático. El adaptador de gráficos muestra imágenes y otra información en la pantalla. El adaptador de red acopla el sistema informático a una red de área local o amplia.

Como se conoce en la técnica, un ordenador puede tener componentes diferentes y/o diferentes a los descritos anteriormente. Además, el ordenador puede carecer de ciertos componentes. Además, el dispositivo de almacenamiento puede ser local y/o remoto desde el ordenador (tal como al estar incorporado dentro de una red de área de almacenamiento (SAN)).

Como se sabe en la técnica, el ordenador está adaptado para ejecutar módulos de programas informáticos para proporcionar la funcionalidad descrita en este documento. Como se usa en este documento, el término "módulo" se refiere a la lógica del programa de ordenador utilizada para proporcionar la funcionalidad especificada. Por lo tanto, un módulo puede implementarse en hardware, firmware y/o software. En una realización, los módulos de programa se almacenan en el dispositivo de almacenamiento, se cargan en la memoria y son ejecutados por el procesador.

Las realizaciones de las entidades descritas en este documento pueden incluir módulos distintos y/o diferentes a los descritos aquí. Además, la funcionalidad atribuida a los módulos puede ser realizada por otros o diferentes módulos en otras realizaciones. Además, esta descripción ocasionalmente omite el término "módulo" por razones de claridad y conveniencia.

#### Kits

Un kit puede incluir un polinucleótido, una biblioteca de polinucleótidos, un vector y/o una célula huésped descrita en el presente documento e instrucciones de uso. Los kits pueden comprender, en un recipiente adecuado, un polinucleótido, una biblioteca de polinucleótidos, un vector y/o una célula huésped descrita en el presente documento, uno o más controles, y varios reguladores, reactivos, enzimas y otros ingredientes estándar bien conocidos en la técnica.

El contenedor puede incluir al menos un pozo en una placa que comprende uno o más pozos. El contenedor puede incluir al menos un vial, tubo de ensayo, matraz, botella, jeringa u otro medio de contenedor, en el cual se puede colocar un polinucleótido, una biblioteca de polinucleótidos, un vector y/o una célula huésped, y en algunos casos, adecuadamente alícuotas. Cuando se proporciona un componente adicional, el kit puede contener contenedores adicionales en los que se puede colocar este componente. Los kits también pueden incluir un medio para contener el polinucleótido, una biblioteca de polinucleótidos, un vector y/o una célula huésped y cualquier otro contenedor de reactivos en confinamiento cercano para la venta comercial. Dichos recipientes pueden incluir recipientes de plástico

moldeados por inyección o soplado en los cuales se retienen los viales deseados. Los contenedores pueden incluir etiquetado con instrucciones de uso y/o advertencias.

#### Ejemplos

5 Los ejemplos se ofrecen solo con fines ilustrativos, y no pretenden limitar el alcance de ninguna realización de la presente invención de ninguna manera. Se han realizado esfuerzos para garantizar la precisión con respecto a los números utilizados (por ejemplo, cantidades, temperaturas, etc.), pero, por supuesto, se deben permitir algunos errores y desviaciones experimentales.

10 Diversos métodos pueden emplear, a menos que se indique lo contrario, métodos convencionales de química de proteínas, bioquímica, técnicas de ADN recombinante y farmacología, dentro de los conocimientos de la técnica. Tales técnicas se explican completamente en la literatura. Véase, por ejemplo, T.E. Creighton, *Proteins: Structures and Molecular Properties* (W.H. Freeman and Company, 1993); A.L. Lehninger, *Biochemistry* (Worth Publishers, Inc., current addition); Sambrook, et al., *Molecular Cloning: A Laboratory Manual* (2nd Edition, 1989); *Methods In Enzymology* (S. Colowick and N. Kaplan eds., Academic Press, Inc.); *Remington's Pharmaceutical Sciences*, 18th Edition (Easton, Pennsylvania: Mack Publishing Company, 1990); *Carey and Sundberg Advanced Organic Chemistry* 3rd Ed. (Plenum Press) Vols A and B(1992); *Current Protocols in Molecular Biology* (2002- ; Wiley; Online ISBN: 9780471142720; DOI: 10.1002/0471142727); *Current Protocols in Immunology* (2001-; Wiley; Online ISBN: 9780471142737; DOI: 10.1002/0471142735).

#### Materiales y Métodos Generales

##### Recolección de sangre y aislamiento de PBMCs.

20 Todas las muestras humanas se recolectaron después de consentimiento informado y bajo los protocolos para sujetos humanos aprobados por la Investigational Review Board (IRB). La sangre se recogió en tubos de heparina (Beckton Dickinson and Company, catálogo # BD366664) o en tubos de CPT (Beckton Dickinson and Company, catálogo BD362761). Para el procesamiento de los tubos de heparina, se transfirió un mililitro de sangre a un tubo de microcentrifuga y se centrifugó a 12,000 rpm durante 3 minutos, se recogió el plasma y se congeló a -80°C (para su posterior análisis de las reactividades de anticuerpos), el resto de la sangre se colocó sobre Ficoll y se centrifugó en una centrifuga de sobremesa Beckman Coulter Allegra X-15R con un rotor de cubeta giratoria SX4750 a 800 g para tubos de heparina durante 20 minutos a temperatura ambiente, con una aceleración mínima y sin uso del freno y la sangre periférica se recogió la capa de células mononucleares (PBMC). Alternativamente, los tubos CPT se centrifugaron directamente a 1,500 g durante 20 minutos a temperatura ambiente, con una aceleración mínima y sin uso del freno, y se recogió la capa de PMBC. Las PBMC recogidas se lavaron luego dos veces con PBS antes de su uso.

35 Las PBMC también pueden congelarse para su uso y aislamiento futuros de células B, células B de memoria, plasmoblastos, células plasmáticas u otras poblaciones de células B. Un método para congelar PBMC implica volver a suspender las PBMC en 90% de suero bovino fetal (FBS) y 10% de dimetilsulfóxido (DMSO) en crioviales, y luego congelar lentamente las células contenidas en los viales durante la noche a -80°C en un Mr. Frosty (Sigma C1562-1EA). Luego, los viales de células congeladas se transfirieron para el almacenamiento a largo plazo en nitrógeno líquido, y se pueden descongelar en una fecha posterior para el aislamiento de células B individuales y para la secuenciación de alto rendimiento de genes de inmunoglobulina pareados. Las células descongeladas se incubaron en un medio que contenía un exceso de DNasa I, generalmente 25 ug/ml(Sigma D4513) hasta el final de la primera clase para evitar la aglomeración de células.

##### Aislamiento y enriquecimiento de células y subpoblaciones celulares.

45 *Plasmoblastos.* Para algunas muestras, las PBMC se enriquecieron primero para plasmoblastos utilizando un Kit de Aislamiento de Células Plasmáticas modificado II (Miltenyi 130-093-628). Este es un paso opcional. Esto produjo menos células totales para la clasificación posterior, lo que resultó en tiempos de clasificación más cortos. Esto se usó principalmente cuando varias muestras debían clasificarse en una sola célula el mismo día. También es posible usar diferentes kits para enriquecer las diferentes poblaciones de células B (véase más abajo). Por cada  $5 \times 10^7$  PBMC, las células se suspendieron en 200  $\mu$ L de regulador MACS helado (PBS con 0.5% de FBS). Se agregaron 50  $\mu$ L de cóctel de anticuerpos de biotina y células no plasmáticas, y las células se incubaron en el refrigerador (4°C) durante 10 minutos. Se agregaron 100  $\mu$ L de regulador MACS, 100  $\mu$ L de cóctel de microperlas de células no plasmáticas y 50  $\mu$ L de microperlas CD56 y se incubaron en la nevera durante 10 minutos adicionales. Luego, las células se lavaron con 7 mL de regulador MACS, se centrifugaron a 300 g durante 5 minutos a 4°C, se resuspendieron en 500  $\mu$ L de regulador MACS y se ejecutaron en una columna de LS equilibrada en un campo magnético. La columna se lavó con 4 x 3 mL de regulador MACS y las células enriquecidas estaban en la fracción negativa.

55 *Células B de memoria.* Las microperlas CD19+ (Miltenyi 130-050-301) y las microperlas CD27+(130-051-601) pueden usarse para enriquecer las células B de memoria antes de la clasificación celular, para acortar los tiempos de clasificación. También se pueden usar otros métodos de enriquecimiento, como el kit de aislamiento de células B de

Memoria (Miltenyi 130-093-546), siempre que se enriquezcan para células CD19<sup>+</sup>CD27<sup>+</sup>. Para cada 5x10<sup>7</sup> PBMC, se utilizan 300 µL de regulador MACS helado para la resuspensión. Luego se agregan 100 µL de microperlas CD19 y 100 µL de microperlas CD27, y la muestra se incuba a 4°C durante 15 minutos. Las células se lavan luego con 7 mL de regulador MACS, se centrifugan a 300g durante 5 minutos a 4°C y se resuspenden en 500 µL de regulador MACS.

5 Luego, las células se ejecutan a través de una columna LS equilibrada en un campo magnético y se lavan con 2x 3 mL de regulador MACS. La columna LS se retira del campo magnético y las células se lavan con 5 mL de regulador MACS para eluir las células enriquecidas.

10 Células B totales. Las microperlas de CD19<sup>+</sup> (Miltenyi 130-050-301) se pueden usar para enriquecer las células B totales antes de la clasificación celular, por ejemplo, para acortar los tiempos de clasificación. También se pueden usar otros métodos de enriquecimiento, siempre que enriquezcan células CD19<sup>+</sup>. Por cada 5x10<sup>7</sup> PBMC, resuspender las células en 400 µL de regulador MACS helado. Se agregan 100 µL de microperlas CD19<sup>+</sup> e incuban en la nevera (4°C) durante 15 minutos. Las células se lavan luego con 7 mL de regulador MACS, se centrifugan a 300 g durante 5 minutos a 4°C y se resuspenden en 500 µL de regulador MACS. Luego, las células se ejecutan a través de una columna de LS equilibrada en un campo magnético y se lavan con 2x 3 mL de regulador MACS. La columna LS se retira del campo magnético y las células se eluyen con 5 mL de regulador MACS, lo que produce las células enriquecidas.

20 Otros tipos de células. Aunque no es necesario, el enriquecimiento de MACS de la población celular deseada puede acortar los tiempos de clasificación. Otras poblaciones de células, incluidas las células plasmáticas, otras poblaciones de células B y poblaciones de células-no-B, también pueden enriquecerse utilizando MACS u otros sistemas utilizando los reactivos apropiados. Por ejemplo, las células T totales pueden enriquecerse utilizando microperlas CD3<sup>+</sup>, y las células T efectoras y las células-T auxiliares se pueden aislar usando microperlas CD8<sup>+</sup> y CD4<sup>+</sup>, respectivamente. Las microperlas CD45RO se pueden usar para aislar células-T de memoria y, junto con las cuentas CD8<sup>+</sup> o CD4<sup>+</sup>, se pueden usar para aislar células-T efectoras de memoria o auxiliares de memoria, respectivamente.

#### Selección de células individuales

25 El enriquecimiento de MACS no se requiere para la selección, pero el enriquecimiento de MACS para plasmoblastos puede realizarse para acortar los tiempos de selección. Si las PBMC han sido enriquecidas con MACS, también se analiza una porción de PBMC no enriquecidas (~1 millón de células) en tándem, lo que permite determinar el porcentaje de plasmoblastos de referencia en la muestra. Para la selección de los plasmoblastos, las células se tñeron con volúmenes recomendados por el fabricante de CD3-V450 (BD 560365), IgA-FITC (AbD Serotec STAR142F), IgM-FITC (AbD Serotec STAR146F) o IgM-PE (AbD Serotec STAR146PE), CD20 PerCP-Cy5.5 (BD 340955), CD38-PE-Cy7 (BD 335808), CD19-APC (BD 340437) y CD27-APC-H7 (BD 560222) en 50 µL de regulador FACS (PBS o HBSS con 2%) FBS) en hielo durante 20 minutos en la oscuridad. Algunas células también pueden tñirse con IgG-PE (BD 555787), CD138-PE (eBioscience 12-1389-42), o HLA-DR-PE (BD 555812) junto con IgM-FITC en su lugar. Para la selección simultánea de plasmoblastos, memoria y células B vírgenes, se utilizó el siguiente esquema de tñición: IgD-FITC (Biolegend 348205), IgG-PE (BD 555787), CD20-PerCP-Cy5.5, CD38-PECy7, IgM- APC (BD 551062), CD27-APC-H7, IgA-biotina (AbD Serotec 205008) seguido de Streptavidin-eFluor710 (eBioscience 49-4317-82) y CD19-BV421 (Biolegend 302233). Las células B de memoria también se han clasificado como CD19<sup>+</sup>CD27<sup>+</sup>IgG<sup>+</sup> o CD19<sup>+</sup>CD20<sup>+</sup>IgG<sup>+</sup>, las células B ingenuas se han seleccionado como CD19<sup>+</sup>IgD<sup>+</sup>IgM<sup>+</sup>. También se han clasificado los plasmoblastos IgA<sup>+</sup>, y se definen como CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>CD38<sup>+</sup>IgA<sup>+</sup>IgM<sup>+</sup>. También se pueden usar otros marcadores de superficie celular, siempre que la población de células B u otra célula celular sea identificable fenotípicamente utilizando marcadores de superficie celular, la población puede ser seleccionada con respecto a células individuales. Véase abajo. Las células se lavaron una vez con 2 mL de regulador FACS y se resuspendieron a un volumen apropiado para FACS. Las células se clasificaron primero en un BD Aria II en un tubo de fondo redondo de 5 mL. Típicamente, se lograron purezas de >80% desde la primera selección. Las células individuales se seleccionaron en las primeras 11 columnas de una placa de PCR de 96 pozos que contenía 6.65 µL de un regulador hipotónico (Tris-HCl 10 mM, pH 7.6) que contenía dNTP 2 mM (NEB N0447L), oligo 5 µM (dT)<sub>20</sub>VN y 1 unidad de Ribolock (Fermentas EO0384), un inhibidor de la AARNasa. Como control negativo, la última columna se dejó sin células. Para los plasmoblastos IgG<sup>+</sup>, la estrategia de selección (selección de células) fue CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>CD38<sup>+</sup>IgA<sup>+</sup>IgM<sup>+</sup>. Las placas clasificadas se sellaron con selladores de placas de aluminio (Axygen PCR-AS-600) y se congelaron inmediatamente en hielo seco y se almacenaron a -80°C.

#### Estrategias de selección por restricción de células individuales

55 **Células B.** Para las células B, la metodología de restricción comprende la selección de uno o más de los siguientes marcadores: IgM, IgG, IgA, IgD, CD19 o CD20. Para las células B de IgG<sup>+</sup> totales, la metodología de restricción comprende la selección de IgG<sup>+</sup>. Para las células B de IgA<sup>+</sup> totales, la metodología de restricción comprende la selección de IgA<sup>+</sup>. Para las células B de IgM<sup>+</sup> totales, la metodología de restricción comprende la selección de IgM<sup>+</sup>.

**Células B activadas.** Las células B activadas incluyen células B que se han estimulado mediante la unión de su receptor de antígeno de membrana a su antígeno afín y/o han recibido ayuda de células T de células T que reconocen epítopos derivados del mismo antígeno macromolecular. Las células B activadas se pueden identificar por una variedad de propiedades que incluyen un aumento del tamaño celular (por ejemplo, "Células B blásticas"; véase a continuación),

la expresión del marcador o marcadores de la superficie celular, la expresión del marcador o marcadores intracelulares, la expresión del factor o factores de transcripción, salir de la fase de brecha 0 (G0) del ciclo celular, que progresa a través del ciclo celular, la producción de citoquinas u otros factores, y/o la regulación descendente de ciertos marcadores o marcadores de la superficie celular, marcadores intracelulares o marcadores, factor de transcripción u otro factor. Un método para identificar una célula B activada es combinar la detección de un marcador de células B como CD19 o inmunoglobulina con un marcador de activación tal como un aumento del tamaño o volumen celular, el marcador de activación de la superficie celular CD69 o la progresión a través del ciclo celular con base en tinción de ADN de acridina naranja permeable a las células u otro análisis del ciclo celular.

*Células blásticas B.* "Células blásticas B" son células B que se activan y aumentan de tamaño en relación con las células B en reposo. Las células B blásticas incluyen la población de plasmoblastos, así como otras poblaciones de células B activadas, y las células B blásticas son físicamente de mayor tamaño que las células B en reposo. Las células B blásticas pueden seleccionarse con una sola célula utilizando varias metodologías diferentes, incluida la restricción (selección) de células B en función de su tamaño físico con base en el diámetro de la célula, el volumen de la célula, la impedancia eléctrica, el FSC, la integral (área) de un pulso FSC (FSC-A), FSC altura (FSC-H), ancho de pulso de dispersión de avance (FCS-W), dispersión lateral (SSC), área de pulso de dispersión lateral (SSC-A), altura de dispersión lateral (SSC-H), lado ancho de dispersión (SSC-W), autofluorescencia y/u otras mediciones del tamaño celular.

En la citometría de flujo, la dispersión de avance (FSC) se mide utilizando un haz de luz en línea con el flujo de células y proporciona información sobre el tamaño proporcional y el diámetro de cada célula. Usando FSC, se pueden seleccionar células B con FSC mayor que la mediana de FSC de células B en reposo, por ejemplo, un FSC-A o FSC-H 5% mayor que las células B en reposo, 10% mayor que las células B en reposo, 15% mayor que en reposo Células B, 20% más que las células B en reposo, 30% más que las células B en reposo, 40% más que las células B en reposo, 50% más que las células B en reposo, 60% más que las células B en reposo. Al analizar las perlas de calibración de tamaños específicos, se puede usar el FSC para determinar el tamaño relativo de las células B en relación con las perlas de calibración. Al hacerlo, se puede ingresar específicamente y así seleccionar células B que poseen diámetros de aproximadamente 8um, >8 um, >9 um, >10 um, >11 um, >12 um, >13 um, > 14 um, >15 um, > 16 um, >17 um, >18 um, >19 um, o >20 um.

Otra medida del tamaño celular es el volumen celular. El "estándar de oro" para el volumen celular utiliza el principio de Coulter que se basa en una medición electrónica (Tzuret al, PLoS ONE, 6(1):e16053. doi:10.1371/journal.pone.0016053, 2011). Aunque el método de clasificación por carga de gotas y deflexión se utilizó por primera vez en un dispositivo que medía el volumen celular por impedancia, los citómetros de flujo disponibles actualmente solo realizan mediciones ópticas. Las mediciones de FSC, específicamente el FSC-A (área integral de FSC) se usan comúnmente para evaluar el tamaño de la célula, aunque las mediciones de FSC pueden verse influidas por las diferencias del índice de refracción entre las partículas y el fluido (Tzur et al, PLoS ONE, 6(1): e16053. doi:10.1371/journal.pone.0016053, 2011). Algunos han demostrado que la estimación del volumen se puede mejorar combinando parámetros ópticos, incluidos FSC-W, SSC y 450/50-A fluorescencia automática (Tzur et al, PLoS ONE, 6(1): e16053. Doi: 10.1371/journal.pone.0016053, 2011).

Por ejemplo, la selección de células B activadas según el aumento de tamaño se puede lograr mediante la identificación de células B con un marcador como CD19 y la evaluación del tamaño a través de FSC o FSC-A. Otros marcadores de células B y/o parámetros para la evaluación del tamaño se describen aquí.

*Plasmoblastos.* Para el aislamiento de plasmoblastos, la metodología de restricción comprende la selección de células B CD19<sup>+</sup>CD38<sup>++</sup>. Para el aislamiento de los plasmoblastos IgG<sup>+</sup>, la metodología de restricción comprende la selección de las células B CD19<sup>+</sup>CD38<sup>++</sup>IgA-IgM<sup>-</sup>. Para el aislamiento de plasmoblastos IgA<sup>+</sup>, la metodología de restricción comprende la selección de células B CD19<sup>+</sup>CD38<sup>++</sup>IgA<sup>+</sup>. Para el aislamiento de plasmoblastos IgM<sup>+</sup>, la metodología de restricción comprende la selección de células B CDM<sup>+</sup>CD38<sup>++</sup>IgM<sup>+</sup>. Además, se pueden usar otras estrategias de activación para aislar un número suficiente de plasmoblastos para llevar a cabo los métodos descritos en este documento. Los plasmoblastos también se aislaron utilizando los siguientes patrones de expresión de marcadores CD19<sup>bajo/+</sup>, CD20<sup>bajo/-</sup>, CD27 y CD38<sup>++</sup>. Aunque el uso de todos estos marcadores generalmente resulta en la población de plasmoblastos más pura de la selección de células individuales, no es necesario usar todos los marcadores anteriores. Por ejemplo, los plasmoblastos también se pueden aislar utilizando las siguientes estrategias de activación: alta dispersión de avance (FSC<sup>hi</sup>) para células más grandes, células FSC<sup>hi</sup>CD19<sup>lo</sup>, células FSC<sup>hi</sup> y CD27<sup>+</sup>, CD38<sup>++</sup>, o CD20<sup>-</sup>. La combinación de cualquiera de estos marcadores u otros marcadores que puedan distinguir los plasmoblastos de otras células B generalmente aumentará la pureza de los plasmoblastos clasificados, sin embargo, cualquiera de los marcadores anteriores (incluido FSC<sup>hi</sup>) puede distinguir los plasmoblastos de otros linfocitos B, aunque con menor pureza.

*Para las células B de memoria.* Para las células B de memoria IgG<sup>+</sup>, la metodología de restricción comprende la selección de CD19<sup>+</sup> CD27<sup>+</sup>IgG<sup>+</sup> o CD19<sup>+</sup> CD20<sup>+</sup>IgG<sup>+</sup>. Para las células B de memoria IgA<sup>+</sup>, la estrategia de activación comprende CD19<sup>+</sup>CD27<sup>+</sup>IgA<sup>+</sup> o CD19<sup>+</sup>CD20<sup>+</sup>IgA<sup>+</sup>. Para las células B de memoria IgM<sup>+</sup>, la estrategia de activación comprende CD19<sup>+</sup>CD27<sup>+</sup>IgM<sup>+</sup> o CD19<sup>+</sup>CD20<sup>+</sup>IgM<sup>+</sup>.

*Para otros tipos de células.* Siempre que la población de células B, células T, u otra célula celular sea identificable fenotípicamente utilizando marcadores de células, puede clasificarse en una sola célula. Por ejemplo, las células-T pueden identificarse como CD3<sup>+</sup> o TCR<sup>+</sup>, las células-T no identificadas como CD3<sup>+</sup>CD45RA<sup>+</sup>, las células-T de memoria identificadas como CD3<sup>+</sup>CD45RO<sup>+</sup>. Las células-T efectoras y auxiliares pueden identificarse como CD3<sup>+</sup> CD8<sup>+</sup> y CD3<sup>+</sup>CD4<sup>+</sup>, respectivamente. Las poblaciones celulares pueden subdividirse aún más mediante el uso de combinaciones de marcadores, como CD3<sup>+</sup>CD4<sup>+</sup>CD45RO<sup>+</sup> para las células-T auxiliares de memoria.

#### *Secuenciación de genes de inmunoglobulina de cadenas liviana y pesada apareadas de células B individuales*

Transcripción reversa con moléculas adaptadoras.

- Las placas clasificadas de una sola célula se descongelaron en hielo y se centrifugaron brevemente antes de su uso. Las placas se incubaron en el termociclador a 55°C durante 3 minutos, 42°C durante 2 minutos, y de forma indefinida a 4°C. Las placas se volvieron a centrifugar brevemente y se abrieron cuidadosamente para evitar la formación de aerosoles. Se añadió 1 µL de una solución 10 µM de la molécula adaptadora apropiada (cada molécula adaptadora generalmente tiene una región de identificación de la muestra (ID de la muestra)) se agregó a cada pozo, con todos los pozos de control negativo (que contiene ARN regulador conservante solo, o células-no-B) recibiendo moléculas adaptadoras idénticas. 2.35 µL de una mezcla que contiene 0.75 µL de H<sub>2</sub>O, 1 µL de 10x M-MuLV RT buffer (NEB B0253S), 0.6 µL de 50 mM MgCl<sub>2</sub>, 0.25 µL de Ribolock (40 U/µL) y 0.125 µL de Superscript III (200 U/µL) (Invitrogen 18080-085) se añadió y se mezcló mediante pipeteo. Las placas se centrifugaron brevemente y se incubaron a 42°C durante 120 minutos a 8 horas usando un agitador térmico de placas y luego se mantuvieron a -20°C. Después de la reacción, los productos de RT de todos los pozos se agruparon en un tubo de microcentrífuga. Los productos de RT reunidos se extrajeron luego con fenolcloroformo-alcohol isopropílico con ~0,1% de 8-hidroxiclороquina (Sigma 77617) y luego se extrajeron por extracción con cloroformo en tubos de fase de bloqueo de gel (5 PRIME 2302820). Los productos de RT luego se concentraron y se desalaron en giros de 5 minutos a 14 000 g con Amicon Ultra-0.5 30kDa (Millipore UFC503096) o Ultra-0.5 100kDa (Millipore UFC510096), seguido de un giro de 5 minutos a 14 000 g con TE (10 mM Tris HCl pH 7.6 con EDTA 1 mM) y un centrifugado final de 5 minutos a 14 000 g con EB (Qiagen 19086). Los productos de RT se eluyeron invirtiendo la columna Amicon Ultra en un nuevo tubo de centrífuga y centrifugando a 1000 g durante 2 minutos. En este punto, los productos de RT se mantuvieron a -20°C o -80°C.

Touchdown PCR

- Para 454 ejecuciones de secuencia 1 y 2, se usó el método de Touchdown PCR de la siguiente manera. Para algunas muestras en las pruebas de PCR 3 y 4, se cambió el método de PCR, lo que llevó a un mayor número de cadenas livianas y pesadas emparejadas. Este cambio se detalla en la subsección "PCR no Touchdown" más adelante.

- Tanto para la 1ª PCR como para la 2ª PCR anidada, se utilizó la ADN polimerasa Phusion Hot Start II (NEB F-549L) en el regulador GC proporcionado. Para IgG, los cebadores y las moléculas adaptadoras se muestran en la Tabla 1. Las secuencias de ID de muestra se muestran en la Tabla 2. Las secuencias de ID de placa se muestran en la Tabla 3. *Consúltese* también las Figuras 3 y 9. Las condiciones de reacción incluyeron una concentración final de MgCl<sub>2</sub> de 1.8 mM, 200 µM de dNTP, 0.2 µM para todos los cebadores, 0.2 U de polimerasa de Phusion, cantidades variables de DMSO como aditivo y 2 µL de plantilla en un volumen final de 25 µL. Para la 1ª PCR, las cadenas livianas lambda y kappa y la cadena pesada gamma se amplificaron en diferentes pozos, y el DMSO se usa a una concentración final del 8%, 5% y 10%, respectivamente. Los cebadores directos utilizados en la 1ª PCR fueron el cebador largo FW 1 y el cebador corto 1 FW. Debido a que el cebador 1 largo de FW agregó una región de identificación de la placa (placa-ID) al extremo 5' de las regiones del amplicón (amplicones), el cebador largo de FW que contiene diferentes ID de la placa se agregó a diferentes muestras. Se usaron cebadores reversos específicos de gen para amplificar las cadenas kappa, lambda y gamma, kappa GSP1, lambda GSP1 y gamma GSP1, respectivamente. Las condiciones del ciclo para la primera PCR incluyeron un paso de desnaturalización inicial a 98°C durante 30", seguido de 2 ciclos de 98°C durante 10", 72°C durante 25"; 7 ciclos de toma de contacto de 98°C durante 10", 71.5°C a 68.5°C durante 15" y 72°C durante 20" con una caída de 0.5°C en cada etapa de fusión subsiguiente; 30 ciclos de 98°C durante 10", 68°C durante 15" y 72°C durante 20", seguidos de una extensión final a 72°C durante 5' y mantenidos a 4°C de forma indefinida. Los productos de 1ª PCR se diluyeron 100x en TE y se utilizaron 2 µL para la 2ª PCR anidada. Para la 2ª PCR, se usó un 5% de DMSO como aditivo en todas las muestras. El cebador delantero es el cebador FW 2 y los cebadores reversos fueron el cebador RV 2 y el cebador largo GSP 2. Kappa GSP largo cebador2, lambda GSP cebador largo 2 y gamma cebador largo 2 se utilizaron para amplificar sus respectivos amplicones. Debido a que el GSP cebador largo 2 también agregó la placa-ID al extremo 3' de los amplicones, un cebador largo diferente al GSP2 con placa específica para la placa. Se agregaron ID a cada muestra de placa combinada. Las condiciones de los ciclos para la segunda PCR anidada incluyeron un paso inicial de desnaturalización de 98°C durante 30", 30-40 ciclos de 98°C durante 10", 67°C durante 15", y 72°C durante 20", seguido de una extensión final de 72°C durante 5' y mantener a 4°C por tiempo indefinido.

PCR no Touchdown

Para la PCR no Touchdown, las condiciones fueron idénticas a la PCR Touchdown, a menos que se indique lo contrario. Los primeros parámetros del ciclo de PCR fueron una desnaturalización inicial de 95°C durante 5', 15-25



## ES 2 716 013 T3

ciclos de 98°C 30", 62°C 30", 72°C 30", una extensión final de 72°C 5' y se mantuvo a 4°C por tiempo indefinido. La 1ª PCR fue una PCR multiplex, donde los 3 cebadores reversos específicos del gen, los cebadores reversos de las regiones kappa, lambda y gamma se utilizaron junto con 0.2, 0.2 y 0.24 µM, respectivamente. Los cebadores utilizados fueron los mismos que en la PCR Touchdown. Los cebadores específicos del gen pueden ser los utilizados en la PCR Touchdown y también cualquiera de los designados como adecuados para la primera PCR (Tabla 6). El DMSO se utilizó a una concentración final del 5%; También se pueden agregar 0.1 mg/ml de BSA (NEB B9001S) y ET-SSB (NEB M2401S) a 1:100 para la reacción de PCR. Durante la 1ª PCR, se usaron 4-6 µl de la plantilla de ADNc en un total de 80 o Volumen de reacción de 90 µl. Cada reacción de PCR1 se dividió en ocho o nueve reacciones de 10 µl, cada una de las cuales se produjo en un pozo diferente. La primera PCR se agrupó nuevamente después de la PCR y se diluyó 100x en TE0.1, y se utilizaron 2 µl para la 2ª PCR. La 2ª PCR es una reacción separada para cada cebador específico del gen (no multiplex), y la mezcla de reacción fue idéntica a la 2ª PCR de aterrizaje, excepto por lo siguiente: cualquiera de los cebadores de región constante específicos del gen designados como que funcionan para la 2ª PCR puede (Tabla 6), los cebadores se utilizaron a 0.2 µM o 0.4 µM en todo el tiempo, o los cebadores específicos del gen se utilizaron a 0.2 µM y el resto se utilizó a 0.4 µM. Se añadió 0.1 mg/ml de BSA a la reacción y ET-SSB también se puede usar a 1:100. Los parámetros de ciclo de la 2ª PCR semianidados fueron una desnaturalización inicial de 95°C durante 5', 20-35 ciclos de 98°C 30", 67°C 30", 72°C 30", una extensión final de 72°C 5' y se mantienen a 4°C por tiempo indefinido. El número total de ciclos de PCR para la 1ra y 2ª PCR combinados fue típicamente entre 50-60 ciclos para la PCR sin toma de contacto. Dado que los diferentes pozos agrupados que se someten al ciclo de la PCR tienden a usar un número diferente de ciclos para obtener una cantidad razonable de producto de ADN (típicamente entre 1-12 ng/ul), se llevaron a cabo 4 ciclos de PCR diferentes para cada 2ª PCR, por ejemplo, ciclos de 23, 26, 30 y 33, 5µl desarrollados en un gel de agarosa al 2% y En base al juicio cualitativo de la cantidad de producto de PCR, solo se usó el producto de PCR de uno de los números de ciclos de la 2ª PCR para cada 2ª PCR de pozos agrupados en la preparación para la secuencia de secuencia 454.

Para la PCR de otras cadenas pesadas de inmunoglobulina en humanos, las cadenas pesadas y livianas de inmunoglobulina en ratones y las cadenas de TCR en humanos y ratones, las condiciones de PCR son idénticas a las de la sección de PCR no Touchdown, excepto que la primera PCR no es multiplexada, siendo cada ADNc Amplificado individualmente. Los siguientes cebadores en 3' en las Tablas 10 y 11 se usan en PCR1 y 2.

### Preparación para la ejecución de la secuenciación 454 de XLR70

Para la 1ª y la 2ª serie de 454, se agregaron cebadores de secuenciación (cebadores de titanio A y B, respectivamente) para una serie de secuenciación 454 de titanio en los amplicones durante la 1ª y la 2ª PCR anidadas. Se corrieron 5 µL de cada amplicón en un gel de agarosa con una escalera de ADN en masa (Fermentas SM0383), se tomó una imagen y se analizaron las intensidades de la banda y se cuantificaron con el software AlphaFC Imager (Cell Biosciences). Se agruparon por separado 5 ng de cada uno de los amplicones kappa, lambda y gamma, se ejecutaron en un gel de agarosa al 0.8% y se visualizaron con GelGreen (Biotium 41005). Las bandas de los tamaños apropiados (~600 pb para kappa y lambda, y ~750 pb para gamma) se cortaron y se purificaron con el kit de Extracción de Gel MinElute (Qiagen 28606), según las instrucciones del fabricante con livianas modificaciones. Brevemente, el gel de agarosa se fundió en regulador QG sin calentamiento, y se realizó la etapa de lavado de QG adicional. El regulador de lavado de PE se dejó reposar durante 5 minutos antes de girar. También se realizó una etapa de lavado de PE adicional. Las muestras se eluyeron con 25 µL de regulador EB. Las muestras también se limpiaron una vez con perlas SPRI utilizando una proporción de 1:0.65 para el volumen de ADN: volumen de perlas para 454 2ª ejecución. La concentración de ADN se determinó con el kit de ensayo de ADN Picogreen (Invitrogen P11496) y las muestras se agruparon de manera que la concentración de ADN de gamma: kappa: lambda es 2:1:1. Las muestras agrupadas se encontraban en una concentración de >0.5ng/µL y se enviaron a un centro de secuenciación de ADN 454 para la secuenciación 454.

Para la 3ª y posteriores secuencias de 454, se cambió el protocolo. Los amplicones todavía se agruparon por separado para normalizar las cantidades de ADN de cada reacción de PCR, pero primero se sometieron a una limpieza de perlas SPRI para eliminar pequeños fragmentos de ADN de acuerdo con las instrucciones del fabricante. Los amplicones se ejecutaron en un gel de agarosa al 3%, y las bandas apropiadas se cortaron y purificaron con el kit de extracción de gel MinElute como antes. A partir de entonces, los amplicones se sometieron a otras 2 rondas de limpieza de perlas SPRI para eliminar aún más pequeños fragmentos de ADN, y se cuantificaron con Picogreen, se verificó la calidad con Nanodrop para asegurar que la relación OD260/280 fuera >1.8 y 1µl se ejecutó en un gel para asegurar que no hubiera Pequeños fragmentos de ADN. Los amplicones Lambda y kappa se agruparon en una proporción de 1:1, y se usó gamma tal como está. Luego, el ADN se diluyó a 1 x 10<sup>9</sup> copias según las instrucciones de 454, y se envió al centro de secuenciación (Roche) para emPCR a 1 cpb y se secuenció; cadena pesada gamma en una región y las cadenas livianas agrupadas en la otra región de la placa de picotitulación.

### Preparación para la ejecución de la secuenciación de 454 XL+

Actualmente, la ejecución de secuenciación 454 XL+ no es compatible con el kit de secuenciación Lib-A que se usó para la ejecución XLR70. XL+ actualmente solo admite el kit Lib-L, que es una secuenciación unidireccional. Para adaptar nuestro protocolo para hacer la secuenciación XL+, se sigue el protocolo para la ejecución de XLR70, pero después de la etapa de limpieza del gel, cada amplicón (kappa, lambda y gamma) se sometió a 2 PCR por separado,

5 cada una de 5 ciclos para agregar el Lib-L A y adaptadores B. Las condiciones de la PCR son las siguientes: se usa la polimerasa de fusión, con 5x de regulador GC y una concentración final de 5% de DMSO. Los cebadores se utilizan a 0.2uM. Se añaden 0.1 mg/ml de BSA a la reacción. Los parámetros del ciclo de PCR son una desnaturalización inicial de 95°C durante 5', 20-35 ciclos de 98°C 30", 67°C 30", 72°C 30", una extensión final de 72°C 5' y mantener a 4°C de forma indefinida. Se realizan dos PCR para cada amplicón: 5LIB-LA y 3LIB-LB en una PCR, y 5LIB-LB y 3LIB-LA en la otra PCR. Los adaptadores se agregan de manera que cada amplicón se convierte en 5' -LibA-amplicon-LibB-3' o 5'-LibB-amplicon-LibA-3'. Estos amplicones tienen los adaptadores LibA "A" o "B" en el extremo 5' (y la correspondiente "B" o "A" adaptador en el extremo 3'), que permite la secuenciación bidireccional. Los amplicones con los nuevos adaptadores Lib A luego se someten a 3 rondas de limpieza de perlas SPRI antes de seguir el protocolo para XLR70 para cuantificar y verificar la calidad del ADN antes de diluirlo a  $1 \times 10^9$  copias y envío a un servicio de secuenciación 454 (Roche) para emPCR a 1cpb y secuenciación.

#### Preparación para la ejecución de la secuenciación de PacBio

15 Para la ejecución de la secuenciación de PacBio, se empleó la PCR Touchdown como se indicó anteriormente. La agrupación de ADN y la limpieza se realizaron como en la sección anterior sobre "preparación para 454 XLR70". Para obtener suficiente ADN (500 ng) para los requisitos de secuenciación, se agrupó un mínimo de 1 µg de ADN para la limpieza con gel y SPRI. La cuantificación de Picogreen y las diluciones  $1 \times 10^9$  no se realizaron, ya que no fue necesario para la secuenciación de PacBio. Si no se obtuvo ADN suficiente de la 2.<sup>a</sup> PCR, se repitieron la 2.<sup>a</sup> PCR y los pasos de agrupación hasta que se obtuvo suficiente ADN. Se envió un mínimo de 500 ng de ADN limpiado a las instalaciones de secuenciación de PacBio para la secuenciación.

#### 20 Otras metodologías de secuenciación

Los métodos descritos en este documento no dependen de la secuenciación 454 o PacBio. Las cadenas livianas Lambda y kappa son de ~600 pb y la cadena pesada gamma es de ~700 pb. Por lo tanto, lo que generalmente se desea es la capacidad de tener lecturas de secuencia más largas de modo que las lecturas de secuencia directa e inversa se superpongan lo suficiente como para permitir la reconstrucción de la secuencia completa, de aproximadamente 600 pb, de las cadenas livianas (LC) (la longitud exacta de la secuencia depende de la longitud de la región 5' no traducida (UTR)), y aproximadamente 700 pb de la secuencia de las cadenas pesadas (HC). Por lo tanto, se puede utilizar cualquier tecnología de secuenciación que pueda producir lecturas de secuenciación de al menos aproximadamente 350-400 pb y, de ese modo, lograr la superposición utilizada para el ensamblaje de secuencias, y tecnologías de secuenciación que permiten la aplicación. Las lecturas de 600-700+ pb permitirían una secuencia utilizando solo el cebador directo (Fw) (secuenciación desde el extremo 5').

#### Secuencias

Los datos de secuencia para las ejecuciones anteriores se recibieron de la instalación relevante y se procesaron como se describe a continuación.

#### Nomenclatura de Secuencias

35 Cada secuencia en la lista de secuencias que corresponde a una lectura de secuenciación, ensamblaje de secuencia o traducción de aminoácidos de una secuencia tiene un identificador. Cada uno de estos identificadores tiene 9 campos separados por un punto, ".". Los campos están numerados del 1 al 9 y proporcionan la siguiente información:

1. ID de lectura. Una ID de lectura asignada por el software asociado con la tecnología de secuenciación utilizada para determinar la lectura, o "NA" si la secuencia no es una lectura en bruto.
- 40 2. Número de placa. Un número de placa con el que está asociada la secuencia. Consúltese la Tabla 12 (tabla de mapeo de placa a muestra) para obtener la información correspondiente de la muestra biológica.
3. ID de muestra. ID de muestra que indica el pozo con el que está asociada la secuencia. Los números de identificación de la muestra están entre 1 y 89 inclusive. *Consúltese* la Tabla 2 para la correspondencia entre la identificación de la muestra y el nombre del pozo.
- 45 4. Nombre de pozo. Nombre de pozo que contiene el pozo con el que está asociada la secuencia. El nombre del pozo corresponde al nombre habitual de la placa de 96 pozos, por ejemplo D07. El nombre del pozo y la identificación de la muestra son formas equivalentes de especificar un pozo particular en una placa.
5. Cónfigo ID. El ID de cónfigo distingue diferentes secuencias asociadas con un pozo de un ensamblaje y tipo de cadena dados.
- 50 6. Plataforma. El campo de la plataforma indica la tecnología de secuenciación de la que se deriva la secuencia. Los valores posibles para la plataforma son 454, Sanger y PacBio.

7. Tipo de cadena. El campo de tipo de cadena indica si la secuencia está asociada con un conjunto de secuencias de anticuerpos de cadena pesada, secuencias de anticuerpos de cadena liviana, o un conjunto que contiene secuencias de anticuerpos de cadena tanto pesada como liviana. Los valores posibles son "pesada", "liviana" o "CMB".

5 8. ID de ejecución. Un identificador para un conjunto de lecturas en una plataforma particular.

9. Tipo de secuencia. El tipo de la secuencia. Los valores posibles son "en bruto" para lecturas en bruto de la tecnología de secuenciación, "nb", "urt", "multim50", "zerom50" o "pb" para lecturas ensambladas (véase la sección de Ensamblaje de Secuencias.), o "nb-aa", "urt-aa", "multim50-aa", "zerom50-aa" o "sanger-aa" para secuencias de aminoácidos derivadas de varias secuencias de consenso de ensamblaje nt.

10 Preparación de secuencias para análisis.

Los datos generados a partir de la secuenciación 454 se analizaron mediante el software de análisis de datos GS FLX 454 y se retornaron secuencias de alta calidad aprobadas por el filtro. Debido a la restricción del filtro de amplicón predeterminado utilizado por el software de análisis de datos 454 GS FLX, es posible que la restricción del filtro deba relajarse para obtener lecturas largas suficientes. Una forma es seguir las sugerencias en 454 boletín técnico APP No. 001-2010. Cambiar <vfScanAllFlows> de "TiOnly" a "False" del filtro de amplicón puede provocar un gran aumento en las secuencias de filtro de buena calidad aprobadas. Otra opción es cambiar el <vfTrimBackScaleFactor> a un número más bajo. Para la Ejecutar 1 de 454, se utilizó el procesamiento de disparo estándar y para la Ejecutar 2, <vfScanAllFlows> se cambió a "Falso", y se utilizó el procesamiento de la tubería de amplicón estándar para las ejecuciones 3 y 4.

20 Los datos generados a partir de la secuenciación de Pacific Biosciences se recibieron de Pacific Biosciences a la vez que la Secuencia de Consenso Circular se lee con los puntajes de calidad asociados.

#### Asignación de Secuencias a Pozos

El ADNc de las muestras se secuenció con la tecnología de secuenciación 454 o Pacific Biosciences. Las lecturas son las del Listado de Secuencias cuyo Tipo de Secuencia están "sin procesar". Las lecturas de secuenciación se analizaron y se asignaron a una placa y pozo fuente y se descartaron.

Las asignaciones de placas y pozos para las lecturas se realizaron comparando la secuencia de lectura observada con la posible región de identificación de la placa, la región del cebador universal y las secuencias de la región de identificación de la muestra utilizando expresiones regulares. La comparación se realizó en tres etapas utilizando las expresiones regulares enumeradas en las Tablas 13, 14 y 15.

30 En la etapa 1, análisis de posibles regiones de identificación de la placa, se verificó una lectura con todas las expresiones regulares enumeradas en la columna "Expresión regular de la región de identificación de la placa" en la Tabla 13, lo que requiere una coincidencia para comenzar con el primer nucleótido de la secuencia. Si no se encontró ninguna coincidencia, la lectura se descartó y la asignación de placa/pozo continuó con la siguiente lectura disponible para procesar, si la hubiera. Si se encontró una coincidencia, se asignó a la secuencia la ID correspondiente de la columna "ID de Placa" como su ID de placa. Los nucleótidos de la lectura que coinciden con la expresión regular de la placa se registraron para su uso durante las etapas posteriores de la coincidencia y durante el ensamblaje.

40 En la etapa 2, análisis de la región del cebador universal, se comparó una lectura con la "expresión regular del cebador universal" "CACGACCGGTGCTCGATT+AG", que requiere una coincidencia de lectura para comenzar con el primer nucleótido después de la última concordancia del nucleótido leído con la Expresión Regular de la Placa. Si la lectura no coincide con la expresión regular del cebador universal, la lectura se descartó y la asignación de placa/ pozo continuó con la siguiente lectura disponible para procesar, si la hubiera. De lo contrario, los nucleótidos de la lectura que coinciden con la expresión regular del cebador universal se registraron para su uso durante la última etapa de coincidencia y durante el ensamblaje.

45 En la etapa 3, análisis de posibles regiones de identificación de la muestra, se verificó una lectura frente a todas las expresiones regulares enumeradas en la columna "Expresión Regular de la Región de Identificación de la Muestra" en la Tabla 14, lo que requiere una coincidencia para comenzar con el primer nucleótido después de la última comparación de nucleótidos de lectura a la expresión regular del cebador universal. Si no se encontró ninguna coincidencia, la lectura se descartó y la asignación de placa/pozo continuó con la siguiente lectura disponible para procesar, si la hubiera. Si se encontró una coincidencia y la columna de ID de muestra contenía solo un único identificador, la ID de muestra de la lectura se asignó como la ID que se encuentra en la columna de ID de muestra. Si la columna de ID de muestra contenía más de un solo identificador, esos identificadores se consideraron "ID de muestra candidatas". Luego, la lectura se verificó de forma secuencial con todas las expresiones regulares enumeradas en la columna "Expresión regular de la región de identificación de muestra" de la Tabla 15, donde al menos una de las ID de muestra correspondientes en la columna "ID de muestra" de la Tabla 15 coincidió con una ID de muestra candidata. Si la lectura coincidía con la expresión regular, y la coincidencia comenzó con el primer

nucleótido después de la última lectura de nucleótidos coincidente con la expresión regular del cebador universal, el identificador más a la derecha de los ID de muestra candidatos se asignó como el ID de muestra de la lectura. De lo contrario, el identificador más a la derecha se eliminó de la lista de ID de muestras candidatas y el proceso se repitió con la lista más pequeña de ID de muestras candidatas hasta que se encontró una coincidencia, o, si no se encontró una expresión regular coincidente en la lista de expresiones regulares en la Tabla 15, a continuación, se asignó la última ID de muestra candidata (es decir, la que está más a la izquierda en la lista original de ID de muestra candidata) como la ID de muestra para la lectura.

Las lecturas que se descartaron durante la ID de placa y el proceso de asignación de ID de muestra no se incluyeron en la lista de secuencias.

## 10 Ensamblaje de Secuencias

Todas las lecturas de secuencia asignadas a una ID de muestra asociada con un pozo se ensamblaron para producir secuencias de consenso. Estas secuencias de consenso corresponden a las secuencias de ARNm de cadena pesada y liviana expresadas en las células clasificadas.

15 Las secuencias se ensamblaron con Newbler 2.5 (Ejecuciones 1 y 2), y Newbler versión 2.6 y/o Mira versión 3.4.0 para otras secuencias.

Las secuencias en el listado con un campo de Plataforma de "454", un campo de Tipo de Cadena de "mixto", un ID de Ejecución de "1" o "2" y un Tipo de Secuencia de "nb" son contiguos resultantes de un conjunto que utiliza newbler. Para ensamblar estas secuencias, se leyeron en Python los archivos de salida sff de la secuenciación 454, que contienen secuencias y puntuaciones de calidad para cada nucleótido, utilizando el paquete Biopython y las secuencias se subdividen de acuerdo con sus códigos de barras compuestos (-ID de muestra + placa-ID) como se describe arriba y la salida en archivos sff separados. Luego, estos archivos se volvieron a analizar mediante sffile (proporcionado por el software de análisis de datos GS FLX) en archivos sff con encabezados de archivos entendidos por Newbler, un ensamblador de secuencias proporcionado en el paquete de software de análisis de datos GS FLX, usando "-force", "-ADNc" y opciones "-urt". Newbler luego reunió las lecturas avanzadas con códigos de barras compuestos compartidos. Debido a que las lecturas reversas tienen solo un ID de placa 3', es posible que se produzca un ensamblaje de secuencias entre las lecturas de secuencias de diferentes células de avance y de retroceso. Para evitar este problema potencial, el uso de V(D)J de la cadena pesada y liviana de las lecturas reversas ensambladas tanto de avance como hacia abajo se puede identificar primero utilizando HighV-QUEST (<http://imgt.cines.fr/HighV-QUEST/index.action>). Luego, las secuencias se pueden agrupar según su uso de V(D)J antes de volver a ensamblarse con Newbler utilizando una lectura adelantada y las lecturas reversas que comparten el mismo uso de V(D)J. Esto se puede repetir para todas las lecturas avanzadas ensambladas. El ensamblaje de la secuencia también se puede hacer para ser intolerante con los desajustes de nucleótidos, evitando así el ensamblaje de lecturas directas e reversas de diferentes células que comparten el mismo uso de V(D)J. De esta manera, se puede evitar en gran medida el ensamblaje de secuencias inapropiadas de lecturas reversas entre secuencias altamente similares de diferentes células.

Las secuencias en el listado con un campo de Plataforma de "454", un campo de Tipo de Cadena de "pesado" o "ligero", un ID de Ejecución de "3" o "4" y un Tipo de Secuencia de "nb" son contiguos resultantes de ensamblajes de lecturas en 454, ejecutando Newbler con esta línea de comando: `runAssembly -ADNc -o salida seqs.fastaq` donde `seqs.fastaq` contenía lecturas recortadas de un solo pozo en formato FastQ.

40 Todos los pozos para las lecturas ID de Ejecutar 3 o ID de Ejecutar 4 que Newbler no creó exactamente un contig de cadena pesada o exactamente un contig de cadena liviana se volvieron a analizar al ensamblarlas con Mira. Las secuencias en el listado con un campo de Plataforma de "454", un campo de Tipo de Cadena de "pesada" o "liviana", un ID de Ejecución de "3" o "4" y un Tipo de Secuencia de "multim50" o "zerom50" son contiguos resultantes de estos ensamblajes, ejecutando Mira con esta línea de comando: `mira --project=seqs-job=denovo,est,precision, 454 454_SETTINGS -ED:as=yes -AL:egp = no -CL:pvlc=yes--fastq - notraceinfo`

Un archivo llamado `seqs_in.454.fastq` contenía lecturas recortadas de un solo pozo en formato FastQ.

Para los pozos de las lecturas de ID de Ejecutar 3 o ID de Ejecutar 4 donde ni Newbler ni Mira crearon contiguos utilizando los comandos de ensamblaje anteriores, se ejecutó un comando de Newbler diferente. Las secuencias en el listado con un campo de Plataforma de "454", un campo de Tipo de Cadena de "pesada" o "ligera", un ID de Ejecución de "3" o "4" y un Tipo de Secuencia de "urt" son contiguos resultantes de estos ensamblajes, donde Newbler se ejecutó con esta línea de comando: `runAssembly -ADNc -ud -urt -o salida seqs.fastaq` donde `seqs.fastaq` contenía lecturas recortadas de un solo pozo en formato FastQ.

55 Las secuencias en el listado con un campo de Plataforma de "PacBio", un Tipo de Cadena de "pesada" y un Tipo de Secuencia de "pb" son contiguos resultantes de ensamblajes de lecturas de la plataforma de PacBio, ejecutando Mira con la línea de comando: `mira --proyecto=seqs-job=denovo,est,exacto,454 454_SETTINGS -ED:ace=yes -AL:egp=no -CL:pvlc=yes --fastq - notraceinfo`

Un archivo llamado seqs\_in.454.fastq contenía lecturas recortadas de un solo pozo en formato FastQ.

#### Secuencias de aminoácidos

5 Secuencias en el listado con un Campo de Plataforma de "454" y un Tipo de Secuencia de "nb-aa", "urt-aa", "multim50-aa" o "zerom50-aa" son secuencias de aminoácidos determinadas mediante la traducción de las secuencias de nucleótidos de ensamblajes de lecturas en 454 como se describe en "Ensamblaje de Secuencias".

Las secuencias en el listado con un Campo de Plataforma de "PacBio" y un Tipo de Secuencia de "pb-aa" son secuencias de aminoácidos determinadas mediante la traducción de las secuencias de nucleótidos de los ensamblajes de Pacific Biosciences como se describe en "Ensamblaje de Secuencias".

10 Las secuencias en el listado con un Tipo de Secuencia de "sanger-aa" son secuencias de aminoácidos determinadas mediante la traducción directa de lecturas determinadas por la secuenciación de Sanger.

#### Otras opciones de análisis de datos de secuenciación.

El flujo de trabajo del análisis de datos descrito anteriormente se puede usar para determinar con precisión las secuencias de las cadenas pesada y liviana de cada célula. Sin embargo, esta información no es absolutamente necesaria para nuestra metodología de "criba de selección" (véase "Selección de anticuerpos humanos expresados").

15 Para que la criba de selección funcione, primero se agruparon las secuencias de anticuerpos pareados en familias clonales, en función de su uso de V(D)J de cadena pesada y el uso de VJ de cadena liviana. Por lo tanto, no se requirió la secuencia completa de las cadenas pesada y liviana de inmunoglobulina, y se puede usar suficiente información de secuencia para determinar el uso de V(D)J. Por lo tanto, se pueden tolerar errores de secuenciación y se pueden usar lecturas de menor calidad generadas por 454 o cualquier otra tecnología de secuenciación. El ensamblaje de

20 secuencias de la lectura directa no es generalmente un problema porque todas las secuencias se pueden agrupar primero de acuerdo con su código de barras compuesto antes de ensamblarse. Debido a que cada código de barras compuesto proviene de una muestra/célula, solo hay una secuencia "correcta" para cada cadena de inmunoglobulina con el mismo código de barras compuesto. Los errores de secuencia en diferentes cadenas pueden promediarse porque es poco probable que todos los errores de secuencia se produzcan en las mismas bases, lo que significa que

25 tomar la secuencia de base de consenso proporcionará la secuencia más precisa. En casos de ambigüedad, generalmente se elegirán las bases con altos puntajes de calidad de Phred. Debido a que 454 recorta lecturas de secuencia desde el extremo 3' hasta que solo quedan lecturas de mayor calidad, esto puede dar como resultado lecturas muy cortas. Con nuestro método, se pueden tolerar lecturas de menor calidad y, por lo tanto, utilizar lecturas mucho más largas generadas por 454 (400-500pb). Con estas lecturas más largas, se pueden identificar el uso de

30 V(D)J sin requerir el ensamblaje de la lectura directa con la lectura reversa, lo que hace que la identificación de la placa 3' no sea esencial en algunos aspectos. Además, la última generación de secuencias en 454 puede secuenciar hasta una media de 746pb y un modo de 800pb. Por lo tanto, la secuenciación de solo las lecturas avanzadas puede ser suficiente para cubrir la totalidad de los amplicones de inmunoglobulina de cadena pesada y liviana, lo que también

35 hace que la identificación de la placa 3' no sea esencial, en algunos aspectos, ya que no es necesario unir las lecturas reversas.

#### Selección y clonación de anticuerpos.

Después del ensamblaje, se analizaron las secuencias de las cadenas pesada y liviana para seleccionar los anticuerpos para la caracterización. Los anticuerpos se seleccionaron en función del uso previsto de la línea germinal V(D)J y la inspección de árboles evolutivos derivados de las secuencias de anticuerpos. Los anticuerpos seleccionados

40 se clonaron, expresaron y caracterizaron en diferentes ensayos.

#### Asignación de V(D)J

Las secuencias de las cadenas pesada y liviana de los ensayos 1 y 2 se analizaron con V-QUEST (Brochet, X. y otros, Nucl. Acids Res. 36, W503-508 (2008)), un software que compara una secuencia de cadena de anticuerpos con una base de datos de los alelos conocidos de las secuencias de la línea germinal y predice los alelos de la línea germinal

45 utilizados en el anticuerpo, cómo se recombinaron las secuencias de la línea germinal y las mutaciones en el anticuerpo en relación con la línea germinal. La Tabla 18 muestra los resultados de una asignación V-QUEST V (D)J para los anticuerpos que se eligieron para una caracterización adicional; los mismos datos se obtuvieron para todos los demás ensamblajes de secuencia de Ejecutar 1 y de Ejecutar 2 también. Algunas secuencias de los ensamblajes de Ejecutar 3 y Ejecutar 4 se analizaron con SoDA (Volpe, Cowell y Kepler, Bioinformatics (2006) 22 (4): 438-444), software similar a V-QUEST.

50

Si el genoma del paciente ha sido secuenciado, los datos de la secuencia del genoma se pueden usar como secuencias de la línea germinal para el análisis de asignación de VDJ, mejorando aún más la capacidad de identificar de manera confiable las hipermutaciones somáticas en las secuencias de anticuerpos del paciente.

#### Árboles evolutivos y familias clonales

Las secuencias de nucleótidos correspondientes a los péptidos maduros de las cadenas pesadas de los ensayos 1 y 2 se separaron en conjuntos que correspondían al paciente del que se derivaron. A partir de estos conjuntos individuales, el software clustalx2 (Larkin MA, Blackshields G, Brown NP, Chenna R, McGettigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R, Thompson JD, Gibson TJ, Higgins DG. (2007) *Bioinformatics*, 23, 2947-2948) se utilizó para generar una alineación y un árbol utilizando la configuración predeterminada para todos los parámetros.

Los árboles evolutivos de secuencias de un paciente también pueden construirse a partir del conjunto de secuencias de familias clonales individuales. Las secuencias de cadena pesada y liviana del supuesto anticuerpo progenitor putativo para la familia pueden inferirse y agregarse al conjunto de secuencias si aún no están en el conjunto. Se puede construir un árbol evolutivo para el conjunto utilizando, por ejemplo, Parsimonia Máxima, Probabilidad Máxima o cualquier algoritmo adecuado, y el árbol puede estar enraizado en la secuencia del anticuerpo progenitor. El árbol se puede construir sobre la base de las cadenas pesadas solas, o las cadenas livianas solas, o preferiblemente mediante la construcción del árbol con base en las cadenas pesadas y livianas simultáneamente. De modo que el árbol represente la coevolución de las cadenas pesadas y livianas.

#### 15 Selección de anticuerpos

Para cada paciente, la tabla de resultados de V-QUEST se revisó junto con los árboles construidos a partir de las secuencias Ejecutar 1 o Ejecutar 2 (que se ve en TreeViewX: <http://darwin.zoology.gla.ac.uk/~rpage/treeviewx>). Las secuencias representativas se seleccionaron para cubrir las diferentes familias de VDJ presentes en base a los datos V-QUEST, inspeccionando las secuencias correspondientes en el árbol para elegir las secuencias que parecían ser representativas del clado. Típicamente se seleccionó una secuencia de cada clado. Algunas de las secuencias seleccionadas provienen de familias con muchos miembros, pero algunas también fueron seleccionadas de familias con pocos miembros o un miembro. Las secuencias seleccionadas se describen en la Tabla 18. Para cada anticuerpo en la Tabla 18, la columna "Anticuerpo" es la misma que el texto que sigue a "-" en el campo ID de Cóntigo del nombre asociado con la secuencia en el Listado de Secuencias.

#### 25 Clonación y expresión de pares de inmunoglobulina de cadena liviana y pesada clonados

##### Vectores

Un sistema es un sistema de vector seleccionable de neomicina y dihidrofolato reductasa (DHFR) modificado de los vectores Invitrogen pADNc3.3 y pOptivec. Un sistema alternativo es el sistema Lonza GS, en donde el marcador seleccionable y amplificable es la glutamina sintetasa (GS) (véase más abajo). Las secuencias que codifican la cadena liviana kappa de la inmunoglobulina, la cadena liviana lambda y la cadena pesada gamma se insertan en vectores. Las secuencias de consenso y líder de Kozak ya están presentes en los clones y, por lo tanto, no es necesario que se diseñen en los vectores. Las regiones constantes se sintetizan para contener sitios de restricción flanqueantes en 5' y uno o más sitios de restricción internos. Para facilitar la clonación de cadenas pesadas y livianas de inmunoglobulinas variadas, los insertos están diseñados con múltiples sitios de restricción que aumentan la posibilidad de que el clon en sí no contenga el sitio de restricción y, por lo tanto, no se corte internamente. Los insertos tienen dos sitios de restricción de 8 cuchillas diferentes en el extremo 5' de la región del inserto y en dos sitios de restricción diferentes diseñados en regiones constantes. Los sitios de restricción 5' son FseI y PacI para ambas cadenas livianas, y AclI y AsiSI para la cadena pesada gamma. Los sitios de restricción diseñados en la región constante en sí mismos son NheI y XhoI para ambas cadenas livianas, y EcoRI y SacII para la cadena pesada gamma. Consúltense la Tabla 16 para ver la secuencia de las inserciones de región constante que contienen los sitios de restricción. Los clones de cadena pesada o liviana de la 1.<sup>a</sup> reacción de PCR se someten luego a una 2.<sup>a</sup> ronda de PCR con cebadores de clonación, que tienen sitios de restricción flanqueantes 5' que se incorporan a los clones. Las enzimas de restricción apropiadas se utilizan para cortar los vectores de expresión y los clones, que ahora tienen extremos complementarios y se ligan entre sí utilizando la ADN ligasa T4. Los sistemas de vectores Invitrogen y Lonza GS contienen un marcador de selección amplificable. Este marcador es DHFR en el sistema Invitrogen y GS en el sistema Lonza GS. Bajo la presión de selección del selector apropiado (metotrexato para DHFR y L-metionina sulfoximina para la glucosa sintetasa (GS)), los genes unidos al marcador de selección se amplifican junto con él. Con más copias de los genes de inmunoglobulina, hay una mayor secreción de anticuerpos. Esto es útil cuando es necesario purificar grandes cantidades de anticuerpos para su posterior selección *in vivo* de anticuerpos neutralizantes.

#### 50 Clonación y expresión

Suponiendo que los plasmoblastos de mayor afinidad se seleccionan durante la maduración del centro germinal, se espera que la familia clonal de mayor afinidad también tenga el mayor número de clones. Además, el clon de mayor afinidad dentro de cada familia clonal también será el clon más frecuente dentro de esa familia. Sobre la base de estos supuestos, se eligió expresar el clon de frecuencia más alta de las 5 familias clonales más grandes de cada muestra de paciente en algunos aspectos. Los clones se amplifican a partir del 1.<sup>er</sup> ADNc de PCR, donde todas las muestras de la misma placa, cada una con una sola célula, se han agrupado. El cebador directo contiene la identificación de la muestra y, por lo tanto, solo amplifica el código de barras del ADN con esa identificación de la muestra en particular. Esto es altamente específico porque las ID de muestra contienen diferencias de nucleótidos entre sí. Algunas ID de

muestra pueden tener cebadores directos de clonación idénticos y los clones amplificados por estos cebadores deben distinguirse entre sí mediante la selección de colonias bacterianas. Tanto los cebadores directos como los reversos contienen sitios de restricción flanqueantes (primera y segunda regiones de sitio de restricción) que permiten que el clon se integre (con el marco de codificación alineado) en un vector que ya contiene una región constante kappa, lambda o gamma. Véase Tablas 4 y 5 para clonar secuencias de cebadores. Las cadenas livianas se clonan en pADNc3.3 modificado y las cadenas pesadas en pOptivec modificado, o ambas cadenas se clonan en el vector de expresión dual Lonza GS. Las células de mamífero se transfectan doblemente con vectores de expresión separados que codifican los genes de las cadenas pesada y liviana de inmunoglobulina, o bien se transfectan individualmente con un vector de expresión dual que contiene los genes de las cadenas pesada y liviana. Los sobrenadantes que contienen los anticuerpos secretados se recogen y se criban en cuanto a las propiedades deseadas.

En algunos casos, las regiones variables de los genes de Ig se pueden clonar por síntesis de ADN e incorporando el ADN sintetizado en el vector que contiene la región constante apropiada utilizando enzimas de restricción y biología molecular estándar. Durante la síntesis, no es necesario seguir la secuencia exacta de nucleótidos siempre que la secuencia de aminoácidos no cambie, a menos que se desee la mutagénesis. Esto permite la optimización del codón que puede resultar en niveles de expresión más altos. Esto también permite agregar sitios de restricción con el propósito de clonar. Las secuencias no traducidas, tales como 5' UTR y secuencias de códigos de barras, no necesitan sintetizarse y las secuencias líderes también pueden intercambiarse por otras secuencias peptídicas de señal conocidas por niveles de expresión más altos. Esto da como resultado una secuencia de nucleótidos de Ig que puede ser muy diferente de las lecturas de alto rendimiento pero que dan la secuencia de aminoácidos idéntica cuando se expresa.

En algunos casos, el adaptador de código de barras de ID de muestra agregado durante la transcripción reversa ya puede incorporar un sitio de enzimas de restricción. Esto da como resultado un adaptador con un sitio de restricción 3' del código de barras de ID de muestra en el conjunto de amplicones de PCR. Durante la clonación con cebadores de clonación, los amplicones deseados se amplifican a partir de un grupo de amplicones específicos de placa utilizando cebadores en 5' que son complementarios a las secuencias de códigos de barras de ID de muestra, y cebadores 3' específicos de cadena (para las cadenas kappa, lambda y gamma). Los cebadores en 3' agregarán en 3' sitios de restricción. Los cebadores en 5' no necesitan agregar sitios de restricción, ya que el cebador en 5' ya contiene un sitio de restricción 3' del código de barras de ID de pozo. Después de esta amplificación, se utilizan enzimas de restricción para cortar el amplicón para la ligación en el vector que contiene el inserto de la región constante. Durante la digestión con enzimas de restricción, las secuencias agregadas al extremo 5' de las secuencias del gen Ig, como los códigos de barras y las secuencias universales, se cortan, ya que están en 5' del sitio de restricción 5'.

#### Métodos alternativos de clonación y expresión.

En otro aspecto, las regiones variables de los genes de Ig pueden clonarse por síntesis de ADN, e incorporando el ADN sintetizado en el vector que contiene la región constante apropiada utilizando enzimas de restricción y biología molecular estándar. Durante la síntesis, no es necesario seguir la secuencia exacta de nucleótidos siempre que la secuencia de aminoácidos no cambie, a menos que se desee la mutagénesis. Esto permite la optimización del codón que puede resultar en niveles de expresión más altos. Esto también permite agregar sitios de restricción con el propósito de clonar. Las secuencias no traducidas, tales como 5' UTR y secuencias de códigos de barras, no necesitan sintetizarse, las secuencias líderes también pueden intercambiarse por otras secuencias peptídicas de señal conocidas por niveles de expresión más altos. Esto da como resultado una secuencia de nucleótidos de Ig que puede ser muy diferente de las lecturas de alto rendimiento pero que dan la secuencia de aminoácidos idéntica cuando se expresa.

En otro aspecto, el adaptador de código de barras de -ID de pozo agregado durante la transcripción reversa ya puede incorporar un sitio de enzima de restricción. Esto da como resultado un adaptador con un sitio de restricción 3' del código de barras de ID de pozo en el conjunto de amplicones de PCR. Durante la clonación con cebadores de clonación, los amplicones deseados se amplifican a partir de un grupo de amplicones específicos de placa utilizando cebadores en 5' que son complementarios a las secuencias de código de barras de ID de pozo, y cebadores 3' específicos de cadena (para las cadenas kappa, lambda y gamma). cebadores en 3' agregarán en 3' sitios de restricción. Los cebadores en 5' no necesitan agregar sitios de restricción, ya que el cebador en 5' ya contiene un sitio de restricción 3' del código de barras de ID de pozo. Después de esta amplificación, se utilizan enzimas de restricción para cortar el amplicón para la ligación en el vector que contiene el inserto de la región constante. Durante la digestión con enzimas de restricción, las secuencias agregadas al extremo 5' de las secuencias del gen de Ig, como los códigos de barras y las secuencias universales, se eliminan, ya que están en 5' del sitio de restricción 5'.

#### Clonación de cadenas pesadas y livianas en vectores Lonza.

##### 55 *Clonación de las regiones constantes de inmunoglobulina.*

Los vectores de Lonza se obtuvieron a través del acuerdo de licencia académica de la Universidad de Stanford con Lonza. Las cadenas livianas kappa y lambda se insertaron en el vector pEE12.4 y la cadena pesada gamma se insertó en el vector pEE6.4. Las secuencias de las cadenas pesada y liviana se clonaron en dos etapas: primero, se clonaron

las regiones constantes, seguidas del extremo 5' de las cadenas de inmunoglobulina (las secuencias líder y V(D)J). Las inserciones de las regiones constantes fueron sintetizadas genéticamente por las tecnologías de ADN integrado (IDT) y contenían mutaciones silenciosas apropiadas para la optimización de genes y la incorporación de sitios de restricción. Los insertos se obtuvieron de IDT en su vector pIDTSmart patentado. Las secuencias de inserción están en la Tabla 17. Se usó IgG1 como la región constante de la cadena pesada gamma. Los alelos utilizados fueron Km3, Mcg-Ke-Oz y G1m3 para las cadenas kappa, lambda y gamma, respectivamente. Para incorporar las regiones constantes en los vectores Lonza, los vectores Lonza y los insertos de región constante pIDTSmart se transformaron individualmente en dam-dcm-*E. coli* competentes y los plásmidos se purificaron usando el kit miniprep Qiagen siguiendo las instrucciones del fabricante. Luego, los plásmidos se digirieron con HindIII y BclI a 37°C durante 1 hora y se aplicaron en un gel de agarosa al 1.2% a 150 V durante 1 hora. Los vectores Lonza digeridos y las inserciones de la región constante se purificaron en gel y se ligaron (pEE12.4 con Km3 o cadenas livianas de Mcg-Ke-Oz y pEE6.4 con la cadena pesada gamma1 G1m3) utilizando T4 ADN ligasa durante 10 minutos a temperatura ambiente en una proporción de 3:1 inserto:vector. La ADN ligasa de T4 se inactivó luego a 70°C durante 8 minutos y 5 ul de la mezcla de ligación se transformaron en células TOP10 competentes para el choque térmico utilizando técnicas estándar de biología molecular. Las colonias se seleccionaron y la inserción se verificó mediante la secuenciación de Sanger.

#### Clonación de las regiones variables de inmunoglobulina.

A continuación, se digirió pEE12.4 que contenía cadena liviana lambda o kappa usando Ascl y XmaI durante 5 horas a 37°C y se purificó en gel. El pEE6.4 que contenía la cadena pesada gamma 1 se digirió usando Ascl y Agel durante 5 horas a 37°C y se purificó en gel. Los amplicones seleccionados se amplificaron selectivamente usando cebadores directos específicos de ID de pozo y cebadores reversos específicos de región constante (Tablas 4 y 5) de la dilución de placa específica 100x ID de la 1ª PCR. Los cebadores directos tenían el sitio de restricción Ascl en el extremo 5' del cebador, y los cebadores reversos contenían el sitio de restricción XmaI o Agel para los cebadores de la región constante de los cebadores de cadena liviana y pesada respectivamente. El ciclo de la PCR se realizó utilizando una desnaturalización inicial a 98°C durante 30 segundos y de 35 a 45 ciclos a 98°C durante 10 segundos, 68°C durante 15 segundos y 72°C durante 20 segundos. La extensión final fue de 72°C a los 5 minutos y en espera a 4°C por tiempo indefinido. Los productos de la PCR se purificaron utilizando placas de ultrafiltración PCR<sub>0.96</sub> de Millipore, siguiendo las instrucciones del fabricante. Después de esto, los productos de la PCR se sometieron a una doble digestión durante 3 horas a 37°C utilizando Ascl y XmaI para las cadenas livianas y Ascl y Agel para la cadena pesada gamma 1. Los productos digeridos se procesaron luego en una agarosa de bajo punto de fusión al 2.5% con gelgeen (Biotium) y se visualizaron bajo luz azul. Se cortaron secciones de gel que contenían bandas en los tamaños apropiados. La ligación en gel se realizó fundiendo las secciones de gel a 65°C y añadiendo los vectores Lonza tratados con fosfatasa Antartic (NEB) digeridos y apropiados que contenían el inserto de la región constante, e incubando con ADN ligasa T4 durante 1-3 horas a temperatura ambiente. Luego, las bacterias competentes para el choque térmico se transformaron y se colocaron en placas sobre agar de ampicilina. Se seleccionaron 6 clones por constructo y se cultivaron en 2xLB (2x concentrado de caldo Luria-Bertani). Miniprep se realizó utilizando las placas de filtro Multiscreen de 96 -pozos de Millipore de acuerdo con las instrucciones del fabricante ([www.millipore.com/techpublications/tecg1/tn004](http://www.millipore.com/techpublications/tecg1/tn004)) para obtener el ADN plasmídico. La PCR de colonias se realizó utilizando una desnaturalización inicial de 95°C durante 5 minutos y 40 ciclos de 95°C durante 1 minuto, 50°C durante 2 minutos, 72°C durante 1 minuto y una extensión final a 72°C durante 5 minutos y manteniendo a 4°C indefinidamente. Se enviaron clones con el inserto apropiado para la secuenciación de Sequetech, Mountain View, CA, USA. La identificación del VDJ del clon se realizó utilizando IMGT HIGHV-Quest y los clones correctos se mantuvieron como stock bacteriano (almacenado a -80°C con glicerol al 15%) y plásmidos.

#### Expresión de anticuerpos monoclonales en 293T.

Se realizaron transfecciones duales transitorias de construcciones pareadas pEE12.4-cadena liviana y pEE6.4-cadena pesada usando Lipofectamine 2000 siguiendo el protocolo del fabricante. Las transfecciones se han realizado en placas de 48 pozos, 24 pozos, 6 pozos, 60mm y 100mm. En resumen, las células 293T se cultivaron en DMEM+ FBS IgG ultrabajo al 10% (Invitrogen) para evitar que la IgG bovina compita con la IgG humana secretada en la etapa de purificación de la proteína A aguas abajo. Se cultivaron células 293T durante 20 pases antes de descongelar y utilizar nuevas partes alícuotas de N<sub>2</sub> líquido. Para las transfecciones de placas de 48 pozos, cada pozo se sembró el día anterior con 8x10<sup>4</sup> células y se dejó crecer hasta una confluencia del ~90% al día siguiente. Se incubaron 50 ng de construcciones de cadena liviana y pesada en medio óptimo para un volumen final de 50 ul, y Lipofectamine 2000 también se incubó por separado con 50 ul de medio óptimo. Ambas incubaciones fueron de 5 a 25 minutos. Lipofectamine 2000 y los constructos se mezclaron luego mediante pipeteo suave y se incubaron durante 20 minutos antes de agregar a las células 293T y se mezclaron suavemente. Los medios se cambiaron al día siguiente y los sobrenadantes de cultivo se recolectaron cada dos días (por ejemplo, Lunes, Miércoles y Viernes) durante 2 semanas. Para las transfecciones de otros tamaños, se utilizaron las siguientes cantidades de construcciones y Lipofectamine 2000: para las transfecciones en placas de 24-pozos, se utilizaron 100 ng de cada construcción con 1.25ul de Lipofectamina 2000. Para placas de 60 mm, se utilizaron 625 ng de cada construcción con 12.5 ul de Lipofectamine 2000. Para las transfecciones de 100 mm, se utilizaron 3 µg de cada construcción en 37.5ul de Lipofectamine 2000.

60 ELISA anti-IgG humana



En algunos casos, se realizó ELISA de IgG humana en la muestra para cuantificar la cantidad de IgG expresada en el sobrenadante de cultivo, y el sobrenadante de cultivo se usó directamente en aplicaciones posteriores después de normalizar la cantidad de anticuerpo. El kit de cuantificación ELISA anti-IgG humana se adquirió de Bethyl Laboratories y se realizó de acuerdo con las instrucciones del fabricante. En resumen, se recubrieron 100µl de anticuerpo de captura en placas Nunc Maxisorp durante la noche a 4°C y se lavaron 5 veces con PBST (PBS con Tween20 al 0.05%). Los pozos se bloquearon con BSA al 1% en PBS durante una hora a temperatura ambiente, luego se lavaron 5 veces con PBST. Los pozos se incubaron luego con las diluciones estándar apropiadas del kit o los sobrenadantes de cultivo diluidos durante 1 hora a temperatura ambiente, luego se lavaron 5 veces con PBST. Se añadieron 100µl de anticuerpo de detección de HRP diluido a cada pozo y se incubaron a temperatura ambiente durante una hora, luego se lavaron 5x con PBST. Se agregaron 50 µl de solución de sustrato TMB y la reacción se detuvo con 50 µl de solución de parada. La absorbancia se leyó en un espectrofotómetro SpectraMax M5 a 450 nm y se generaron curvas estándar con una curva de 4 parámetros. Los anticuerpos se mantuvieron a 4°C en PBS con azida de sodio al 0.1% como conservante.

#### Purificación de la proteína A-IgG de los anticuerpos monoclonales expresados

En otros casos, los anticuerpos se purificaron primero a partir del sobrenadante de cultivo y se cuantificaron utilizando BCA antes de su uso. En resumen, los sobrenadantes de cultivo se recogieron 3 veces por semana durante 2 semanas en tubos de 50 ml y se almacenaron a 4°C con azida de sodio al 0.1% como aditivo hasta la purificación de la proteína A-IgG. Los sobrenadantes de cultivo se hicieron girar y se decantaron para eliminar cualquier agregado celular. Se añadió 1M pH 9.0 Tris a los sobrenadantes de cultivo para asegurar que el pH está entre 7.5 y 8.0 según lo determinado por las tiras indicadoras de pH. Las perlas de proteína A más agarosa (Pierce) se lavaron 2 veces con PBS antes de agregar 400 µl de una suspensión al 50% a los sobrenadantes de cultivo y se incubaron a 4°C durante la noche en un rotador para asegurar una mezcla uniforme de las perlas. Las perlas se recuperaron haciendo girar el sobrenadante de cultivo a 1000 g durante 5 minutos y pipeteando las perlas desde el fondo de los tubos en columnas de flujo por gravedad de 5 ml. Las perlas se lavaron con 4x2ml de PBS antes de la elución con 2x1.5ml de regulador de elución IgG (Pierce), que es un regulador de elución de bajo pH, en una columna concentradora Amicon-4 de 100 kDa. El anticuerpo eluido se neutralizó inmediatamente con 400 µl de Tris 1M pH 8.0. Los anticuerpos se concentraron luego girando a 1000 g en concentradores Amicon-4 de 100 kDa durante 10 minutos, seguido de un lavado de 2mL de PBS y un lavado de 2ml de PBS con azida de sodio al 0.1%. Las concentraciones de anticuerpos se determinaron mediante un ensayo de BCA y se ajustaron a 0.5 mg/ml. La proteína A Plus Agarosa se regeneró lavando con 1x2ml de PBS, 3x2ml de regulador de elución IgG y un lavado con 3 ml de PBS con azida de sodio al 0.1% y se mantuvo a 4°C. Las columnas se pueden regenerar hasta 5 veces.

#### Cribado de anticuerpos humanos expresados

##### Cribas para la unión de anticuerpo-antígeno

Los anticuerpos seleccionados (Consúltese la sección de clonación y expresión anterior) se analizan primero por su capacidad para unirse al antígeno de interés, y luego se expresan los anticuerpos de toda la familia clonal y se analizan por su capacidad que bloquea o neutraliza el antígeno (Consúltese "cribas funcionales" abajo). La concentración de IgG en los sobrenadantes que contienen anticuerpos de interés se determina primero mediante ELISA de IgG, de modo que se pueda usar la misma cantidad de IgG para cada muestra en el análisis de unión antígeno-anticuerpo. En otros casos, la IgG se purificó del sobrenadante usando perlas de agarosa de proteína A y se cuantificó con un ensayo de BCA usando inmunoglobulinas bovinas como estándar. Las IgG purificadas se normalizaron a la misma concentración antes de su uso en la selección de anticuerpos. Para detectar anticuerpos que se unen al antígeno, realizamos un ELISA indirecto. Una placa de 96 pozos se recubre durante la noche con el antígeno de interés, y luego se elimina el exceso de antígeno. Los sobrenadantes que contienen anticuerpos de interés se agregan a los pozos y se incuban durante cuatro horas, después de lo cual se lavan los pozos. Como control positivo, se agregan cantidades conocidas de anticuerpos disponibles comercialmente (de especies no humanas) que son específicas para el antígeno a pozos separados que contienen el antígeno. Como control negativo, los anticuerpos disponibles comercialmente específicos para un antígeno irrelevante se agregan a pozos separados que contienen el antígeno de interés. Luego se agrega un anticuerpo secundario conjugado con HRP a los pozos, se incuba durante 30 minutos y se lava el exceso. Luego se agrega tetrametilbencidina (TMB) a la placa y se deja que la reacción continúe hasta que se observe el color en los pozos de control positivo. La reacción se detiene con ácido y se mide la absorbancia. Se considera que los sobrenadantes de pozos específicos contienen anticuerpos que se unen al antígeno de interés si la lectura de absorbancia que producen es significativamente mayor que la del control negativo.

#### ELISA de Fluzone

A los voluntarios se les administró la vacuna contra la gripe de la temporada 2010/2011 de Fluzone, que consta de 3 cepas de virus inactivados, las cepas A/California/7/2009, A/Perth/16/2009, B/Brisbane/60/2008. Los ELISA de Fluzone se realizaron para determinar si los anticuerpos monoclonales derivados de voluntarios vacunados se unen a la vacuna contra la gripe en sí misma como una prueba inicial para detectar anticuerpos expresados con actividad de unión. La vacuna de Fluzone se diluyó 100 veces en un regulador de carbonato de pH9 y se recubrió en placas Nunc Maxisorp a temperatura ambiente durante 1 hora o toda la noche a 4°C. Las placas se lavaron 5 veces con PBST (PBS w/0.05% Tween20) y se bloquearon con PBS w/1% BSA durante 1 hora a temperatura ambiente. Luego se

agregaron 100  $\mu$ l de 100 ng/ml de anticuerpos contra la gripe expresados a los pozos a temperatura ambiente durante 1 hora, antes de lavar 5 veces con PBST y agregar el anticuerpo de detección de HRP diluido (del kit de cuantificación ELISA de IgG humana de Bethyl Labs) durante 1 hora a temperatura ambiente. Las placas se lavaron 5x con PBST y se agregaron 50  $\mu$ l de sustrato TMB. Se dejó que el color se desarrollara hasta 30 minutos antes de detener la reacción con 50  $\mu$ l de solución de parada. Las placas se leyeron en un espectrofotómetro SpectroMax M5 a una absorbancia de 450 nm. Los anticuerpos utilizados en este ensayo se describen en la Tabla 19. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

Determinación de resonancia de plasmón superficial de afinidades de anticuerpos de gripe.

La unión de anticuerpos monoclonales (mAbs) a moléculas de HA se analizó a 25°C utilizando un biosensor de resonancia de plasmón de superficie ProteOn (BioRad Labs). Los anticuerpos monoclonales contra la gripe expresados derivados de los plasmoblastos de un donante vacunado contra la gripe (25 nM en regulador de acetato a pH 4.5) se acoplaron a un chip sensor GLC con acoplamiento de amina a una densidad objetivo de 800 unidades de resonancia (RU) en las células de flujo de prueba usando Química de EDAC-NHS. Los grupos de éster activos sin reaccionar se inactivaron con etanolamina. La hemaglutinina recombinante purificada H3 (HA( $\Delta$ TM)(H3N2/Perth/16/2009) y H1 (HA( $\Delta$ TM)(A/California/07/2009)(H1N1)) se adquirieron de Immune Technology Corp. (New York, NY) y se diluyeron a 100, 50, 25, 12.5, 6.25, nM, junto con un control de regulador en blanco se inyectaron a un caudal de 30  $\mu$ L/min con 120 segundos de tiempo de contacto y 2000 segundos de tiempo de disociación. La cinética de unión y el análisis de datos se realizaron con el software administrador BioRad ProteON. Las mediciones de afinidad se calcularon utilizando el algoritmo de analito bivalente, ya que la HA consistía en varias unidades de repetición. La precisión de las curvas ajustadas se verificó al verificar que los valores de  $\chi^2$  de la bondad de cada ajuste eran inferiores al 10% del valor máximo de unión (Rmax). Los anticuerpos utilizados en este ensayo se describen en la Tabla 20. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

Reactividades de anticuerpos de RA en micromatrices de antígeno de RA

Para imprimir las micromatrices de antígeno, los antígenos se diluyeron en solución salina tamponada con fosfato a 0.2 mg/mL y se unieron a los portaobjetos ArrayIt SuperEpoxy utilizando un sistema ArrayIt NanoPrint Protein LM210. Los portaobjetos se marcaron con un marcador hidrofóbico y se bloquearon durante la noche en PBS con suero bovino fetal al 3% y Tween 20 al 0.05% a 4°C, meciéndose suavemente a 30 rpm. Las matrices se sondaron con 400  $\mu$ l de 40  $\mu$ g/mL de anticuerpo monoclonal durante una hora a 4°C, se sacudieron suavemente a 30 rpm. Luego se lavaron las matrices y se incubaron en un anticuerpo secundario anti-IgG/IgM humano conjugado con Cy3 diluido a 1:2500 durante 45 minutos a 4°C, se sacudió suavemente a 30 rpm. Después de otro lavado, las diapositivas se escanearon utilizando un escáner de micromatrices GenePix 4300A. Se usó el software GenePix 7 para encontrar la intensidad de la fluorescencia media de cada característica y fondo.

Para analizar los datos, se restaron las intensidades de fluorescencia de fondo de cada característica y se expresaron como el valor mediano de las cuatro características de antígeno en cada matriz. Las intensidades medianas se transformaron logarítmicamente con un valor de corte de 10. Estos valores se sometieron a agrupación jerárquica utilizando el software Cluster para organizar antígenos en función de las similitudes entre sí. Las relaciones se mostraron como un mapa de calor utilizando el software Java TreeView. Los anticuerpos utilizados en este ensayo se describen en la Tabla 21. Las secuencias de los anticuerpos se pueden consultar en la tabla principal, Tabla 18.

Anti-histona 2A ELISA

Se utilizó un ELISA directo para la detección de anticuerpos contra la histona 2A. Las placas de microtitulación (Nunc Maxisorp) se recubrieron con 100  $\mu$ l de H2A recombinante en regulador de carbonato, a una concentración de 20  $\mu$ g/ml, y se incubaron a 4°C durante la noche. Después del bloqueo en PBS que contenía un 1% de albúmina de suero bovino (BSA), se utilizaron anticuerpos derivados de pacientes con AR en una titulación de 15  $\mu$ g/ml a 250  $\mu$ g/ml en regulador de dilución (PBS que contenía BSA al 0.1% y Tween-20 al 0.1%). se agregó a la placa por duplicado a 100  $\mu$ l/pozo y se incubó durante 2 horas a temperatura ambiente. Las muestras se incubaron luego durante 1 hora a temperatura ambiente con una dilución 1:5000 de un anticuerpo monoclonal de cabra anti-humano marcado con peroxidasa de rábano picante. La reacción se desarrolló mediante la aplicación de sustrato 3,3',5,5'-tetrametilbencidina (TMB) (Sigma-Aldrich) durante 15 minutos y se detuvo mediante la adición de 50  $\mu$ l de 2N H<sub>2</sub>SO<sub>4</sub>. La cuantificación relativa de anticuerpos contra se realizó mediante densitometría óptica a 450 nm utilizando un suero de seropositivo RA conocido como control positivo. Los anticuerpos utilizados en este ensayo se describen en la Tabla 22. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

ELISA de Anti-CCP2

Anti-CCP2 ELISA se realizó de acuerdo con las instrucciones del fabricante (Eurodiagnostica, Malmo, Suecia). En resumen, los anticuerpos derivados de pacientes con AR se diluyen a aproximadamente 125  $\mu$ g/ml en regulador de dilución (PBS que contiene BSA al 0.1% y Tween-20 al 0.1%), se agregan a la placa de ELISA CCP2 comercial previamente bloqueada a 100  $\mu$ l/pozo y se incuban durante 2 horas a temperatura ambiente. Las muestras se incubaron luego durante 1 hora a temperatura ambiente con una dilución 1:5,000 de un anticuerpo monoclonal de

5 cabra anti-humano marcado con peroxidasa de rábano picante. La reacción se desarrolló mediante la aplicación de sustrato 3,3', 5,5'-tetrametilbencidina (TMB) (Sigma-Aldrich) durante 15 minutos y se detuvo mediante la adición de 50 µl de 2N H<sub>2</sub>SO<sub>4</sub>. La cuantificación relativa de anticuerpos contra se realizó mediante densitometría óptica utilizando los estándares proporcionados por el vendedor y el suero de RA positivo conocido. Los anticuerpos utilizados en este ensayo se describen en la Tabla 22. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

#### ELISA de Factor Anti-Reumatoide

10 Para la detección de anticuerpos contra el factor reumatoide (RF), las placas de microtitulación (Nunc Maxisorp) se recubrieron con 10 µg/ml de IgG de conejo en regulador de carbonato y se incubaron a 4°C durante la noche. Después del bloqueo en PBS que contenía un 1% de albúmina de suero bovino (BSA), los anticuerpos derivados de pacientes con AR fueron de 5 µg/ml de regulador de dilución (PBS que contenía un 0.1% de BSA y un 0.1% de Tween-20), se agregaron a la placa por duplicado a 100 µl/pozo, y se incubó durante 2 horas a temperatura ambiente. Las muestras se incubaron luego durante 1 hora a temperatura ambiente con una dilución 1:5,000 de un anticuerpo monoclonal de  
15 sustrato 3,3', 5,5'-tetrametilbencidina (TMB) (Sigma-Aldrich) durante 15 minutos y se detuvo mediante la adición de 50 µl de 2N H<sub>2</sub>SO<sub>4</sub>. La cuantificación relativa de anticuerpos contra se realizó mediante densitometría óptica a 450nm utilizando dos controles de control de RF+ conocidos como controles positivos. Los anticuerpos utilizados en este ensayo se describen en la Tabla 23. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

20 Inmunohistoquímica de anticuerpos de pacientes con adenocarcinoma de pulmón en arreglos de tejidos de cáncer de pulmón

Se compraron dos tipos diferentes de portaobjetos de micromatrices tisulares de US Biomax. Eran VLC 12 y BS0481. Se incluye una variedad de núcleos de tejido de carcinomas de pulmón en los portaobjetos, que incluyen adenocarcinomas de pulmón y también controles de tejido pulmonar normal. Los portaobjetos se calentaron en un regulador de recuperación de antígeno citrato pH 6.0 a 95-99°C durante 40 minutos antes de dejar enfriar a temperatura ambiente. Los portaobjetos se trataron previamente con Triton-X al 0.02% y H<sub>2</sub>O<sub>2</sub> al 0.6% durante 20 minutos. Luego se bloquearon los portaobjetos con suero de cabra normal al 10% en TBST (TBS con Tween20 al 0.05%) durante 2 horas antes de bloquear más en 100µg/ml de IgG anti-humana de cabra F(ab) (Jackson Immunoresearch) durante la noche a 4°C. Luego, los portaobjetos se sometieron a un bloque de avidina/biotina (Vector Laboratories) de acuerdo con las instrucciones del fabricante. Los portaobjetos se incubaron luego en 5 o 10 µg/ml de anticuerpo pulmonar expresado durante 1 hora a temperatura ambiente, se lavaron 3x5 minutos con TBST y luego se incubaron con un anticuerpo secundario biotinilado de cabra anti-humano durante 20 minutos a temperatura ambiente. Los portaobjetos se lavaron luego 3x5 minutos con TBST y se incubaron con el reactivo Vectastain ABC preparado durante 30 minutos a temperatura ambiente. Luego, los portaobjetos se lavaron 3x5 minutos de TBST y se tiñeron con Vector Red (Vector Laboratories) y se siguió el desarrollo de color con un microscopio óptico. Después del tiempo de tinción apropiado, la reacción se detuvo lavando con agua destilada y se tiñe de nuevo con hematoxilina. Los portaobjetos se montaron en agua y se fotografiaron con un microscopio BX-51. Los anticuerpos utilizados en este ensayo se describen en la Tabla 24. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

40 Determinación por citometría de flujo de la unión de los anticuerpos expresados en pacientes de adenocarcinoma de pulmón a líneas celulares de cáncer de pulmón

Las líneas celulares de cáncer de pulmón utilizadas fueron A549, H226, H441, H23, H1975, H1437, H2126, H1650 y H2009. Las células HEK 293T también se usaron como control negativo. Las células se separaron incubando células en PBS sin Ca<sup>2+</sup> y Mg<sup>2+</sup> con EDTA 2 mM durante 1 hora a 37°C. Esto es para evitar que los antígenos de la superficie celular se dañen con la tripsinización o cualquier otro producto proteolítico para separar las células. Las células se lavaron una vez con regulador FACS (HBSS con 2% de FCS) antes de suspenderlas en 50 µl de regulador FACS y se incubaron con 10 µg/ml, 3µg/ml, 1µg/ml, 0.2µg/ml de anticuerpos pulmonares expresados para valorar la dosis. Se encontró que la concentración óptima era del rango de 0.2-1µg/ml. Por lo tanto, se utilizaron 1 µg/ml, 0.5 µg/ml, 0.25 µg/ml de anticuerpos pulmonares posteriormente. Los anticuerpos pulmonares se incubaron durante 30 minutos a 4°C antes de lavar con 2x 200 µl de regulador FACS en placas de 96 pozos. Luego se añadió IgG-PE antihumana y se incubó durante 15 minutos a 4°C en la oscuridad. Las muestras se lavaron luego 2 x 200 µl con regulador FACS y se resuspendieron en 200 µl de regulador FACS y se analizaron en un BD LSR II o LSR Fortessa. Sytox blue se usó como tinción viva/muerta. Los anticuerpos utilizados en este ensayo se describen en la Tabla 24. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

55 Citometría de Flujo de Estafilococos

Se obtuvieron partículas fijadas de *S. aureus* (cepa Wood) de Invitrogen. La cepa de madera es una cepa que expresa la proteína A mínima por parte de algunas de las bacterias. Las partículas se suspendieron en 50 µl de regulador FACS a 10x10<sup>6</sup> células/50 µl, y se incubaron con una titulación de 10 µg/ml, 5 µg/ml o 1 µg/ml de anticuerpos

expresados derivados de individuos estafilococos durante 1 hora a 4°C. Las partículas de estafilococos fijas se lavaron luego dos veces con regulador FACS antes de incubar durante 15 minutos en la oscuridad a 4°C con anticuerpo IgG-FITC antihumano. Las partículas se lavaron luego con 1 ml de regulador FACS y se resuspendieron en 200 ul de regulador FACS para su análisis en un BD LSR II o LSR Fortessa. Los anticuerpos utilizados en este ensayo se describen en la Tabla 25. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

#### Cribas funcionales

##### Bloqueo de anticuerpos a interacciones receptor-ligando

Para detectar la capacidad de los anticuerpos de bloquear las interacciones ligando-receptor (por ejemplo, las interacciones citoquina-receptor), se transfectaron las células 293T con un vector que codifica el receptor apropiado. Estas células 293T también se transfectan de forma estable con un indicador de luciferasa dependiente de NF-κB, de modo que estas células 293T transfectadas de forma estable expresan luciferasa cuando se activa NF-κB. Luego cultivamos las células 293T transfectadas con el ligando apropiado en presencia o ausencia de los anticuerpos candidatos. Las células 293T finalmente se analizan para determinar la expresión de luciferasa midiendo la emisión de luz dependiente de luciferasa. Interacción entre el ligando y su receptor, por ejemplo la interacción entre IL-17A y IL-17R, activa NF-κB. Los anticuerpos bloqueantes evitan la señalización de NF-κB por la unión del ligando-receptor y, por lo tanto, anulan la expresión de luciferasa. En los casos en que la interacción ligando-receptor no activa el NF-κB, se utilizan otros elementos de respuesta transcripcional para impulsar el promotor del gen de la luciferasa, por ejemplo elementos de respuesta AP-1, etc.

##### Selección de anticuerpos por su capacidad para inhibir la función de las citoquinas o inhibir un ensayo funcional

Los ensayos funcionales también pueden usarse para detectar anticuerpos anti-citoquinas en sueros de pacientes o en anticuerpos clonados y expresados. En esta metodología, los anticuerpos humanos expresados se analizan para determinar su capacidad para inhibir la inducción de citoquinas u otros mediadores inmunitarios de una respuesta celular.

##### Anticuerpos contra bacterias, células infectadas por virus, parásitos o células cancerosas

Para detectar anticuerpos que matan o neutralizan bacterias, células infectadas con virus, parásitos o células cancerosas, cultivamos el tipo de célula apropiado en presencia o ausencia de anticuerpo, junto con suero no inactivado por calor (que contiene factores del complemento). Si el anticuerpo es un anticuerpo neutralizante, opsonizará las bacterias, otros microbios o células cancerosas y activará los componentes del complemento que forman el complejo de ataque a la membrana (MAC), que induce la muerte celular. Para probar la neutralización, realizamos un ensayo fluorescente vivo/muerto (Invitrogen), en donde las células vivas y muertas se tiñen con diferentes fluoróforos. Las células se pueden analizar para determinar los porcentajes de células vivas y muertas mediante el uso de citometría de flujo. El anticuerpo que da como resultado el mayor porcentaje de células muertas será un buen anticuerpo neutralizador candidato que se analizará más a fondo en cribas *in vivo*.

##### Anticuerpos que neutralizan los virus.

Para detectar los anticuerpos que neutralizan los virus, realizamos un ensayo estándar de reducción de placa u otros ensayos de infección celular *in vitro*. Se espera que los anticuerpos neutralizantes disminuyan la infección viral de las células. Los anticuerpos candidatos se prueban en un modelo *in vivo*.

##### Ensayo de microneutralización de influenza

Algunos anticuerpos contra la gripe expresados que mostraron actividad de unión al ELISA de Fluzone se enviaron a un CRO externo, Virapur, LLC para los ensayos de microneutralización. En resumen, se mezclaron diluciones dobles de cada anticuerpo, a partir de 100 ug/ml, con un volumen igual de aproximadamente 100 TCID<sub>50</sub> unidades infecciosas de virus de reserva titulado en cuadruplicados en pozos de una placa de 96 pozos. Las soluciones de virus/anticuerpo se incubaron durante 2 horas y luego la mezcla se transfirió a una placa de 96 pozos que contenía 80% de células MDCK confluentes. Las células, el anticuerpo y el virus se incubaron durante 2 horas adicionales a 37°C, después de lo cual se eliminó el virus, se enjuagaron las monocapas y se agregaron medios de crecimiento virales a cada pozo. Los pozos se observaron al microscopio después de 72 horas para detectar la presencia de infección por el virus de la influenza. Los anticuerpos utilizados en este ensayo se describen en la Tabla 26. Las secuencias de los anticuerpos se pueden consultar en la tabla principal, Tabla 18.

##### Ensayo de Inhibición de Estafilococos

Se utilizó *S. aureus* cuando en crecimiento de fase logarítmica. Se agregaron a placas de polipropileno de 96 pozos y se añadió anticuerpo anti-estafilococo de pacientes con estafilococos a 10 ug/ml. Se añadió complemento de conejo bebé (Cedarlane) a la cantidad recomendada por el fabricante y se mezcló bien. Las placas se incubaron a 37°C durante 45 minutos antes de diluirse 1:10, 1:100 y 1:1000 y se sembraron en placas de agar con sangre con TSA al

5% y se cultivaron durante la noche. Las UFC bacterianas se contaron y se tabularon al día siguiente. Los anticuerpos utilizados en este ensayo se describen en la Tabla 27. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

#### 5 Inmunoprecipitación de antígenos de estafilococos con anticuerpos derivados de pacientes infectados por estafilococos

Se preparó lisado de estafilococo lisando *S. aureus* utilizando B-Per reactivo de extracción de proteína bacteriana (Pierce) con 100 ng/ml de lisostafina durante 30 minutos a temperatura ambiente junto con 1x inhibidor de proteasa Halt, y se separó de la fracción insoluble mediante centrifugación a 15000 rpm en una microcentrífuga. El lisado se limpió por incubación con proteína G Dynabeads durante 1 hora a temperatura ambiente. 5 ug de anticuerpo derivado del paciente Staph se unieron a proteínas G Dynabeads incubando durante 1 hora a temperatura ambiente. Los anticuerpos unidos a la proteína G se incubaron luego con lisado de estafilococo pre-limpiado durante la noche a 4°C. Luego, las perlas se lavaron 3x con PBST (PBS con Tween20 al 0.1%) y se calentaron con 5x de regulador de muestra de carril reductor (Thermo Scientific) a 95°C durante 5 minutos antes de ejecutar una SDS-PAGE en un gel Bis-Tris de 4-12% de Criterion. Las proteínas se visualizaron con RAPIDStain Reagent (Calbiochem). Los anticuerpos utilizados en este ensayo se describen en la Tabla 28. Las secuencias de anticuerpos se pueden consultar en la tabla principal, Tabla 18.

Identificación de péptidos por espectrometría de masas.

Las bandas de proteína teñidas de interés se cortaron de los geles, se sumergieron en bicarbonato de amonio 10 mM que contenía DTT 10 mM y yodoacetamida 100 mM, se trataron con acetonitrilo al 100% y luego se digirieron durante la noche a 37°C con 0.1 mg de tripsina (Sigma-Aldrich) en acetato de amonio 10 mM que contiene acetonitrilo al 10%. Las proteínas tripsinizadas se identificaron con LCMS utilizando el sistema Agilent 1100 LC y la trampa de iones ultra XCT de Agilent (Agilent Technologies, Santa Clara, CA) como se describió anteriormente (Lopez-Avila V, Sharpe O, Robinson WH: Determination of ceruloplasmin in human serum by SEC-ICPMS. Anal Bioanal Chem 2006, 386:180-7.). Los datos de LCMS se escanearon contra las bases de datos SwissProt o NCBIInr utilizando el software SpectrumMill (Agilent) para la detección de péptidos utilizados para identificar proteínas. Los anticuerpos utilizados en este ensayo se describen en la Tabla 29.

Ejemplo 1: Secuenciación de alto rendimiento de secuencias pareadas de cadenas pesadas y livianas de células B individuales.

Se desarrolló un método para agregar códigos de barras compuestos (ID de muestra + placa-ID) a las secuencias con el fin de identificar inequívocamente qué secuencias se originaron del mismo pozo en una placa. Utilizamos este método para secuenciar genes de inmunoglobulina de cadena liviana y cadena pesada de células B individuales. Las células B individuales se pueden clasificar por citometría de flujo de sangre, células mononucleares de sangre periférica (PBMC), células B de volumen, plasmoblastos, células plasmáticas, células B de memoria u otras poblaciones de células B (Figura 1).

Primero, las células B se clasificaron en células individuales en placas de PCR de 96 pozos, dejando una columna de pozos vacía, como control negativo. Los oligonucleótidos que contienen diferentes códigos de barras de ID de muestra se agregaron a diferentes pozos durante la transcripción reversa (RT). Después de la transcripción reversa del ARNm, la MMLV H- transcriptasa reversa cambia las plantillas y transcribe el oligonucleótido, incorporándolo y el código de barras de la ID de la muestra en el extremo 3' del ADNc de la 1.<sup>a</sup> cadena (Figura 2a). Todos los ADNc (código de barras con un ID de muestra) de una placa se agruparon y se sometieron a dos rondas de PCR. Durante la PCR, se agregaron 454 cebadores de secuenciación (primera y segunda regiones de secuenciación) y códigos de barras de ID de placa en los extremos 5' y 3' del amplicón mediante el uso de cebadores de PCR con secuencias de código de barras flanqueantes en 5'. Los amplicones (regiones de amplicón) de diferentes placas ahora tienen n ID de placa diferentes, y el código de barras compuesto que comprende una ID de placa y una ID de muestra identifica inequívocamente las secuencias como provenientes de una célula en particular, permitiendo el emparejamiento de genes de cadena pesada -y liviana- secuenciados (Figura 2b-c).

La Figura 3 describe la metodología general utilizada y las secuencias asociadas. Los cebadores y las moléculas adaptadoras se muestran en la Tabla 1. Las secuencias de ID de muestra se muestran en la Tabla 2. Las secuencias de ID de placa se muestran en la Tabla 3. Los cebadores de clonación se muestran en la Tabla 4, con la secuencia 3' de los cebadores directos de clonación que se muestran en la Tabla 5.

Se obtuvieron productos de PCR en los tamaños esperados: ~600 pb para las cadenas livianas kappa y lambda, y ~700 pb para la cadena pesada gamma (Figura 4a). A continuación, se envió el material para la secuenciación de Sanger. Obtuvimos secuencias que fueron identificadas por NCBI BLAST como cadenas kappa, lambda y gamma (datos no mostrados). La investigación adicional del cromatograma de ADN mostró una mezcla de varios picos que comienzan en los códigos de barras de ID de muestra, lo que demuestra que agregamos con éxito códigos de barras de ID de muestra a ADNc de células en diferentes pozos durante la RT y los amplificamos con éxito en dos rondas posteriores de PCR (Figura 4b). La cadena de secuenciación de Sanger del extremo 3' también mostró una mezcla

de varios picos a partir de la unión VJ, debido a la amplificación de genes de diferentes células, que diferían después de la unión VJ como resultado de las inserciones y eliminaciones y la recombinación aleatoria de diferentes V y J genes. Además, cuando realizamos la PCR con cebadores de clonación específicos para el ID de muestra del pozo A1, obtuvimos un solo pico en lugar de una mezcla de varios picos, lo que demuestra que de hecho se pueden amplificar secuencias de una célula específica dentro del conjunto (Figura 4c-d).

Ejemplo 2: Esquema de restricción para la selección de células individuales de plasmoblastos.

Los plasmoblastos se definieron como CD19<sup>+</sup>CD20<sup>-</sup>CD27<sup>+</sup>CD38<sup>++</sup> para este experimento. La Figura 22 muestra un esquema de selección para la clasificación por citometría de flujo de células de plasmoblastos individuales en placas de 96 pozos.

10 Se prepararon PBMC individuales y se tiñeron como se describe anteriormente. Las células se clasificaron por primera vez según su perfil FSC y SSC (datos no mostrados). Las células CD19<sup>+</sup> B vivas se clasificaron (panel izquierdo), y se redujeron aún más a células B CD20<sup>-</sup> (2º panel desde la izquierda), y se refinaron a células CD27<sup>+</sup>CD38<sup>++</sup>. A partir de esto, se determinaron los plasmoblastos IgG<sup>+</sup> como IgA<sup>-</sup> e IgM<sup>-</sup>, ya que los plasmoblastos IgG<sup>+</sup> no expresan IgG en la superficie celular. Esta población se clasificó en células individuales en placas de 96 pozos.

15 Ejemplo 3: Los plasmoblastos están presentes en sujetos sometidos a desafío inmunológico.

20 Los plasmoblastos generalmente representan aproximadamente el 0.15% de las células B en donantes sanos, pero pueden variar entre aproximadamente el 3.3% y el 16.4% en sujetos que se enfrentan a una variedad de desafíos inmunológicos que incluyen infecciones (por ejemplo, infecciones por *S. aureus* y *C. diff*), cáncer asociado con no-progresión (por ejemplo, melanoma metastásico y adenocarcinoma metastásico del pulmón en donde los pacientes después de una intervención (quimioterapia en el caso del paciente con adenocarcinoma pulmonar y tratamiento con ipilimumab en el caso del paciente con melanoma metastásico) se convirtieron en no progresivos a largo plazo asociados con una respuesta activa de células B) y vacunas (por ejemplo, influenza).

25 La Figura 23 muestra que los plasmoblastos estaban presentes y se podían obtener de una variedad de sujetos para la secuenciación de alto rendimiento del repertorio de anticuerpos pareados y la caracterización del repertorio de anticuerpos con respuesta humoral activa. Esto demuestra que las metodologías descritas en el presente documento pueden usarse para obtener árboles evolutivos de cadenas pesadas y livianas (H&L) y usar esta información para clonar y/o expresar anticuerpos para, por ejemplo: a) descubrimiento de antígenos novedosos; b) para informar el diseño de la vacuna, por ejemplo, el uso del sistema inmunitario para informarnos de cuáles son los antígenos conocidos y novedosos es probable que sean útiles para la opsonización y la fagocitosis y/o la muerte/ inhibición de un patógeno u objetivo de interés y, opcionalmente, poner eso en el diseño de la vacuna; c) hacer anticuerpos monoclonales neutralizantes, por ejemplo, a partir de vacunas; d) hacer anticuerpos monoclonales de unión; e) fabricar anticuerpos contra patógenos microbianos; y f) fabricar anticuerpos contra el cáncer. Los ejemplos de estos se describen con más detalle a continuación.

Ejemplo 4: La actividad de la enfermedad en CCP+ RA se correlaciona con los plasmoblastos circulantes.

35 Habiendo demostrado que nuestro método se puede usar para la secuenciación de genes de inmunoglobulina en pares identificables, utilizamos nuestro método para investigar el repertorio de anticuerpos de plasmoblastos en pacientes con CCP+ RA. Se obtuvieron muestras de sangre de pacientes con AR con consentimiento y se tiñeron para plasmoblastos mediante citometría de flujo (Figura 5a). Los plasmoblastos circulantes se expresaron como un porcentaje del total de PBMC. Encontramos que los pacientes con CCP+ RA tienen porcentajes de plasmoblastos de sangre periférica significativamente más altos que los pacientes con CCP-RA (Figura 5b). Además, los porcentajes de plasmoblastos en pacientes con CCP+, pero no en pacientes con CCP-, se correlacionaron con la actividad de la enfermedad ( $r = 0.35$  y  $p = 0.028$ ) (Figura 5c).

Ejemplo 5: Los plasmoblastos produjeron anticuerpos anti-CCP.

45 Aunque los pacientes con CCP+ tienen porcentajes de plasmoblastos que se correlacionan con la actividad de la enfermedad, estos pacientes podrían tener una infección en curso u otros factores que elevaron sus porcentajes de plasmoblastos circulantes. Para determinar la especificidad de los plasmoblastos circulantes en pacientes con CCP+, se cultivaron células B enriquecidas con RosetteSep de pacientes en RPMI suplementado con FBS al 10%. Otros suplementos de medios, como anti-IgM, IL-6, BAFF, etc., no se usaron, por lo que los plasmoblastos serían las únicas células que secretan anticuerpos (mientras que otras células B permanecen inactivadas). Para confirmar que solo los plasmoblastos producen anticuerpos, agotamos algunas de las muestras de plasmoblastos (Figura 5d). Las células B se cultivaron luego durante siete días antes de recoger el sobrenadante y ejecutarlo en una matriz de péptidos Luminex. La matriz evalúa la reactividad del anticuerpo a péptidos citrulinados. La reactividad de los anticuerpos estuvo ausente en los sobrenadantes de muestras empobrecidas en plasmoblastos en comparación con los sobrenadantes de células B empobrecidas en forma simulada, lo que sugiere que los plasmoblastos secretan cantidades significativas de autoanticuerpos peptídicos anti-citrulina (Figura 5e). Además, cuando se contaron los péptidos con una intensidad fluorescente media (MFI) superior a 60 para cada muestra, se encontró una fuerte correlación entre los porcentajes

de plasmoblastos circulantes y el número de péptidos a los que reaccionaron los anticuerpos ( $r = 0.90$  y  $p = 0.0139$ ). Se eligió un MFI de 60, ya que este era el umbral por debajo del cual >99% de la reactividad del péptido cae en los sobrenadantes de muestras agotadas en plasmoblasto.

Ejemplo 6: secuenciación 454 y análisis de secuencias.

- 5 Los plasmoblastos de los pacientes se clasificaron en una sola célula en placas de 96 pozos como se describió anteriormente, y su ARN se sometió a transcripción reversa y se amplificó por PCR de acuerdo con "Touchdown PCR" en la sección de materiales y métodos, de modo que contenían ID de muestra (región de identificación de la muestra) y placa -Códigos de barras de ID (región de identificación de la placa), como se describió anteriormente. *Consúltase* la Figura 3. Las secuencias del ADNc se obtuvieron luego de una instalación de -secuenciación 454 (DNA Sequencing Center, Brigham Young University and 454 sequencing center 454, Roche).

- 15 Las secuencias se obtuvieron a partir de una primera secuencia de secuencia de 454 utilizando la tubería de disparo. Se obtuvieron secuencias de calidad aceptable a partir de una segunda secuenciación de 454 a través de un filtro de amplicón modificado de la serie de análisis de datos 454 GS FLX. El filtro de amplicón se modificó para que `<vfScanAllFlows>` se estableciera en "falso", y `<vfBadFlowThreshold>` se cambió a "6". Las secuencias de una tercera y cuarta ejecución se obtuvieron utilizando un filtro de amplicón estándar 454. Las secuencias pasadas por el filtro luego se procesaron como se describe en la sección "Asignación de secuencias a pozos" en los materiales y métodos, y las secuencias en cada pozo se ensamblaron individualmente como se describe en "Ensamblaje de secuencias" en los materiales y métodos. Las secuencias ensambladas se analizaron luego con IMGT HighV-Quest para obtener la identificación de las regiones VDJ utilizadas.

- 20 Después del ensamblaje de la secuencia y la identificación del uso de V(D)J, se utilizó ClustalX para agrupar las secuencias en familias clonales (Figura 6). Alternativamente, las secuencias se pueden ensamblar usando lecturas tanto de avance como de retroceso. En este caso, las secuencias directas subdivididas se ensamblan primero como anteriormente. Las secuencias de avance y de retroceso V(D)J se identifican luego utilizando HighV-QUEST y el subconjunto de secuencias de avance y de retroceso de acuerdo con el uso de la placa ID y V(D)J, y las secuencias de avance y de retroceso se ensamblan utilizando Newbler. Debido a que las secuencias de inmunoglobulina son muy similares, el ensamblaje de un subconjunto más pequeño de secuencias evita que los problemas potenciales de secuencias de diferentes células se emparejen incorrectamente.

Ejemplo 7: Agrupación de secuencias en árboles evolutivos

- 30 Se aislaron células mononucleares de sangre periférica (PBMC) de sujetos humanos con los diagnósticos indicados o después de la vacunación. Los plasmoblastos se clasificaron en -células individuales en pozos individuales en placas de 96 pozos, creando muestras de células individuales en cada pozo, el ARNm de cada pozo se sometió a transcripción reversa y luego los contenidos de los pozos se agruparon y se sometieron a dos rondas de PCR para amplificar el Inmunoglobulina ADNc de cadena pesada y liviana. La transcripción reversa agregó una identificación identificativa de la muestra a todos los ADNc generados a partir de cada muestra individual, y la primera ronda y la segunda ronda de PCR agregaron las -identificaciones de la placa y luego los cebadores A y B de titanio 454 a cada amplicón, respectivamente, como se describe en "Touchdown PCR y PCR no Touchdown" en la sección de materiales y métodos. La metodología general se describe en la Figura 3. Los amplicones agrupados se secuenciaron con la tecnología de secuenciación 454, las lecturas de calidad aceptable se obtuvieron como se describió anteriormente. Las lecturas se asignaron a los pozos y se ensamblaron como se describe en las secciones "Asignación de lecturas a pozo" y "Ensamblaje de secuencias" en los materiales y métodos. Los segmentos V(D)J en las secuencias ensambladas se identificaron luego utilizando HighV-QUEST. Las secuencias de cadenas pesadas y livianas identificadas con códigos de barras compuestos compartidos pueden emparejarse simplemente al juntar las secuencias ensambladas con los códigos de barras compuestos correspondientes.

- 45 Los amplicones de sujetos humanos individuales se agruparon en función de estos segmentos V(D)J, y las secuencias que expresaban los mismos segmentos V(D)J se clasificaron como pertenecientes a la misma familia clonal (Figura 7). Cada gráfico circular representa el porcentaje de clones derivados de plasmoblastos individuales de un sujeto humano individual que expresa segmentos génicos V(D)J idénticos (es decir, porcentaje de clones en cada familia clonal). Los sujetos humanos incluyeron aquellos con sepsis (2 sujetos), artritis reumatoide (3 sujetos), cáncer de pulmón (1 sujeto) y después de la vacunación contra la influenza (1 sujeto). Se eligió a estos sujetos para mostrar que las familias clonales pueden aislarse de los plasmoblastos de los sujetos sometidos a afecciones agudas (sepsis y vacuna contra la gripe) y crónicas (artritis reumatoide y cáncer de pulmón).

- 55 Las secuencias V(D)J de la cadena pesada de inmunoglobulina de los sujetos humanos individuales de la Figura 7 se agruparon usando ClustalX y se mostraron usando Treeview como árboles radiales sin raíz (Figura 8). Cada árbol radial representa las secuencias de la cadena pesada derivadas de un sujeto humano individual. Para cada árbol radial, los extremos terminales representan una secuencia única. Las ramas principales representan familias clonales, y las ramas más pequeñas representan subfamilias clonales que difieren entre sí por mutaciones que surgieron a través de la diversidad de unión (adición de P-nucleótidos o N-nucleótidos, o eliminación de nucleótidos), hipermutación somática y maduración de la afinidad.

Ejemplo 8: Clonación, expresión y purificación de anticuerpos.

5 Todos los anticuerpos seleccionados se clonaron, expresaron y aislaron como se describe anteriormente en la sección de Materiales y Métodos (*Consúltese* las secciones: Clonación de cadenas livianas y pesadas en vectores Lonza; Expresión de anticuerpos monoclonales en 293T; Anti-human IgG ELISA; y Purificación de proteína A-IgG de anticuerpos monoclonales expresados). Los anticuerpos purificados se utilizaron para estudios posteriores, como se explica más abajo.

Ejemplo 9: Caracterización de anticuerpos de sujetos después de la vacunación contra la influenza.

10 Los anticuerpos de humanos a los que se administró una vacuna contra la influenza se seleccionaron y aislaron como se describió anteriormente. Los anticuerpos seleccionados para una caracterización adicional a continuación se indican en las secciones apropiadas.

#### *ELISA de Fluzone*

15 Se les administró a voluntarios la vacuna contra la gripe de la temporada 2010/2011 de Fluzone, que consta de 3 cepas de virus inactivados, las cepas A/California/7/2009, A/Perth/16/2009, B/Brisbane/60/2008. Los ELISA de Fluzone se realizaron como se describió anteriormente para determinar si los anticuerpos monoclonales derivados de voluntarios vacunados se unen a la vacuna contra la gripe en sí como una prueba inicial de anticuerpos expresados con actividad de unión. 14 de los 31 anticuerpos unidos a Fluzone ELISA (Figura 26) y un subconjunto de estos se seleccionaron posteriormente y se probaron para determinar la actividad de unión a la hemaglutinina utilizando resonancia de plasmón de superficie. Los anticuerpos caracterizados por Fluzone ELISA fueron: Flu14 - Flu23, Flu25 - Flu27, Flu29, Flu30, Flu34, Flu35, Flu37, Flu39 - Flu41, Flu43 - Flu46. S1 y S2 se utilizaron como controles negativos.

20 *Determinación por resonancia de plasmón superficial de afinidades de anticuerpos contra la gripe.*

25 La unión de anticuerpos monoclonales (mAbs) a moléculas de HA se analizó a 25°C utilizando un biosensor de resonancia de plasmón de superficie ProteOn (BioRad Labs) como se describió anteriormente. De los 14 anticuerpos que se unieron a Fluzone ELISA, 10 se unieron a H3 y 1 se unió a H1, mientras que 3 no se unieron (Figura 27). Uno de los no aglutinantes, el aglutinante H1 y otros 4 aglutinantes H3 elegidos al azar fueron seleccionados y enviados a una organización de investigación contratada (CRO) para probar la actividad de neutralización en un ensayo de microneutralización. Los anticuerpos caracterizados por SPR fueron: Flu14 - Flu22, Flu26, Flu29, Flu34, Flu35, Flu46.

#### *Ensayo de microneutralización de influenza*

30 Algunos de los anticuerpos contra la gripe expresados que mostraron actividad de unión en el ELISA de Fluzone se enviaron a un CRO externo, Virapur LLC, para los ensayos de microneutralización como se describió anteriormente. Los resultados de los ensayos mostraron que el anticuerpo que se unía a H1 en ensayos anteriores neutralizaba H1, mientras que los anticuerpos que se unían a H3 en ensayos anteriores neutralizaban H3. El no aglutinante no neutralizó el virus de la influenza (Figura 28). Los anticuerpos caracterizados por el ensayo de microneutralización fueron Flu15, Flu16, Flu18, Flu19, Flu20, Flu21.

#### *Variación de CDR*

35 Los anticuerpos contra la gripe se obtuvieron como se describe anteriormente. La Figura 25 muestra un dendrograma parcial en explosión para mayor claridad (a). Las familias clonales son claramente visibles y la familia clonal sombreada tiene la V(D)J asignada como se muestra en la caja gris. La secuencia de aminoácidos a través de las CDR (región en caja) para las cadenas pesada y liviana se muestran en la Figura 25 (b) y (c) respectivamente, mostrando algunas diferencias de residuos entre las cadenas.

40 Estos resultados anteriores demuestran que los árboles evolutivos se pueden obtener utilizando las composiciones y los métodos descritos aquí. Se pueden aislar anticuerpos monoclonales completamente humanos a partir de células B activadas, tales como plasmoblastos, de sujetos que experimentan afecciones agudas usando las composiciones y métodos descritos en este documento. Estos anticuerpos monoclonales completamente humanos también pueden ser anticuerpos neutralizantes usando las composiciones y métodos descritos en este documento. Los resultados también demuestran que las composiciones y los métodos descritos en el presente documento pueden usarse para aislar mAbs dirigidos contra antígenos extraños.

Ejemplo 10: Caracterización de anticuerpos de sujetos con AR.

50 Los anticuerpos de humanos que padecen artritis reumatoide (AR) se seleccionaron y aislaron como se describió anteriormente. Los anticuerpos seleccionados para una caracterización adicional a continuación se indican en las secciones apropiadas.

#### *Reactividades de anticuerpos de AR en micromatrices de antígeno de AR*



Los anticuerpos derivados de pacientes con AR se probaron en una matriz de antígenos de AR y se analizaron las fluorescencias con una máquina GenePix como se describe en la sección "Reactividades de los anticuerpos de la RA en matrices de antígenos de la AR" en materiales y métodos. Las relaciones identificadas se mostraron como un mapa de calor utilizando el software Java TreeView (Figura 37). Los anticuerpos caracterizados por este ensayo fueron:

5 RA1, RA2, RA3, RA4, RA8 - RA13, RA16, RA19, RA22 y RA23. Flu14 y Flu26 se utilizaron como controles negativos.

ELISA de Antihistona 2A

Para la detección de anticuerpos contra H2A, se realizó un ELISA directo como se describe en "Anti-histona 2A ELISA" en la sección de materiales y métodos. La Figura 35a muestra los valores de absorbancia detectados para cada anticuerpo probado. Los anticuerpos caracterizados en la Figura 35a y en el ELISA de anti-CCP2 a continuación

10 fueron: RA1, RA2, RA4-RA16, RA19, RA23-RA24. La Figura 36 muestra los anticuerpos seleccionados (RA1, RA2, RA8, RA9) en otro ELISA independiente utilizando 30 ug/ml de anticuerpos.

ELISA de Anti-CCP2

Anti-CCP2 ELISA se realizó como se describe en "Anti-CCP2 ELISA" en la sección de materiales y métodos. La Figura 35b muestra los valores de absorbancia detectados para cada anticuerpo probado.

15 ELISA de Factor Anti-Reumatoide

Los anticuerpos derivados de pacientes con AR se utilizaron como el anticuerpo primario en un ELISA directo y se usó anti-IgG-HRP humana como el anticuerpo secundario, y se visualizaron con el sustrato TMB. Para la detección de anticuerpos contra el factor reumatoide (RF), el ELISA Anti-RF se realizó como se describe en el ELISA "factor anti-reumatoide" en la sección de materiales y métodos. La Figura 34 muestra que los anticuerpos RA2 y RA3 mostraron

20 reactividad. Los anticuerpos caracterizados aquí fueron: RA1-RA6, RA8-RA12, RA14.

Estos resultados anteriores demuestran que los anticuerpos se pueden aislar de células B activadas, como los plasmoblastos, de sujetos que experimentan afecciones crónicas utilizando las composiciones y los métodos descritos en este documento. Los resultados también demuestran que las composiciones y los métodos descritos en el presente documento se pueden usar para aislar mAbs dirigidos contra antígenos propios.

25 Ejemplo 11: Caracterización de anticuerpos de sujetos con cáncer de pulmón.

Los anticuerpos de un humano no progresivo a largo plazo que sufría de adenocarcinoma de pulmón metastásico se seleccionaron y aislaron como se describió anteriormente. Este humano desarrolló un adenocarcinoma de pulmón metastásico y se esperaba que sucumbiera al cáncer; sin embargo, después de la quimioterapia, este paciente entró en un estado de no progresión a largo plazo durante más de 4 años que se asoció con plasmoblastos que constituyen

30 el 3.1% de todas las células B de sangre periférica. Los niveles elevados de plasmoblastos en sangre periférica en este paciente indicaron que una respuesta inmune en curso podría estar contribuyendo a su no progresión a largo plazo. Los siguientes anticuerpos se seleccionaron para una caracterización adicional a continuación: LC1, LC5-LC7, LC9-LC18. Se utilizó Flu16 como control negativo.

35 *Inmunohistoquímica de anticuerpos de pacientes con adenocarcinoma de pulmón en arreglos de tejidos de cáncer de pulmón*

La inmunohistoquímica utilizando dos tipos diferentes de portaobjetos de micromatrices de tejido se realizó como se describe en "Inmunohistoquímica de anticuerpos de pacientes con adenocarcinoma de pulmón en arreglos de tejidos de cáncer de pulmón" en la sección de materiales y métodos. Nuestros resultados demostraron uno de los anticuerpos expresados unidos al adenocarcinoma de pulmón (Figura 32).

40 *Determinación por citometría de flujo de la unión de los anticuerpos expresados en pacientes de adenocarcinoma de pulmón a líneas celulares de adenocarcinoma de pulmón*

La unión de anticuerpos a varias líneas celulares de cáncer de pulmón se realizó como se describe en "Determinación por citometría de flujo de la unión de anticuerpos expresados desde pacientes de adenocarcinoma de pulmón a líneas celulares de cáncer de pulmón" en la sección de materiales y métodos. Nuestros resultados mostraron que un anticuerpo se unió a las líneas celulares de adenocarcinoma de pulmón y puede ser específico para los adenocarcinomas de pulmón (Figura 33).

Estos resultados anteriores demuestran que los anticuerpos pueden aislarse de células B activadas, como los plasmoblastos, de sujetos que experimentan afecciones crónicas como el cáncer utilizando las composiciones y los métodos descritos en este documento. Los resultados también demuestran que las composiciones y los métodos descritos en el presente documento se pueden usar para aislar mAbs dirigidos contra antígenos propios.

50

Ejemplo 12: Caracterización de anticuerpos de sujetos con infección por *Staphylococcus aureus*.

Los seres humanos con infecciones por *S. aureus*, incluido un humano con osteomielitis crónica por *S. aureus* con control inmune mediado de la infección en ausencia de antibióticos, se utilizaron como fuentes de sangre periférica a partir de la cual se tiñeron y clasificaron los plasmoblastos de sangre periférica. El procesamiento de ADNc con código de barras, secuenciación 454 y análisis bioinformático generó árboles evolutivos de repertorios de anticuerpos en humanos que montan respuestas inmunitarias efectivas contra *Staph. aureus*. Los anticuerpos de humanos que montan respuestas inmunitarias efectivas contra la infección por *S. aureus* se seleccionaron y aislaron como se describió anteriormente. Los anticuerpos seleccionados para una caracterización adicional a continuación se indican en las secciones apropiadas.

5

Citometría de Flujo de Estafilococos

10 Se usaron anticuerpos anti-estafilococos para teñir *S. aureus* fijado como se describe en "Citometría de flujo de estafilococos" en los materiales y métodos. Nuestros resultados mostraron que los anticuerpos S6 y S11 se unen a la superficie de *S. aureus* y pueden ser candidatos para opsonización, lo que resulta en fagocitosis y muerte/ inhibición de *S. aureus* (Figura 29). Los anticuerpos caracterizados en este ensayo fueron: S1 - S4, S6 - S13, con F26 como control negativo.

15 Ensayo de Inhibición de Estafilococos

*S. aureus* en el crecimiento de la fase logarítmica se combinó con un anticuerpo anti-estafilococo para determinar la actividad inhibitoria de los anticuerpos como se describe en "Ensayo de inhibición de estafilococos" en los materiales y métodos. Nuestros resultados demuestran que varios de los anticuerpos clonados y expresados mostraron una potente actividad de inhibición/muerte en *S. aureus* (Figura 30). Los anticuerpos caracterizados por este ensayo fueron S6 y S9, con LC1 como control negativo.

20

Inmunoprecipitación de antígenos estafilococos con anticuerpos derivados de pacientes infectados por estafilococos

Se usaron anticuerpos para inmunoprecipitar diversos antígenos estafilococos como se describe en "Inmunoprecipitación de antígenos estafilococos con anticuerpos derivados de pacientes infectados con estafilococos" de los materiales y métodos. Las proteínas inmunoprecipitadas se identificaron luego con espectrometría de masas como se describe a continuación. Los anticuerpos caracterizados por este ensayo fueron S1 - S13.

25

Identificación por espectrometría de masas de péptidos.

Las bandas teñidas de proteínas de interés se seleccionaron y se sometieron a espectrometría de masas como se describe en "Identificación por espectrometría de masas de péptidos" de los materiales y métodos. Los resultados identificaron el péptido modulín alfa 1 soluble en fenol o la delta-hemolisina como la diana de unión probable para el anticuerpo S4. Esto demuestra que los métodos descritos en este documento se pueden usar para realizar un nuevo descubrimiento de antígenos (Figura 31). El anticuerpo caracterizado por este ensayo fue S4.

30

Estos resultados anteriores demuestran que los anticuerpos pueden aislarse de células B activadas, como los plasmoblastos de sujetos que sufren afecciones agudas, como una infección bacteriana, utilizando las composiciones y los métodos descritos en este documento. Los resultados también demuestran que las composiciones y métodos descritos en el presente documento se pueden usar para aislar mAbs dirigidos contra antígenos extraños y para determinar la identidad de los antígenos unidos por anticuerpos seleccionados.

35

Ejemplo 13: Células blásticas y caracterización de plasmoblastos.

Las secuencias de inmunoglobulina de las células B que son activadas por una respuesta inmune en curso pueden usarse para producir un árbol evolutivo de una respuesta inmune en curso, como se describió anteriormente. Este árbol evolutivo se caracteriza típicamente por múltiples familias clonales que representan células B activadas desde múltiples líneas de descendencia. Las secuencias de las células B ingenuas generalmente no podrán usarse para producir un árbol evolutivo de este tipo, ya que no se han activado y, por lo tanto, proporcionan poca o ninguna información sobre la respuesta inmune activa y en curso. Las células B activadas primero se convierten en células blásticas, que se activan y son de mayor tamaño. Estas células blásticas luego se convierten en células B de memoria o en células plasmáticas. En los seres humanos, aunque las células de memoria B y las células plasmáticas son el resultado de una respuesta inmune, se unen a grandes grupos de células de memoria B y células plasmáticas que han resultado de las respuestas a insultos inmunológicos anteriores, lo que dificulta la distinción entre las células de memoria B y las células plasmáticas. o respuestas inmunitarias previas. Por lo tanto, en los seres humanos, las células blásticas son un candidato preferido para la secuenciación para obtener árboles evolutivos de una respuesta inmune en curso. En la investigación, los animales criados en condiciones controladas (por ejemplo, ratones), sin embargo, las células B blásticas, las células B de memoria y las células plasmáticas son candidatas a la secuenciación para obtener árboles evolutivos a medida que se crían en un ambiente limpio, lo que hace posible que la mayoría de las células de memoria B y las células plasmáticas después de una respuesta inmune rigurosa para estar en contra de la agresión, especialmente después de las inyecciones de refuerzo, ya que no deberían tener una gran memoria o poblaciones de células plasmáticas que hayan visto algún desafío inmunológico importante anteriormente.

40

45

50

55

De manera similar para las células T, en los seres humanos, las células preferidas para secuenciar para obtener un árbol evolutivo de una respuesta inmune en curso serán las células T blásticas. Para ratones, las células T activadas, blásticas y de memoria son todas candidatas preferidas para secuenciar para obtener un árbol evolutivo.

5 Se sabe que las células B blásticas son más grandes que las células B típicas. El tamaño de un linfocito pequeño, de los cuales una célula B en reposo es uno, es típicamente de entre 6 y 8  $\mu\text{m}$  de tamaño. Los linfocitos blásticas (células T y B) suelen tener un tamaño de entre 8 y 14  $\mu\text{m}$ . (Consúltese la Figura 41, también Tzur et al, PLoS ONE, 6(1): e16053. doi:10.1371/journal.pone.0016053, 2011; Teague et al, Cytometry, 14:7 2005). Los plasmoblastos pueden tener el siguiente patrón de expresión: CD19<sup>bajo/+</sup>, CD20<sup>bajo/-</sup>, CD27<sup>+</sup> y CD38<sup>alto</sup>. Aunque el uso de todos estos marcadores resulta en la población más pura para la selección de células individuales, no es necesario usar todos los marcadores anteriores para aislar los plasmoblastos.

10 Como se ejemplifica en la Figura 39, los plasmoblastos pueden activarse mediante el uso de un FSC<sup>hi</sup> para células más grandes, lo que da como resultado una población de plasmoblastos con un 37% de pureza. La activación de las células FSC<sup>hi</sup>CD19<sup>hi</sup> proporciona un 72% de pureza de plasmoblastos. La activación de FSC<sup>hi</sup> y CD27<sup>+</sup>, CD38<sup>hi</sup>, o CD20<sup>-</sup> da 44, 80 y 71 por ciento de pureza de plasmoblastos, respectivamente. Se puede usar la combinación de cualquiera de estos marcadores u otros marcadores que puedan distinguir los plasmoblastos de otras células B para aumentar la pureza de los plasmoblastos clasificados, sin embargo, cualquiera de estos marcadores solo puede distinguir los plasmoblastos de otros linfocitos B, aunque con una menor pureza.

Ejemplo 14: Plataforma alternativa para secuenciación y análisis.

20 Las lecturas de la cadena pesada de una sola placa, la placa 44, se prepararon para la secuenciación de PacBio utilizando los métodos descritos en "Touchdown PCR" en los materiales y métodos para amplificar el ADNc de la cadena pesada gamma. Se realizaron 48 2<sup>o</sup> PCR para obtener suficiente ADN para la ejecución de PacBio. La agrupación y limpieza del ADN se realizó como se describe en "Preparación para la ejecución de la secuenciación de PacBio". El ADN fue enviado a PacBio para la preparación y secuenciación. Las lecturas de CCS se obtuvieron de PacBio y se asignaron a los pozos y se ensamblaron de acuerdo con la "Asignación de secuencias a los pozos" y el "Ensamblaje de secuencias" en los materiales y métodos. Los resultados de la asignación se encuentran en la Figura 38. Esto muestra que nuestros métodos y composiciones no son específicos de la plataforma para la secuenciación de alto rendimiento.

Ejemplo 15: Secuenciación y análisis en ejecuciones 454 XL+

30 La secuenciación puede adaptarse a las ejecuciones 454 XL+ siguiendo el método descrito en "Preparación para la ejecución de la secuencia 454 XL+". Esto se debe hacer ya que las ejecuciones 454 XL+ actualmente solo admiten la química Lib-L, mientras que nuestras ejecuciones 454 XLR70 utilizan la química Lib-A. En general, esto también puede adaptarse a situaciones en las que se prefiere la química Lib-L a la química Lib-A típica para la secuenciación de amplicones en series XLR70. Las lecturas de 454 XL+ aún pueden asignarse a pozos y ensamblarse siguiendo los métodos descritos en "Asignación de secuencias a pozos" y "Ensamblaje de secuencias". Las lecturas de XLR70 y XL+ después del filtrado 454 se pueden usar de manera idéntica, es decir, la selección corriente abajo de los anticuerpos para la clonación y la expresión y el ensayo de las propiedades funcionales del anticuerpo todavía pueden proceder según las Figuras 6 y 9.

Ejemplo 16: Clonación de genes de inmunoglobulina apareados.

40 Suponiendo que cada familia clonal reconoce el mismo epítipo, y que la variación de la secuencia dentro de cada familia se debe a una hipermutación somática, primero se pueden clonar y expresar el clon más frecuente de cada familia clonal para la detección de anticuerpos que se unen al antígeno de interés (Figura 6). Utilizamos el clon más frecuente porque, durante la maduración de la afinidad y la selección en el centro germinal, los centrocitos que se unen al antígeno con la afinidad más alta compiten con otros centrocitos por los factores de supervivencia. Por lo tanto, se espera que el clon de frecuencia más alta también tenga la afinidad de enlace más alta. Una vez que se ha identificado un clon como un anticuerpo capaz de unirse a un antígeno, se clonan, expresan y analizan las secuencias de inmunoglobulina pareadas representativas de la familia clonal completa (Figura 6). Este proceso puede involucrar la clonación y expresión de secuencias que representan múltiples subclones dentro de la familia clonal, o codificar anticuerpos de diferentes isotipos dentro de la familia clonal, para permitir la prueba directa y la comparación de las propiedades de unión y funcionales de clones específicos que representan el espectro de anticuerpos contenidos dentro de la familia clonal. El clon específico que exhibe la unión deseada y las propiedades funcionales se seleccionan luego para una caracterización y consideración adicionales para el desarrollo como un anticuerpo humano terapéutico.

55 Un enfoque alternativo para seleccionar candidatos para la clonación de una familia (o cualquier otro conjunto de anticuerpos de interés) es construir un árbol filogenético para los anticuerpos (enraizado en la secuencia de la línea germinal en el caso de una familia clonal). Los nodos de la hoja en un árbol de este tipo corresponden a los miembros de la familia de anticuerpos. Los candidatos para la clonación se seleccionan descendiendo desde la parte superior del árbol, siempre eligiendo la rama con el mayor número de nodos de hojas debajo (eligiendo aleatoriamente en caso

de un empate), luego en el último nodo sobre las hojas eligiendo la hoja con mayor número de mutaciones, o elección aleatoria en caso de empate. Los candidatos adicionales podrían seleccionarse, si se desea, seleccionando a los candidatos repetidamente hasta que se logre el número deseado, de la siguiente manera. Para cada nodo en el árbol, si no se ha seleccionado ninguna de las hojas que son descendientes del nodo, cuente el número de hojas que son descendientes. Para el nodo con el mayor recuento de este tipo (selección aleatoria en caso de empate), descendiendo, eligiendo siempre la rama con el mayor número de nodos de hojas descendientes (selección aleatoria en caso de empate entre ramas). Luego, en el último nodo sobre las hojas, elija la hoja con el mayor número de mutaciones o elija al azar en caso de un empate.

Otro enfoque más para seleccionar candidatos de una familia de anticuerpos (o cualquier otro conjunto de anticuerpos de interés) es enumerar los anticuerpos por número descendente de mutaciones no silenciosas relativas a la línea germinal y seleccionar de la lista en orden, escogiendo así los anticuerpos que más evolucionan.

Ejemplo 17: Transfección permanente y expresión de anticuerpos humanos candidatos.

Los clones deseados se amplifican selectivamente a partir de una placa combinada de secuencias mediante el uso de cebadores de clonación específicos para un ID de muestra dado; estos cebadores también incorporan diferentes sitios de restricción 5' y 3' en el clon. Los sitios de restricción se utilizan luego para insertar el clon en vectores. Debido a que los clones amplificados pueden contener solo una secuencia de región constante parcial, los vectores ya contienen las regiones constantes kappa, lambda o gamma con los sitios de restricción apropiados necesarios para insertar los clones amplificados en el marco de lectura abierto. Los sitios de restricción múltiple se diseñan en el vector, ya que los clones tienen secuencias variables, para evitar el problema potencial del sitio de restricción que también existe en el propio clon, que luego se reduciría por la enzima de restricción. Esto permite insertar tantos clones como sea posible. Los vectores utilizados son dos vectores separados con diferentes marcadores seleccionables de mamíferos (Invitrogen pADNc3.3 modificado o vectores pOptivec que contienen un gen de región constante con sitios de restricción diseñados por ingeniería genética) o un vector de expresión dual que contiene ambos genes (sistema Lonza GS; pEE6.4 y pEE12.4). Consúltense las Tablas 16 y 17 respectivamente para ver las secuencias de las inserciones de región constante. Los marcadores de selección son amplificables, como la dihidrofolato reductasa (DHFR) en pOptivec o glutamina sintetasa (GS) en el sistema Lonza GS, para permitir la amplificación de genes y la producción eficiente de anticuerpos para propósitos de detección adicionales que requieren grandes cantidades de anticuerpo (por ejemplo, una criba *in vivo*). Las células de mamífero se transfectan utilizando una doble transfección, con una cadena liviana en uno y la cadena pesada en el otro vector (pOptivec modificado y pADNc3.3), o un vector de expresión dual (sistema Lonza GS) que contiene ambos genes.

*Vectores modificados de Invitrogen.* Los vectores son dos vectores separados con diferentes marcadores seleccionables de mamíferos y sitios de restricción diseñados. pADNc3.3 tiene un gen de resistencia a neomicina como marcador seleccionable, y pOptivec tiene un marcador -seleccionable por DHFR. CHO DHFR: las -células se cotransfectan con pADNc3.3 modificado y pOptivec bajo selección de Geneticin. Solo las células DHFR transfectadas con pOptivec, que contiene una copia de DHFR, sobrevivirán, y el gen de resistencia a neomicina en pADNc3.3 confiere resistencia a Geneticin, lo que permite la selección de células que se transfectan con éxito con ambos vectores (que contiene la cadena liviana en un vector y una cadena pesada en el otro vector), y por lo tanto producirá inmunoglobulinas funcionales.

*Sistema Lonza GS.* El sistema Lonza GS utiliza los vectores pEE12.4 y pEE6.4. El vector pEE12.4 contiene el gen GS como marcador de selección amplificable, y pEE6.4 es un vector complementario. La cadena liviana se clonará en uno de los vectores y la cadena pesada en el otro vector. Posteriormente, ambos vectores se cortan con enzimas de restricción y se ligan para formar un solo vector que puede expresar genes de cadena pesada y liviana en promotores separados. Por lo tanto, es un sistema de vector de expresión dual, que permite la expresión de ambos genes de un solo vector. Las células CHO se transfectan con el vector de expresión dual bajo la selección de metionina sulfoximina. Las células transfectadas se seleccionan para.

Amplificación de genes

Tanto la dihidrofolato reductasa (DHFR) como la GS son marcadores de selección amplificables. Bajo la presión de la selección de cantidades crecientes de metotrexato y metionina sulfoximina respectiva, las líneas celulares transfectadas que tienen regiones genómicas duplicadas que contienen los genes DHFR y GS sobrevivirán porque son más resistentes a los reactivos de selección. Los genes cerca de los marcadores de selección, como los genes de inmunoglobulina de cadena pesada y liviana insertados, también se amplifican, lo que da como resultado copias genéticas más altas y mayores tasas de producción de inmunoglobulinas. Los clones que producen anticuerpos que se ha encontrado que tienen propiedades neutralizantes en las cribas *in vitro* (véase más adelante) se amplifican para que se puedan obtener más anticuerpos para estudios *in vivo* posteriores.

Ejemplo 18: Identificación de la especificidad de los anticuerpos humanos expresados

La detección de anticuerpos se produce en dos etapas. Estamos utilizando un novedoso proceso de "detección selectiva", en donde primero seleccionamos las familias clonales apropiadas para utilizarlas en la detección de

anticuerpos neutralizantes. Seleccionamos los 1-3 clones más frecuentes de cada familia clonal para determinar su capacidad para unirse al antígeno. Nuestra criba generalmente toma el formato de un ELISA indirecto, aunque la citometría de flujo puede usarse para identificar anticuerpos de unión a células. Esto comprende primero unir el antígeno apropiado a una placa ELISA, luego incubarlo con sobrenadantes que contienen los anticuerpos expresados. Los anticuerpos que se unen al antígeno son detectados por un anticuerpo secundario específico.

Una vez que se han identificado los anticuerpos de unión, toda la familia clonal de ese clon se clona y se expresa en la etapa de selección de la "criba de selección". Aunque se espera que todos los anticuerpos en una familia clonal se unan al mismo epítipo, pueden diferir ligeramente en la avidéz de la unión al antígeno y en su posicionamiento sobre el antígeno, diferencias que pueden afectar las propiedades de unión y/o la capacidad de neutralización de los anticuerpos; así, en la mayoría de los casos, varios anticuerpos diferentes (que poseen pequeñas diferencias en sus regiones CDR3) se expresan y se analizan para determinar sus propiedades de unión y neutralización.

Para neutralizar los anticuerpos que se dirigen a un par específico de ligando/receptor, las células 293T primero se transfectan de manera estable con un constructo indicador de la vía de señalización, como un plásmido que contiene el gen de la luciferasa unido a elementos de respuesta de transcripción NF-kB. La activación de NF-kB en la célula transfectada induce la expresión de luciferasa, cuyos niveles se pueden determinar en un ensayo de luciferasa. Esto mide la señalización de NF-kB activada por la unión del ligando-receptor. NF-kB es el elemento de señalización elegido porque la mayoría de los eventos de señalización activan NF-kB. Para ensayar otras vías de señalización, la región promotora del gen de luciferasa contiene el sitio de unión transcripcional apropiado, como el de AP-1, por ejemplo. A continuación, las células 293T se transfectan con el receptor diana. Las células 293T luego se incuban con el ligando y los anticuerpos de unión de interés en placas de 96 pozos. Después de 24 o 48 horas, se realiza un ensayo de luciferasa para determinar la expresión del gen de luciferasa. Los pozos con anticuerpos neutralizantes tienen una expresión de luciferasa mínima o nula. Los resultados se verifican mediante inmunotransferencia Western para proteínas de señalización fosforiladas en la ruta de señalización de NF-kB. Un anticuerpo neutralizante previene la señalización del ligando-receptor; y en consecuencia deroga la fosforilación de las proteínas de señalización.

Para los antígenos presentes en las células vivas, como los antígenos del cáncer y los antígenos bacterianos, el ensayo de neutralización *in vitro* toma la forma de un ensayo que detecta células vivas/muertas, y se puede realizar en un formato de alto rendimiento. Las células cancerosas o las bacterias se incuban en placas de 96 pozos con un anticuerpo candidato. Una mancha que distingue las células vivas de las células muertas y es compatible con la citometría de flujo puede aplicarse a cada pozo. Las células vivas y muertas se tiñen con diferentes fluoróforos y se analizan mediante flujo para obtener porcentajes de células vivas y muertas. Los anticuerpos que pasan la criba *in vitro* se examinarán *in vivo* por su actividad neutralizadora.

El ensayo de neutralización *in vitro* del virus se puede realizar utilizando un ensayo de neutralización de placa estándar. Al realizar ensayos de neutralización de placa en placas de 96 pozos, se puede obtener una imagen de cada pozo usando un microscopio y el conteo de placa se puede automatizar con el software de análisis de imágenes. Los anticuerpos neutralizantes reducen la formación de placa. A continuación, estos anticuerpos se exploran adicionalmente *in vivo* para determinar la actividad neutralizadora.

Consúltese la sección de ejemplo 9 "Fluzone ELISA", la sección de ejemplo 11 "Citometría de flujo ..." y la sección de ejemplo 12 "Citometría de flujo de estafilococos" para ensayos exitosos de actividad de unión usando ELISA (ejemplo 7) y citometría de flujo (ejemplos 9 y 10). Consúltese la sección del ejemplo 9 "Ensayo de microneutralización de la influenza" para un ensayo exitoso de anticuerpos con actividad neutralizante.

#### Ejemplo 19: Secuenciación de células B con más de una célula por pozo

Las muestras individuales que tienen múltiples células B se transcriben por separado en contenedores. La transcripción reversa agrega una ID de muestra y una región de cebador universal 5' a todos los ADNc de la primera cadena. El ADNc de todos los contenedores de un conjunto de contenedores se agrupan y se someten a 2 rondas de PCR. Los pasos son como se describen en "Touchdown PCR y PCR no Touchdown", "Preparación para la ejecución de la secuencia XLR70 454" en los materiales y métodos. Las secuencias para los cebadores también se muestran en la Figura 9. Tenga en cuenta que independientemente de qué gen se amplifique, los cebadores directos permanecen constantes (b). Después de RT y 2 PCR, los amplicones de todos los conjuntos de contenedores se agrupan y se secuencian en 454. La asignación a los pozos y el ensamblaje de secuencias siguen el protocolo descrito en "Asignación de secuencias a los pozos" y "ensamblaje de secuencias" en los materiales y métodos. La combinación de ID de placa e ID de muestra permite la identificación de secuencias que se originan en la misma muestra.

Aunque hay varias células en un pozo, se pueden emparejar cadenas pesadas individuales con cadenas livianas. Las cadenas pesadas de las células B derivadas de un progenitor común estarán relacionadas clonalmente, al igual que las cadenas livianas. Por lo tanto, se pueden asociar una familia clonal de cadena pesada a una familia clonal de cadena liviana mediante la observación de la correlación a través de los pozos. Una vez que se establece una asociación entre las cadenas pesadas de una familia clonal y las cadenas livianas de una familia clonal, se asignan pares en cada pozo seleccionando la cadena pesada que es miembro de la familia clonal de la cadena pesada y una cadena liviana que es un miembro de la familia clonal de la cadena liviana. La selección del par no es ambigua cuando

solo una instancia de la familia de cadenas pesadas y una sola instancia de la familia de cadenas livianas están presentes en un pozo. Después de determinar qué cadenas pesadas y livianas están asociadas entre sí, se pueden dibujar árboles evolutivos y seleccionar los anticuerpos para la caracterización posterior de sus propiedades funcionales.

#### 5 Ejemplo 20: Secuenciación de células B con una o más células por pozo

Las muestras podrían clasificarse con una célula B por pozo en algunas placas, y múltiples células B por pozo en otras placas, aunque las cadenas pesadas y livianas aún podrían emparejarse para aquellos pozos que tienen más de una célula B. Examinamos las secuencias generadas por el paciente de vacunación contra la gripe del ejemplo 9 anterior, donde algunos pozos tenían más de un conjunto de secuencia de cadena pesada distinto o se observó más de un conjunto de secuencia de cadena liviana distinto. RT, PCR, secuenciación y asignación a los pozos y el ensamblaje de secuencias siguieron el protocolo del ejemplo 9 anterior. Para determinar qué cadenas pesadas y livianas estaban asociadas entre sí, las cadenas pesadas se asignaron a familias clonales agrupando todas las cadenas pesadas con el mismo uso de los genes V y J, y el mismo número de nucleótidos entre el final del segmento del gen V y el inicio del segmento del gen J. Las cadenas livianas se asignaron a familias clonales agrupando todas las cadenas livianas con el mismo uso de los genes V y J, y el mismo número de nucleótidos entre el final del segmento del gen V y el comienzo del segmento del gen J. Las relaciones de emparejamiento entre cadenas pesadas y livianas se asignaron primero para los pozos con exactamente una cadena pesada y una cadena liviana, basándose en que comparten un pozo (es decir, que tienen los mismos códigos de barras compuestos). Luego, se calculó una puntuación para cada posible emparejamiento de una familia clonal de cadena pesada con una familia clonal de cadena liviana. La puntuación se determinó contando el número de veces que un miembro de la familia de la cadena pesada y la familia de la cadena liviana comparten un pozo. Luego, cada familia de cadenas pesadas se asoció con la familia de cadenas livianas con la que se obtuvo la puntuación más alta, o la familia de cadenas pesadas no se asoció a una familia de cadenas livianas si la puntuación más alta se logró con más de una familia de cadenas livianas. Las cadenas pesadas y livianas individuales se emparejaron comenzando con la familia general de cadenas pesadas con la puntuación más alta, y avanzaron bien a través de la familia que asignaba pares, y luego continuaron con la siguiente familia de cadenas pesadas. Para una familia de cadenas pesadas dada, para cada pozo, si había una sola cadena pesada dentro del pozo que era miembro de la familia de cadenas pesadas, entonces la cadena liviana de ese pozo que pertenecía a la familia de cadenas livianas asociadas de la familia de cadenas pesadas era asignada a ser el par de la cadena pesada. Si existía más de una cadena liviana, no se asignó ningún emparejamiento. Este proceso de asociación de cadenas pesadas con cadenas livianas continuó hasta que todas las familias y todas las cadenas dentro de esas familias hayan sido consideradas. Si, para una cadena pesada o liviana dada, el proceso dio lugar a más de un candidato para el emparejamiento, se descartaron tanto la cadena pesada como la liviana. Se generaron árboles evolutivos a partir de las cadenas emparejadas y se seleccionaron anticuerpos para la caracterización posterior de sus propiedades funcionales. Una parte del árbol evolutivo se muestra en la Figura 25A.

#### 35 Ejemplo 21: Uso de plasmoblastos seleccionados para generar anticuerpos monoclonales humanos

Desde un sujeto con una afección reciente o actual que resulta en una generación aguda, subaguda o en curso de plasmoblastos circulantes, la citometría de flujo se realiza en sangre periférica (sangre entera o células mononucleares de sangre periférica (PBMC)) para identificar la población de plasmoblastos. Esta población de células B se clasifica luego por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la AARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. Durante la RT, se añaden a la reacción oligonucleótidos adaptadores de ID de muestra bien específicos. Estos adaptadores tienen secuencias de códigos de barras bien específicas (ID de muestra) que pueden identificar secuencias como originadas en diferentes pozos. Utilizando la actividad de cambio de plantilla y de cola 3' de las transcriptasas reversas MMLV H<sup>-</sup>, se agregan las ID de muestra al extremo 3' del ADNc de la 1.<sup>a</sup> hebra. ADNc de cada placa se combinan juntos. Durante la primera ronda de PCR, un cebador largo FW específico de la placa 1 agrega la ID de la placa al extremo 5' de los amplicones. Por lo tanto, un cebador largo FW 1 con diferentes ID de placa se agrega a diferentes placas, lo que da a cada producto de PCR una secuencia de código de barras de identificación. Los cebadores reversos específicos del gen se utilizan para amplificar las cadenas kappa, lambda y gamma, son kappa GSP1, lambda GSP1 y gamma GSP1 respectivamente. Estos cebadores se unen a la región constante de los genes de inmunoglobulina. Los productos de la primera ronda de PCR se diluyen y se utilizan para una segunda PCR anidada. Se usa el cebador FW 2 como el cebador directo y los cebadores reversos kappa, lambda y gamma GSP. Los cebadores largos se usan para amplificar sus respectivos amplicones. En particular, el GPS primer 2 para cada placa agrega una ID de placa común al extremo 3' de cada amplicón para cada placa, por lo que cada una tendrá dos ID de placa y un código de barras de ID de muestra. Más detalles para RT, 1.<sup>a</sup> y 2.<sup>a</sup> PCR se encuentran en "PCR no Touchdown" en los materiales y métodos. Las placas múltiples se agrupan luego de acuerdo con el método detallado en "preparación para la secuenciación 454 de secuencia XLR70" y se someten a secuenciación de ADN 454 de alto rendimiento y se identifican secuencias individuales con sus códigos de barras que sirven como identificadores de los cuales se obtienen cadenas pesadas y/o livianas de cada una proporcionando así una guía para hacer coincidir las cadenas pesada y liviana variables individuales derivadas de la misma célula inicial, de acuerdo con los métodos detallados en "Asignación de secuencias a pozos" y "Ensamblaje de secuencias" en los materiales y métodos. Luego se dibujan los árboles evolutivos y se seleccionan los anticuerpos para la clonación, expresión y determinación de la actividad funcional (véase Figuras 6-8).

Los genes candidatos de las cadenas pesada y liviana de células de origen particulares se clonan y expresan para seleccionar las propiedades deseadas como en el ejemplo 8. Una vez que se transfectan de forma estable o transitoria, la expresión de las cadenas pesada y liviana pareadas dará lugar a la generación de anticuerpos monoclonales que recapitulan Especificidad de la célula inicialmente ordenada. Los sobrenadantes que contienen anticuerpos secretados se seleccionan luego para determinar las propiedades deseadas que incluyen, entre otras, la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 demuestran cómo se hizo esto usando las composiciones y los métodos de este documento para obtener anticuerpos monoclonales humanos contra la hemaglutinina de plasmoblastos clasificados de células individuales de un humano vacunado contra la influenza.

Ejemplo 22: Uso de células B de memoria seleccionadas no específicas o específicas de antígeno para generar anticuerpos monoclonales humanos

Desde un sujeto con una exposición documentada o sospechada a un antígeno de interés, el FACS se realiza en sangre periférica (ya sea sangre entera o células mononucleares aisladas de sangre periférica, PBMC) para identificar la población de células B de memoria (definida como CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>). Además, las células de memoria B específicas contra antígenos de interés también pueden clasificarse tiñendo sangre periférica o PBMC con marcadores de superficie de células de memoria B y con antígeno(s) conjugado con fluoróforo (CD19<sup>+</sup>CD20<sup>+</sup>CD27<sup>+</sup>antígeno<sup>+</sup>). Esta población de células se clasifica por FACS como células individuales o como células múltiples en pozos. El proceso descrito en detalle en el Ejemplo 21 se repite en el código de barras y se obtienen secuencias de la secuenciación 454 y se asignan secuencias a los pozos y se ensamblan las secuencias. HighV-QUEST se utiliza para identificar el uso del gen VDJ y algunos miembros de cada familia clonal en un árbol evolutivo seleccionado para clonación y expresión como en el ejemplo 8. La clonación y la expresión se realizan como se detalla en el ejemplo 8. Una vez que se transfecta, la expresión de la totalidad de las cadenas livianas y pesadas emparejadas resultará en la generación de anticuerpos monoclonales que recapitularán la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos se analizarán para determinar la especificidad del antígeno frente al antígeno o antígenos objetivo de interés, así como a la funcionalidad mediante los ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales humanos a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

Ejemplo 23: Uso de células B totales no seleccionadas o específicas de antígeno para generar anticuerpos monoclonales humanos

De un sujeto con o sin exposición previa sospechada o documentada a un antígeno de interés, el FACS se realiza en sangre periférica (sangre entera o células mononucleares de sangre periférica aisladas, PBMC) para identificar la población de células B CD19<sup>+</sup>. Esta población de células se clasifica por FACS como células individuales o como células múltiples en pozos. El proceso descrito en detalle en el Ejemplo 21 se repite. El proceso descrito en detalle en el Ejemplo 21 se repite en el código de barras y se obtienen secuencias de la secuenciación 454 y se asignan secuencias a los pozos y se ensamblan las secuencias. HighV-QUEST se utiliza para identificar el uso del gen VDJ, y algunos miembros de cada familia clonal en un árbol evolutivo seleccionado para clonación y expresión como en el ejemplo 8. Una vez transfectados, la expresión de cadenas livianas y pesadas emparejadas dará lugar a la generación de monoclonales. Los anticuerpos recapitulan la especificidad de la célula inicialmente seleccionada. Los sobrenadantes que contienen anticuerpos expresados se analizarán para determinar la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales humanos a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

Ejemplo 24: Uso de células plasmáticas para generar anticuerpos monoclonales humanos

De un sujeto con o sin exposición previa sospechada o documentada a un antígeno de interés, el FACS se realiza en sangre periférica (ya sea sangre entera o células mononucleares aisladas de sangre periférica, PBMC) o células de médula ósea para identificar la población de células plasmáticas CD138<sup>+</sup>. Esta población de células se clasifica por FACS como células individuales o como células múltiples en pozos. El proceso descrito en detalle en el Ejemplo 21 se repite en el código de barras y se obtienen secuencias de la secuenciación 454 y se asignan secuencias a los pozos y se ensamblan las secuencias. HighV-QUEST se utiliza para identificar el uso del gen VDJ y algunos miembros de cada familia clonal en un árbol evolutivo seleccionado para clonación y expresión. Una vez transfectados, la expresión de cadenas pesadas y livianas emparejadas resultará en la generación de anticuerpos monoclonales que recapitulan la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos expresados se analizarán para determinar la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales humanos a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

## Ejemplo 25: Uso de células B blásticas para generar anticuerpos monoclonales humanos

De un sujeto con o sin exposición previa sospechada o documentada a un antígeno de interés, el FACS se realiza en sangre periférica (ya sea sangre entera o células mononucleares aisladas de sangre periférica, PBMC) para identificar la población de células B que emiten FSC<sup>hi</sup>. Las células blásticas son células B activadas y, por lo tanto, son células que han respondido contra el antígeno y están proliferando activamente. Estas células B están formadas por familias clonales y sus cadenas livianas y pesadas se pueden utilizar para obtener árboles evolutivos. Otros marcadores de activación de células B, como CD69<sup>hi</sup> y CD44<sup>hi</sup> también se pueden usar en conjunto. Además, el contenido de ADN, que se puede teñir usando tinciones de ADN permeables a las células como SYTO Blue (Invitrogen), para determinar las células que están activadas, proliferando y en el ciclo celular, también se puede usar en conjunto para delimitar las células B blásticas. Esta población de células se clasifica por FACS como células individuales o como células múltiples en pozos. El proceso descrito en detalle en el Ejemplo 21 se repite en el código de barras y se obtienen secuencias de la secuenciación 454 y se asignan secuencias a los pozos y se ensamblan las secuencias. HighV-QUEST se utiliza para identificar el uso del gen VDJ y algunos miembros de cada familia clonal en un árbol evolutivo seleccionado para clonación y expresión. Una vez transfectados, la expresión de cadenas pesadas y livianas emparejadas resultará en la generación de anticuerpos monoclonales que recapitulan la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos expresados se analizarán para determinar la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales humanos a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

## Ejemplo 26: Uso de células B murinas para generar anticuerpos monoclonales

Un ratón es desafiado con un antígeno de interés y puede recibir inyecciones de refuerzo varias veces antes de sacrificar el ratón para obtener células B murinas. Las células murinas B se pueden obtener de la sangre, de los esplenocitos o de la médula ósea. La citometría de flujo se realiza para obtener células B CD19<sup>+</sup> o B220<sup>+</sup>. Esta población de células B se clasifica luego por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la AARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. La RT, 1<sup>a</sup> y 2<sup>a</sup> PCR se realizan como se detalla en "PCR no Touchdown" en los materiales y métodos. Los cebadores específicos del gen de ratón se encuentran en la Tabla 11 y otros cebadores usados para la RT y la PCR se encuentran en la Tabla 1. Luego se agruparon las placas múltiples de acuerdo con el método detallado en "preparación para la secuencia de 454 XLR70" y se sometieron a un alto rendimiento 454 La secuenciación de ADN y las secuencias individuales identificadas con sus códigos de barras sirven como identificadores de las cadenas pesadas y/o livianas que se obtienen de cada pozo, proporcionando así una guía para hacer coincidir las cadenas pesadas y livianas variables individuales derivadas de la misma célula inicial, de acuerdo con los métodos detallados en "Asignación de secuencias a pozos" y "ensamblaje de secuencias" en los materiales y métodos. Luego se dibujan árboles evolutivos y se seleccionan los anticuerpos para la clonación, expresión y determinación de la actividad funcional.

Las secuencias para la clonación pueden obtenerse a través de la síntesis de genes sintéticos o amplificarse a partir de los primeros productos de PCR utilizando cebadores de clonación. El cebador de clonación de avance es específico de la muestra-ID y puede amplificar secuencias específicas de un conjunto de amplicones. La secuencia de cada cadena pesada y liviana se clona luego en un vector de expresión que contiene sitios de restricción complementarios para los introducidos por los cebadores de clonación. Los vectores también contienen la región constante de la cadena pesada o liviana, en donde se clonan las secuencias de la cadena pesada o liviana (alineando el marco de lectura) para producir el anticuerpo completo. Los vectores que contienen clones de cadenas pesadas o livianas se transfectan a continuación en un sistema de expresión de mamíferos o, alternativamente, ambos amplicones pueden clonarse en un vector de expresión dual para permitir una única transfección en células de mamíferos.

Los genes candidatos de las cadenas pesada y liviana de células particulares de origen se expresan luego utilizando la selección de las propiedades deseadas como se mencionó anteriormente. Una vez transfectados de forma estable o transitoria, la expresión de las cadenas livianas y pesadas emparejadas dará lugar a la generación de anticuerpos monoclonales que recapitulan la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos secretados se seleccionan luego para determinar las propiedades deseadas que incluyen, entre otras, la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

## Ejemplo 27: Uso de células plasmáticas murinas para generar anticuerpos monoclonales

Un ratón es desafiado con un antígeno de interés y puede recibir inyecciones de refuerzo varias veces antes de sacrificar el ratón para obtener células B murinas. Las células plasmáticas murinas se pueden obtener de la sangre, de los esplenocitos o de la médula ósea, aunque los esplenocitos y la médula ósea se utilizan típicamente. La citometría de flujo se realiza para obtener células plasmáticas CD19<sup>bajo</sup>-B220<sup>bajo</sup>-CD138<sup>+</sup>. Esta población de células



plasmáticas se clasifica por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la AARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. La RT, 1<sup>a</sup> y 2<sup>a</sup> PCR se realizan como se detalla en "PCR no Touchdown" en los materiales y métodos. Los cebadores específicos del gen de ratón se encuentran en la Tabla 11, y otros cebadores usados para la RT y la PCR se encuentran en la Tabla 1. Luego se agruparon varias placas de acuerdo con el método detallado en "preparación para la secuencia de 454 XLR70" y se someten a un alto rendimiento secuenciación 454 de ADN y secuencias individuales identificadas con sus códigos de barras que sirven como identificadores de qué cadena pesada y/o liviana se obtienen de cada pozo, proporcionando así una guía para hacer coincidir cadenas variables, pesadas y livianas individuales derivadas de la misma célula inicial, de acuerdo con los métodos detallados en "Asignación de secuencias a pozos" y "ensamblaje de secuencias" en los materiales y métodos. Luego se dibujan árboles evolutivos y se seleccionan los anticuerpos para la clonación, expresión y determinación de la actividad funcional.

Las secuencias para la clonación se pueden obtener a través de la síntesis de genes sintéticos o se pueden amplificar a partir de los primeros productos de la PCR utilizando cebadores de clonación como se describe en el ejemplo 26. Los genes de cadenas pesadas y livianas candidatos de células particulares de origen se expresan luego para seleccionar las propiedades deseadas como se indica anteriormente. Una vez transfectados de forma estable o transitoria, la expresión de las cadenas livianas y pesadas emparejadas dará lugar a la generación de anticuerpos monoclonales que recapitulan la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos secretados se seleccionan luego para determinar las propiedades deseadas que incluyen, entre otras, la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

Ejemplo 28: Uso de células B de memoria murinas no específicas o específicas de antígeno para generar anticuerpos monoclonales

Un ratón es desafiado con un antígeno de interés y puede recibir inyecciones de refuerzo varias veces antes de sacrificar el ratón para obtener células B murinas. Las células murinas de memoria B se pueden obtener típicamente a partir de esplenocitos o ganglios linfáticos. La citometría de flujo se realiza para obtener células B de memoria CD19<sup>+</sup> o B220<sup>+</sup> y CD38<sup>+</sup>IgG<sup>+</sup>. También se pueden usar otros marcadores, como CD45RO. Las células B de memoria específicas de antígeno también pueden visualizarse mediante tinción con antígeno conjugado con fluoróforo y clasificarse. Esta población de células B de memoria se clasifica luego por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la AARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. La RT, la 1.<sup>a</sup> y la 2.<sup>a</sup> PCR, seguidas de la secuenciación, la asignación de secuencias a los pozos y el ensamblaje de la secuencia se realizan como en el ejemplo 26. Luego se dibujan los árboles evolutivos y se seleccionan los anticuerpos para la clonación, la expresión y la determinación de la actividad funcional.

Las secuencias para la clonación se pueden obtener a través de la síntesis de genes sintéticos o se pueden amplificar a partir de los primeros productos de la PCR utilizando cebadores de clonación como se describe en el ejemplo 26. Los genes de cadenas pesadas y livianas candidatos de células particulares de origen se expresan luego para seleccionar las propiedades deseadas como se indica anteriormente. Una vez transfectados de forma estable o transitoria, la expresión de las cadenas livianas y pesadas emparejadas dará lugar a la generación de anticuerpos monoclonales que recapitulan la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos secretados se seleccionan luego para determinar las propiedades deseadas que incluyen, entre otras, la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 3 y 9 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

Ejemplo 29: Uso de plasmoblastos murinos de vida corta para generar anticuerpos monoclonales

Un ratón es desafiado con un antígeno de interés y puede recibir inyecciones de refuerzo varias veces antes de sacrificar el ratón para obtener células B murinas. Los plasmoblastos de vida corta murinos se pueden obtener típicamente a partir de esplenocitos. Estos plasmoblastos se han descrito de diversas maneras como CD19<sup>bajo/-</sup> B220<sup>bajo/-</sup> y CD22<sup>bajo</sup> o CD11c<sup>+</sup>, y también como CD138<sup>+</sup>. La citometría de flujo se realiza para obtener plasmoblastos. Esta población de plasmoblastos se clasifica luego por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la ARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. La RT, la 1.<sup>a</sup> y la 2.<sup>a</sup> PCR, seguidas de la secuenciación, la asignación de secuencias a los pozos y el ensamblaje de la secuencia se realizan como en el ejemplo 26. Luego se dibujan los árboles evolutivos y se seleccionan los anticuerpos para la clonación, la expresión y la determinación de la actividad funcional.

Las secuencias para la clonación se pueden obtener a través de la síntesis de genes sintéticos o amplificados a partir de los 1.º productos de la PCR utilizando cebadores de clonación como se describe en el ejemplo 26. Los genes de cadenas pesadas y livianas candidatos de células particulares de origen se expresan luego para seleccionar las propiedades deseadas como se indica anteriormente. Una vez transfectados de forma estable o transitoria, la expresión de las cadenas livianas y pesadas emparejadas dará lugar a la generación de anticuerpos monoclonales que recapitulan la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos secretados se seleccionan luego para determinar las propiedades deseadas que incluyen, entre otras, la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

#### Ejemplo 30: Uso de células B blásticas murinas para generar anticuerpos monoclonales

Un ratón es desafiado con un antígeno de interés y puede recibir inyecciones de refuerzo varias veces antes de sacrificar el ratón para obtener células B murinas. Las células B blásticas murina se pueden obtener típicamente a partir de esplenocitos. Las células blásticas son células B activadas y, por lo tanto, son células que han respondido contra el antígeno y están proliferando activamente. Estas células B están formadas por familias clonales y sus cadenas livianas y pesadas se pueden utilizar para obtener árboles evolutivos. Las células B blásticas pueden clasificarse como FSC<sup>hi</sup>, y también pueden identificarse a través de marcadores de superficie celular como CD44<sup>hi</sup> CD69<sup>hi</sup> y, a medida que proliferan las células B blásticas, también pueden identificarse como que han ingresado en el ciclo celular al tener un mayor contenido de ADN al teñirse manchas de ADN de las células permeablemente tal como SYTO Blue. La citometría de flujo se realiza para obtener células B blásticas. Esta población de plasmoblastos se clasifica luego por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la ARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. La RT, la 1.ª y la 2.ª PCR, seguidas de la secuenciación, la asignación de secuencias a los pozos y el ensamblaje de la secuencia se realizan como en el ejemplo 26. Luego se dibujan los árboles evolutivos y se seleccionan los anticuerpos para la clonación, la expresión y la determinación de la actividad funcional.

Las secuencias para la clonación se pueden obtener a través de la síntesis de genes sintéticos o se pueden amplificar a partir de los productos de la 1.ª PCR utilizando cebadores de clonación como se describe en el ejemplo 26. Los genes de cadenas pesadas y livianas candidatos de células particulares de origen se expresan luego para seleccionar las propiedades deseadas como se indica anteriormente. Una vez transfectados de forma estable o transitoria, la expresión de las cadenas livianas y pesadas emparejadas dará lugar a la generación de anticuerpos monoclonales que recapitulan la especificidad de la célula seleccionada inicialmente. Los sobrenadantes que contienen anticuerpos secretados se seleccionan luego para determinar las propiedades deseadas que incluyen, entre otras, la especificidad del antígeno frente al antígeno objetivo de interés, así como la funcionalidad mediante ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

#### Ejemplo 31: Obtención de anticuerpos monoclonales a partir de células B de IgA+ humanas no específicas o específicas de antígeno

De un sujeto con o sin exposición previa sospechada o documentada de un antígeno de interés, el FACS se realiza en sangre periférica (sangre entera o células mononucleares de sangre periférica aisladas, PBMC) o en médula ósea para aislar células IgA+ B. Estas células B pueden ser células B de memoria, células plasmáticas o plasmoblastos. Estas células B IgA+ también pueden ser específicas de antígeno, al clasificar las células B positivas para antígeno utilizando un antígeno conjugado con fluoróforo para teñir las células B de IgA+. Esta población de células se clasifica por FACS como células individuales o como células múltiples en pozos. El proceso descrito en detalle en el Ejemplo 21 se repite en el código de barras y se obtienen secuencias a partir de la secuenciación 454 y se asignan secuencias a los pozos y se ensamblan las secuencias, y los cebadores específicos de la región constante de IgA utilizados para la PCR están en la Tabla 10. Se utiliza HighV-QUEST para identificar el gen VDJ uso, y unos pocos miembros de cada familia clonal en un árbol evolutivo seleccionado para clonación y expresión como en el ejemplo 8. Los sobrenadantes que contienen anticuerpos expresados se someterán a detección de especificidad de antígeno contra el antígeno objetivo de interés, así como a la funcionalidad mediante los ensayos funcionales apropiados. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales humanos a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

#### Ejemplo 32: Obtención de anticuerpos monoclonales a partir de células B IgM + humanas no específicas o específicas de antígeno

De un sujeto con o sin exposición previa sospechada o documentada a un antígeno de interés, el FACS se realiza en sangre periférica (sangre entera o células mononucleares de sangre periférica aisladas, PBMC) para aislar células IgM+ B. Estas células B pueden ser células B de memoria, células plasmáticas o células B blásticas. Estas células

IgM+ B también pueden ser específicas de antígeno, al clasificar las células B positivas para el antígeno utilizando un antígeno conjugado con fluoróforo para teñir las células B de IgM+. Esta población de células se clasifica por FACS como células individuales o como células múltiples en pozos. El proceso descrito en detalle en el ejemplo 21 se repite en el código de barras y se obtienen secuencias a partir de la secuenciación 454 y se asignan secuencias a los pozos y se ensamblan las secuencias, y los cebadores específicos de la región constante IgM utilizados para la PCR se encuentran en la Tabla 10. Se utiliza HighV-QUEST para identificar el gen VDJ uso, y unos pocos miembros de cada familia clonal en un árbol evolutivo seleccionado para clonación y expresión como en el ejemplo 8. Los sobrenadantes que contienen anticuerpos expresados se someterán a detección de especificidad de antígeno contra el antígeno objetivo de interés, así como a la funcionalidad mediante los ensayos funcionales apropiados. Las figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método. Las Figuras 26-27 proporcionan otro ejemplo de obtención de anticuerpos monoclonales humanos a partir de la caracterización funcional de anticuerpos clonados y expresados seleccionados de árboles evolutivos.

Ejemplo 33: Obtención de anticuerpos monoclonales a partir de células B IgA+ murinas no sesgadas o específicas de antígeno

Un ratón es desafiado con un antígeno de interés, y puede recibir inyecciones de refuerzo varias veces antes de sacrificar al ratón para obtener células IgA + B murinas. Estas células B pueden ser células B de memoria, células plasmáticas, plasmoblastos o células B blásticas, y típicamente pueden obtenerse a partir de esplenocitos. Estas células B IgA+ también pueden ser específicas de antígeno, al clasificar las células B positivas para antígeno utilizando un antígeno conjugado con fluoróforo para teñir las células B de IgA+. Esta población de células IgA+ B luego se clasifica por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la ARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. RT, 1ª y 2ª PCR, seguidas de la secuenciación, la asignación de secuencias a los pozos y el ensamblaje de la secuencia se realizan como en el ejemplo 26, y los cebadores específicos de la región constante de IgA utilizados para la PCR están en la Tabla 11. Luego se dibujan los árboles evolutivos y se seleccionan los anticuerpos para la clonación, expresión y determinación de la actividad funcional. Las Figuras 9 y 6 proporciona un ejemplo de una metodología general para llevar a cabo este método.

Ejemplo 34: Obtención de anticuerpos monoclonales a partir de células B IgM+ murinas no sesgadas o específicas de antígeno

Un ratón es desafiado con un antígeno de interés y puede recibir inyecciones de refuerzo varias veces antes de sacrificar el ratón para obtener células IgM+ B murinas. Estas células B pueden ser células B de memoria, células plasmáticas, plasmoblastos o células B blásticas, y típicamente pueden obtenerse a partir de esplenocitos. Estas células IgM+ B también pueden ser específicas de antígeno, al clasificar las células B positivas para el antígeno utilizando un antígeno conjugado con fluoróforo para teñir las células B de IgM+. Esta población de células IgM+ B se clasifica luego por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la ARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. RT, 1ª y 2ª PCR, seguidas de la secuenciación, la asignación de secuencias a los pozos y el ensamblaje de la secuencia se realizan como en el ejemplo 26, y los cebadores específicos de la región constante de IgA utilizados para la PCR están en la Tabla 11. Luego se dibujan los árboles evolutivos y se seleccionan los anticuerpos para la clonación, expresión y determinación de la actividad funcional. Las Figuras 9 y 6 proporciona un ejemplo de una metodología general para llevar a cabo este método.

Ejemplo 35: Secuenciación de más de una secuencia a partir de células T humanas

Desde un sujeto con una afección reciente o actual que resulta en una generación aguda, subaguda o en curso de células T circulantes, la citometría de flujo se realiza en sangre periférica (ya sea sangre entera o células mononucleares de sangre periférica (PBMC)) para identificar la población de células T de interés. Esta población de células T puede ser de células T activadas o células T blásticas. Las células T activadas pueden identificarse utilizando CD44<sup>hi</sup>, CD69<sup>hi</sup>, CD154<sup>+</sup>, CD137<sup>+</sup>, o células T blásticas, que también son células T activadas pueden delimitarse por su tamaño o FSCh<sub>i</sub>, y también pueden identificarse como que están en el ciclo celular utilizando un permeante celular Colorante de ADN como SYTO Blue. Las células T activadas deben consistir en familias clonales que luego se pueden agrupar en un árbol evolutivo, con miembros de la familia idénticos en familias clonales, que se pueden usar para seleccionar clones para el análisis funcional posterior. Las células T luego se clasifican por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la AARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. RT y PCR para codificar con barras los genes TCR se detallan en "PCR no Touchdown" en los materiales y métodos, la preparación de la secuenciación se detalla en "Preparación para la secuenciación 454 de secuencias XLR70 " en los materiales y métodos, y la asignación de secuencias a los pozos y el ensamblaje Las lecturas se detallan en "Asignación de secuencias a los pozos" y "Ensamblaje de secuencias" en los materiales y métodos. Los cebadores específicos del gen TCR se encuentran en la Tabla 10. Luego, se construyen árboles evolutivos y los genes candidatos de células particulares de origen. elegido para ser clonado y expresado para la selección de propiedades deseadas. Las secuencias para la clonación pueden ser sintetizadas por genes o amplificadas a partir de los productos de 1.ª PCR con cebadores de clonación. Los clones específicos se pueden aislar de un grupo de clones teniendo el cebador

de clonación directa específico para la secuencia de código de barras de ID de muestra. Los cebadores de clonación reversa son complementarios para el gen apropiado. Tanto los cebadores directos como los reversos contienen sitios de restricción flanqueantes para integrar el clon (marco de codificación alineado) en un vector. Las células se transfectan doblemente con dos vectores de expresión, cada uno de los cuales contiene un gen de interés, o se transfectan individualmente con un vector de expresión dual que expresa ambos genes de interés, por ejemplo, las cadenas alfa y beta de células T.

Una vez transfectados de forma estable o transitoria, los genes de interés se pueden expresar y analizar las propiedades de la función utilizando las cribas deseadas.

Ejemplo 36: Secuenciación de más de una secuencia a partir de células T murinas

10 Un ratón es desafiado con un antígeno de interés y puede recibir inyecciones de refuerzo varias veces antes de sacrificar al ratón para obtener células T murinas. Las células T son CD3<sup>+</sup>, y las células T auxiliares son CD4<sup>+</sup> y las células T citotóxicas son CD8<sup>+</sup>. Esta población de células T puede ser memoria o células T activadas o células T blásticas. Las células T de memoria pueden identificarse como CD45RO<sup>+</sup>. Las células T activadas pueden identificarse utilizando CD44<sup>hi</sup>, CD69<sup>hi</sup>, o células T blásticas, que también son células T activadas pueden delimitarse por su tamaño o FSC<sup>hi</sup>, y también pueden identificarse por estar en el ciclo celular utilizando un colorante de ADN permeable a las células, como SYTO azul. Todas estas células T en un ratón mantenidas en un ambiente limpio después de la exposición repetida al antígeno deben tener una gran fracción de familias clonales que luego se pueden mostrar como un árbol evolutivo, que luego se puede usar para seleccionar TCR para la clonación y la expresión y el análisis funcional en sentido descendente.

20 Las células T se clasifican por citometría de flujo utilizando los marcadores sugeridos anteriormente. como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la ARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. RT y -PCR para codificar con barras los genes TCR se detallan en "PCR no Touchdown" en los materiales y métodos, la preparación de la secuenciación se detalla en "Preparación para la secuenciación 454 de secuencias XLR70 " en los materiales y métodos, y la asignación de secuencias a los pozos y el ensamblaje Las lecturas se detallan en "Asignación de secuencias a los pozos" y "Ensamblaje de secuencias" en los materiales y métodos. Los cebadores específicos del gen TCR se encuentran en la Tabla 11. Luego, se construyen árboles evolutivos y los genes candidatos de células particulares de origen. elegido para ser clonado y expresado para la selección de propiedades deseadas. Las secuencias para la clonación pueden ser sintetizadas por genes o amplificadas a partir de los productos de 1.<sup>a</sup> PCR con cebadores de clonación. Los clones específicos se pueden aislar de un grupo de clones teniendo el cebador de clonación reversa son complementarios para el gen apropiado. Tanto los cebadores directos como los reversos contienen sitios de restricción flanqueantes para integrar el clon (marco de codificación alineado) en un vector. Las células se transfectan doblemente con dos vectores de expresión, cada uno de los cuales contiene un gen de interés, o se transfectan individualmente con un vector de expresión dual que expresa ambos genes de interés, por ejemplo, las cadenas alfa y beta de células T.

Una vez transfectados de forma estable o transitoria, los genes de interés se pueden expresar y analizar las propiedades de la función utilizando las cribas deseadas.

Ejemplo 37: Secuenciación de más de una secuencia a partir de una muestra

40 Se identifica una muestra única que comprende ácidos nucleicos de interés. La muestra única puede tener una sola célula o una población de células. La muestra se puede clasificar por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la AARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. Esta población de células B se clasifica luego por citometría de flujo como células individuales en pozos que contienen un regulador hipotónico con un inhibidor de la AARNasa. Las células clasificadas pueden congelarse en este momento o usarse inmediatamente para RT-PCR para crear ADNc. La RT, 1<sup>a</sup> y 2<sup>a</sup> PCR se realizan como se detalla en "PCR no Touchdown" en los materiales y métodos. Las placas múltiples se agrupan luego de acuerdo con el método detallado en "preparación para la secuenciación 454 de secuencia XLR70" y se someten a una secuenciación de ADN 454 de alto rendimiento y se identifican secuencias individuales con sus códigos de barras que sirven como identificadores de los cuales se obtienen cadenas pesadas y/o livianas de cada una. proporcionando así una guía para hacer coincidir las cadenas pesada y liviana variables individuales derivadas de la misma célula inicial, de acuerdo con los métodos detallados en "Asignación de secuencias a pozos" y "ensamblaje de secuencias" en los materiales y métodos. Luego se dibujan árboles evolutivos y se seleccionan los anticuerpos para la clonación, expresión y determinación de la actividad funcional.

55 Los genes candidatos de células de origen particulares se eligen para clonarse y expresarse para seleccionar las propiedades deseadas u otras necesidades. Las secuencias para la clonación pueden ser sintetizadas por genes o amplificadas a partir de los productos de 1.<sup>a</sup> PCR con cebadores de clonación. Los clones específicos se pueden aislar de un grupo de clones teniendo el cebador de clonación directa específico para la secuencia de código de barras de

ID de muestra. Los cebadores de clonación reversa son complementarios para el gen apropiado. Tanto los cebadores directos como los reversos contienen sitios de restricción flanqueantes para integrar el clon (marco de codificación alineado) en un vector. Las células se transfectan con dos o más vectores de expresión, cada uno de los cuales contiene un gen de interés, o se transfectan individualmente con un vector de expresión que expresa los genes de interés.

Una vez transfectados de forma estable o transitoria, los genes de interés pueden expresarse y explorarse donde se desee. Las Figuras 9 y 6 proporcionan un ejemplo de una metodología general para llevar a cabo este método.

Ejemplo 38: Clonación de regiones de inmunoglobulina V(D)J por síntesis de ADN.

Las regiones V(D)J de cadena liviana y cadena pesada de inmunoglobulina deseadas pueden generarse sintéticamente mediante síntesis de ADN para clonación en vectores de expresión. La secuencia utilizada para la síntesis puede derivarse directamente de las secuencias 454 de alto rendimiento, o alternativamente, el ADNc que codifica las inmunoglobulinas de las cadenas pesada y liviana de la(s) muestra(s) de interés se puede secuenciar de la muestra individual o de las muestras agrupadas para obtener más información de verificación de la secuencia, y esta secuencia se usa para sintetizar las regiones V(D)J de cadena liviana y cadena seleccionada seleccionadas. Las regiones variables de los genes de Ig se pueden clonar por síntesis de ADN, e incorporar el ADN sintetizado en el vector que contiene la región constante apropiada utilizando enzimas de restricción y biología molecular estándar. Durante la síntesis, no es necesario seguir la secuencia exacta de nucleótidos siempre que la secuencia de aminoácidos no cambie, a menos que se desee la mutagénesis. Esto permite la optimización del codón que puede resultar en niveles de expresión más altos. Esto también permite agregar sitios de restricción con el propósito de clonar. Las secuencias no traducidas, tales como 5' UTR y secuencias de códigos de barras, no necesitan sintetizarse, las secuencias líderes también pueden intercambiarse por otras secuencias peptídicas de señal conocidas por niveles de expresión más altos. Esto da como resultado una secuencia de nucleótidos de Ig que puede ser muy diferente de las lecturas de alto rendimiento pero que dan la secuencia de aminoácidos idéntica cuando se expresa.

En una realización, las regiones V(D)J amplificadas se insertan en vectores que ya contienen las regiones constantes de isotipo kappa, lambda, gamma u otras cadenas pesadas con los sitios de restricción apropiados necesarios para insertar los clones amplificadas en el marco de lectura abierto. En otra realización, la región variable completa puede sintetizarse in gen con región constante y clonarse en un vector de expresión para la expresión y la prueba funcional posterior de las propiedades del anticuerpo.

Ejemplo 39: Clonación de las regiones de inmunoglobulina V(D)J utilizando un sitio de restricción ya presente en el adaptador de identificación de la muestra.

En otro aspecto, las regiones V(D)J de cadena liviana y cadena pesada de inmunoglobulina deseadas se pueden clonar usando un sitio de restricción ya incorporado en el adaptador de ID de muestra agregado durante la transcripción reversa. Esto da como resultado un adaptador con un sitio de restricción 3' del código de barras de ID de pozo en el conjunto de amplicones de PCR. Durante la clonación con cebadores de clonación, los amplicones deseados se amplifican a partir de un grupo de amplicones específicos de placa utilizando cebadores en 5' que son complementarios a las secuencias de código de barras de ID de pozo, y cebadores 3' específicos de cadena (para las cadenas kappa, lambda y gamma). Los cebadores en 3' agregarán en 3' sitios de restricción. Los cebadores en 5' no necesitan agregar sitios de restricción, ya que el cebador en 5' ya contiene un sitio de restricción 3' del código de barras de ID de pozo. Después de esta amplificación, se utilizan enzimas de restricción para cortar el amplicón para la ligación en el vector que contiene el inserto de la región constante. Durante la digestión con enzimas de restricción, las secuencias agregadas al extremo 5' de las secuencias del gen Ig, como los códigos de barras y las secuencias universales, se cortan, ya que están en 5' del sitio de restricción 5'.

Ejemplo 40: Identificación de familias clonales mediante la secuenciación de solo una cadena de inmunoglobulina (cadena pesada o cadena liviana), seguida de la clonación de las regiones V(D)J de cadena liviana y pesada de inmunoglobulina pareadas

Las cadenas pesadas y livianas de anticuerpos se transcriben de manera reversa a partir de ARNm, incorporando distintas ID de muestra en los ADNc generados a partir de cada muestra, y los ADNc de muestra agrupados para amplificar la PCR. Los ADNc de inmunoglobulina se amplifican y la cadena pesada de inmunoglobulina o la cadena liviana se secuencian utilizando secuenciación 454 de alto rendimiento, y las secuencias se agrupan de acuerdo con su uso de la cadena pesada V(D)J de inmunoglobulina o secuencias V(D)J de cadena liviana que muestran el uso de los mismos segmentos V(D)J codificados por genoma. Se utiliza la bioinformática para identificar las familias clonales de interés, y las regiones V(D)J de cadena liviana y pesada de inmunoglobulina deseadas de la misma muestra se amplifican selectivamente para la secuenciación y/o clonación. Para la amplificación por PCR, el cebador directo incluye la identificación de la muestra y el cebador reverso es específico para la cadena constante o la región constante de la cadena pesada. Los cebadores pueden incorporar sitios de restricción en los amplicones. Los amplicones se pueden insertar en los vectores de expresión apropiados que ya contienen una región constante de cadena pesada o liviana. Los anticuerpos se pueden expresar y analizar para las propiedades deseadas.

Ejemplo 41: Identificación de familias clonales de la secuenciación V(D)J de la cadena pesada y liviana de inmunoglobulina para la clonación y expresión de anticuerpos utilizando solo ID de muestra (y no ID de placa)

Las cadenas pesadas y livianas de anticuerpos se transcriben a la reversa del ARNm en cada muestra, incorporando distintas ID de muestra en el ADNc generado a partir de cada muestra. Cada ID de muestra tiene una longitud de al menos 6 nucleótidos y 1 par de bases diferente, lo que da como resultado 4096 ID de muestra potenciales distintas. Se utiliza un ID de muestra distinto para cada muestra, y los ID de muestra únicos identifican el ADNc derivado de diferentes muestras y permiten la secuenciación y clonación pareadas de 2 o más ADNc expresados en una muestra individual. Los amplicones de cadena pesada y liviana se amplifican mediante PCR, que se agrega a los adaptadores A y B de titanio requeridos para secuenciación 454 de alto rendimiento y todas las muestras se envían para su secuenciación. Las secuencias se asignan a los pozos y se ensamblan siguiendo las secciones "Asignación de secuencias a los pozos" y "Ensamblaje de secuencias" en los materiales y métodos. Las asignaciones de V(D)J se realizan utilizando HighV-QUEST y se agrupan en familias clonales en función de su uso de V(D)J. Los clones seleccionados luego se amplifican específicamente con cebadores de clonación, que también agregan sitios de restricción en el amplicón. Los amplicones se insertan luego en el marco en vectores de expresión que ya contienen las regiones constantes pesadas o livianas apropiadas para la expresión de los anticuerpos para la detección de las propiedades deseadas.

Ejemplo 42: Clonación de secuencias pareadas por ligación en la secuencia del cebador universal.

Los genes de cadenas pesadas y livianas de anticuerpos se transcriben a la reversa del ARNm, lo que agrega una secuencia 3' al ADNc recién sintetizado que consiste en una región adaptadora y un código de barras de ID de muestra. Las muestras se agrupan y se agrega una secuencia de cebador universal en el extremo 3' del ADNc de la primera hebra utilizando la ADN ligasa de T4 y un oligonucleótido de cebador universal antisentido 5' fosforilado. Alternativamente, la síntesis de ADNc de 2.<sup>a</sup> hebra se puede hacer para obtener un ADNc de doble cadena en lugar de un híbrido de ARNm/ADNc antes de ligarse en la secuencia del cebador universal. Luego se realizan dos rondas de PCR para amplificar el ADNc y agregar las ID de placa y los cebadores de titanio A y B para la secuenciación de 454. Alternativamente, las ID de placa y los cebadores de titanio también se pueden agregar mediante ligación de ADN en lugar de incorporarse durante la PCR utilizando la ADN ligasa de T4. Después de secuenciación 454, las secuencias se ensamblan y se identifican las familias clonales. Los clones seleccionados de familias clonales pueden clonarse específicamente utilizando cebadores de clonación que agregan sitios de restricción a los amplicones. Luego, las secuencias se insertan en el marco en vectores de expresión que ya tienen las regiones constantes de cadena pesada o liviana apropiadas. Luego se expresan los anticuerpos y se analizan las propiedades deseadas.

Ejemplo 43: Prueba de cebadores específicos de genes para la transcripción reversa (RT-GSP) de cadenas pesadas y livianas de inmunoglobulina.

Se usaron RT-GSP en lugar de oligo (dT) como cebadores en la transcripción reversa de los genes de las cadenas pesada y liviana. El ADNc se amplificó luego por PCR y se visualizó en un gel de agarosa. Los cebadores RT-GSP fueron IgKC\_v3 (a), IgLC\_v5, IgLC\_v6, IgLC\_v7 e IgLC\_v8 en los carriles 1-4, respectivamente (b), IgHGC\_v10, IgHGC\_v11, IgHGC\_v13 e IgGC\_v15 en los carriles 1-4 respectivamente (c) e IgHGCvc, y las cámaras de seguridad. KC, LC y GC en los nombres de los cebadores indican que el cebador es específico para la cadena kappa, la cadena lambda y la cadena pesada gamma, respectivamente. Las bandas blancas en las fotos de gel indican dónde se habían recortado los carriles no relevantes. Véase Figura 10 y Tabla 6.

Ejemplo 44: Prueba de secuencias de región adaptadora.

El ARN se sometió a transcripción reversa con oligonucleótidos que comprenden una región cebadora universal y una región adaptadora en el extremo 3' terminal. El ADNc se amplificó luego utilizando la secuencia de la región del cebador universal como cebador directo y las secuencias específicas del gen como cebadores reversos. Los productos amplificados se visualizaron en un gel de agarosa. La región del adaptador consta de G (a), GGGGG y rGrGrG en los carriles 1 y 2 respectivamente (b). rG indica nucleótidos de ARN en lugar de nucleótidos de ADN. Véase Figura 11 y Tabla 6.

Ejemplo 45: Ensayo de secuencias de cebadores universales.

El ARN se sometió a transcripción reversa con oligonucleótidos que comprendían una secuencia de cebador universal y una región adaptadora en el extremo 3' terminal. El ADNc se amplificó luego por PCR utilizando un cebador directo complementario a la región del cebador universal y un cebador reverso complementario a la secuencia específica del gen. Univ\_seq\_4 (a), univ\_seq\_5 (b) y univ\_seq\_f (c). Las bandas blancas verticales en las fotos de gel indican dónde se han recortado los carriles no relevantes. De lo contrario, los carriles pertenecen a la misma fotografía de gel. Véase Figura 12 y Tabla 6.

Ejemplo 46: Ensayo de secuencias de cebadores específicas de genes para la 1.<sup>a</sup> reacción de PCR.

Se usaron cebadores reversos específicos de gen en la amplificación de secuencias en la primera reacción de PCR. Tanto la primera reacción de PCR como los posteriores productos de PCR anidados se ejecutaron y se visualizaron en un gel de agarosa. Los cebadores reversos utilizados fueron IgKC\_v4, IgLC\_v5, IgHGC\_v13 en los carriles 1-3 respectivamente (a), K\_GSP1, L\_GSP1, G\_GSP1 en los carriles 1-3 respectivamente (b), K\_GSP1c, L\_GSP1c en los carriles 1-2 respectivamente (c), G\_GSP1c, G ), L\_GSP1d, G\_GSP1g en los carriles 1-2 respectivamente (e), G\_GSP1h, G\_GSP1k, L\_GSP1f, L\_GSP1g en los carriles 1-4 respectivamente (h), respectivamente, (f), G\_GSP1d (g) L\_GSP1h-o en los carriles 1-8, respectivamente) -q y G\_GSP1t en los carriles 1-6 respectivamente (K, L y G en los nombres de los cebadores indican que los cebadores son específicos para las regiones constantes de inmunoglobulina kappa, lambda y gamma, respectivamente). Cada gel comienza con un marcador de carril a la izquierda seguido de los carriles de muestra. Las barras blancas entre carriles en la misma fotografía de gel indican dónde se han recortado los carriles intermedios no relevantes. Véase figura. 13. Además, se probaron más cebadores en la Figura 43. Estos cebadores se utilizaron para la primera PCR, y luego se realizó la segunda PCR utilizando los cebadores de la Tabla 1 y los productos de la PCR se realizaron en un gel de agarosa al 2% y se tomó la imagen. Los cebadores utilizados para la 1ª PCR son Kappa GSP1, kappa GSP1e, kappa GSP1f, lambda GSP1, lambda GSPlx y lambda GSPly respectivamente. Consúltense también la Tabla 6 para las secuencias utilizadas.

Ejemplo 47: Prueba de secuencias específicas de genes para la 2ª reacción de PCR.

Se usaron cebadores reversos específicos de gen en la amplificación de secuencias en la 2.ª reacción de PCR. Los productos de PCR se corrieron y se visualizaron en un gel de agarosa. Los cebadores reversos utilizados fueron K\_GSP2, L\_GSP2, G\_GSP2 en los carriles 1-3 respectivamente (a), K\_GSP2v2a, K\_GSP2v2b, L\_GSP2v2 en los carriles 1-3 respectivamente (b), K\_GSP2v2c, K\_GSP2v2c, G\_GSP2v2c1, G\_GSP2v2c2 en los carriles 1-4 respectivamente (c), K\_GSP2v2d-f en los carriles 1-3 respectivamente (d), K\_GSP2v2g, L\_GSP2v2d y G\_GSP2b en los carriles 1-3 respectivamente (e). Regiones constantes de inmunoglobulina lambda y gamma respectivamente. Cada gel comienza con un marcador de carril a la izquierda seguido de los carriles de muestra. Las barras blancas entre carriles en la misma fotografía de gel indican que los carriles intermedios no relevantes se han recortado. Véase Figura 14 y Tabla 6.

Ejemplo 48: Prueba de cebadores específicos de genes para otros genes de regiones variables humanas

La 1ª y 2ª PCR se realizaron utilizando cebadores reversos específicos del gen y los productos se corrieron en un gel de agarosa al 2% y se generaron imágenes. Los carriles son desde la izquierda: marcador, mu, regiones constantes alfa, TCR alfa (a) y marcador, TCR beta (b). las secuencias de los cebadores 3' utilizados se encuentran en la tabla 10. Las barras blancas entre carriles en la misma fotografía de gel indican dónde se han recortado los carriles intermedios no relevantes. Véase Figura 44.

Ejemplo 49: Prueba de cebadores específicos de genes para genes de región variable de ratón

Se realizaron 1.ª y 2.ª PCR y los productos se corrieron en un gel de agarosa al 2% y se generaron imágenes. Los carriles son de izquierda a derecha: marcador, kappa, lambda, lambda, lambda, cadenas livianas lambda y cadena pesada mu (a). Los 4 carriles lambda tenían esta combinación de cebadores utilizados: mouse\_lambda\_GSP1a con mouse\_lambda\_GSP2a, mouse\_lambda\_GSP1a con mouse\_lambda\_GSP2b, mouse\_lambda\_GSP1b con mouse\_lambda\_GSP2a y mouse\_lambda\_GSP1b con mouse\_lambda\_GSP2a. Marcador y cadena pesada alfa (b). Cadenas pesadas de Gamma1, 2a, 2c con 2.ª PCR usando mo\_g12\_GSP2d y mo\_g12\_GSP2e respectivamente, marcador (c). Cadena pesada Gamma 3 con 2.ª PCR usando mo\_g3\_GSP2d, mo\_g3\_GSP2e respectivamente seguido por la cadena pesada gamma 2b con la 2.ª PCR usando mo\_g2b\_GSP2d, mo\_g2b\_GSP2e respectivamente, seguido del marcador (d). Marcador, TCR alfa (e). Marcador, TCR beta (f). Las barras blancas entre carriles en la misma fotografía de gel indican dónde se han recortado los carriles intermedios no -relevantes. Véase la Figura 45 y la Tabla 11.

Ejemplo 50: Generación de pares enlazados de cadenas pesadas y livianas de anticuerpos con un código de barras en un extremo.

Como se muestra en la Figura 1, las células B individuales se pueden seleccionar por citometría de flujo de sangre, células mononucleares de sangre periférica (PBMC), células B de volumen, plasmoblastos, células plasmáticas, células B de memoria u otras poblaciones de células B. Las células B se clasifican en una sola célula en placas de PCR de 96 pozos, dejando una columna de pozos vacía, como control negativo. La Figura 17 describe la metodología general para un método que puede utilizarse para vincular dos secuencias de polinucleótidos de interés y agregar un código de barras en un extremo.

La RT-PCR multiplexada de solapamiento y extensión de un solo paso se puede realizar utilizando un kit de RT-PCR de un solo paso disponible en el mercado (por ejemplo, el kit Qiagen RT PCR de un solo paso) de acuerdo con las recomendaciones del fabricante. En este ejemplo particular, una reacción de síntesis de polinucleótidos, tal como una reacción de transcripción reversa, se usa para generar plantillas de ADNc a partir de una muestra de ARNm. Con referencia a la Figura 17, el cebador específico del gen directo para la reacción de RT-PCR contiene un sitio de enzima de restricción (RE1), un sitio de cebador de secuenciación y un código de barras, para agregar estos elementos a un

primer ADNc de interés. Dos cebadores adicionales (que se muestran por contener RE3) tienen colas complementarias de superposición y extensión. El uso de estos cebadores en una reacción de PCR da como resultado dos ADNc de interés que llevan colas de extensión solapadas, lo que permite que los dos ADNc de interés se alisen y se unan durante la amplificación. En el ejemplo que se muestra, se generaría un producto de la estructura indicada en donde las cadenas LC y HC están unidas físicamente con un código de barras en un extremo.

Los sitios de restricción RE1 y RE2 se pueden usar para clonar el producto de PCR en vectores adecuados para la secuenciación.

Ejemplo 51: Generación de pares enlazados de cadenas pesadas y livianas de anticuerpos con un código de barras interno.

Como se muestra en la Figura 1, las células B individuales se pueden clasificar por citometría de flujo de sangre, células mononucleares de sangre periférica (PBMC), células B de volumen, plasmoblastos, células plasmáticas, células B de memoria u otras poblaciones de células B. Las células B se clasifican en una sola célula en placas de PCR de 96 pozos, dejando una columna de pozos vacía, como control negativo. La Figura 18 describe la metodología general para un método que puede usarse para vincular dos polinucleótidos de interés con un código de barras ubicado en el medio. Los cebadores y oligonucleótidos que pueden usarse para las cadenas pesadas y livianas de anticuerpos se muestran en la Tabla 30. Los sitios de restricción AsiSI y PacI se incluyen en los oligonucleótidos de RT. Las secuencias de -ID de muestra se muestran en la Tabla 2.

El método que se muestra en la Figura 18 se basa en las actividades de cambio de plantilla y de cola de 3' de la transcriptasa reversa durante una reacción de síntesis de ADNc. La cola 3' C agregada al ADNc sintetizado puede usarse para el fusionamiento de una molécula adaptadora que lleva una secuencia de extensión de superposición y un código de barras. Se utilizan dos tipos de moléculas adaptadoras para unir dos ADNc. Un primer adaptador que lleva una extensión de superposición y una secuencia de código de barras se agrega a un primer ADNc. Un segundo adaptador que lleva el complemento reverso de la extensión de superposición sin una secuencia de código de barras se agrega a un segundo ADNc. La propiedad de cambio de plantilla de la transcriptasa reversa agrega estas secuencias a los extremos 3' de sus respectivos ADNc.

En una reacción de PCR, como se muestra en la Figura 18, las secuencias de extensión de superposición complementarias de fusionamiento y las cadenas correspondientes de ADN se sintetizan a partir del sitio de fusionamiento. Rondas subsiguientes de PCR que usan cebadores externos dan como resultado la amplificación de las moléculas de ADNc enlazadas.

Mediante la adición de sitios de restricción apropiados y la adición de sitios de cebadores de secuenciación incorporados en los cebadores para la reacción de amplificación o más tarde mediante ligación, los productos de la PCR se pueden clonar en vectores adecuados para la secuenciación.

Ejemplo 52: Generación de pares enlazados de cadenas pesadas y livianas de anticuerpos con dos códigos de barras internos utilizando cebadores de solapamiento-extensión de secuencia universales.

Como se muestra en la Figura 1, las células B individuales se pueden clasificar por citometría de flujo de sangre, células mononucleares de sangre periférica (PBMC), células B de volumen, plasmoblastos, células plasmáticas, células B de memoria u otras poblaciones de células B. Las células B se clasifican en una sola célula en placas de PCR de 96 pozos, dejando una columna de pozos vacía, como control negativo. La Figura 19 describe la metodología general para un método que puede usarse para introducir dos códigos de barras internos entre dos polinucleótidos de interés unidos. Los cebadores y oligonucleótidos que se pueden usar para las cadenas pesadas y livianas de anticuerpos se muestran en la Tabla 31. Los sitios de restricción AsiSI y PacI se incluyen en el oligonucleótido RT. Las secuencias de ID de muestra se muestran en la Tabla 2.

El método que se muestra en la Figura 19 se basa en las actividades de cambio de plantilla y de cola de 3' de la transcriptasa reversa durante una reacción de síntesis de ADNc. En este ejemplo, la cola 3' C agregada al ADNc cebado con oligo (dT) se puede usar para el fusionamiento de una molécula adaptadora que lleva una secuencia universal y un código de barras a cada uno de los ADNc que se unirán. La propiedad de cambio de plantilla de la transcriptasa reversa agrega estas secuencias a los extremos 3' de sus respectivos ADNc. La PCR posterior de superposición y extensión que utiliza cebadores para la secuencia universal que llevan secuencias complementarias de superposición y combinación con los cebadores específicos de LC y HC externos da como resultado una estructura en donde la LC está unida a HC con dos códigos de barras internos entre ellos, como se muestra en la Figura 19.

Mediante la adición de sitios de restricción apropiados y la adición de sitios de cebadores de secuenciación incorporados en los cebadores para la reacción de amplificación o más tarde mediante ligación, los productos de la PCR se pueden clonar en vectores adecuados para la secuenciación.

Ejemplo 53: Generación de pares enlazados de cadenas pesadas y livianas de anticuerpos con dos códigos de barras internos que utilizan adaptadores de superposición y extensión.



Como se muestra en la Figura 1, las células B individuales se pueden clasificar por citometría de flujo de sangre, células mononucleares de sangre periférica (PBMC), células B de volumen, plasmoblastos, células plasmáticas, células B de memoria u otras poblaciones de células B. Las células B se clasifican en una sola célula en placas de PCR de 96 pozos, dejando una columna de pozos vacía, como control negativo. La Figura 20 describe la metodología general para otro método que puede usarse para introducir dos códigos de barras internos entre dos polinucleótidos de interés unidos. Los cebadores y oligonucleótidos que pueden usarse para las cadenas pesadas y livianas de anticuerpos se muestran en la Tabla 32. Los sitios de restricción AsiSI y PacI se incluyen en los oligonucleótidos de RT. Las secuencias de ID de muestra se muestran en la Tabla 2.

El método que se muestra en la Figura 20 también se basa en las actividades de cambio de plantilla y de cola de 3' de la transcriptasa reversa durante una reacción de síntesis de ADNc. En este ejemplo, la cola 3' C agregada al ADNc sintetizado utilizando cebadores específicos de gen puede usarse para el fusionamiento de moléculas adaptadoras que llevan secuencias de superposición y superposición palindrómicas autocomplementarias y un código de barras para cada uno de los ADNc a unir. La propiedad de cambio de plantilla de la transcriptasa reversa agrega estas secuencias a los extremos 3' de sus respectivos ADNc. El fusionamiento posterior de las secuencias de superposición-extensión agregadas a los ADNc de LC y HC las une en el sitio de superposición. La PCR de solapamiento y extensión que utiliza cebadores externos para LC y HC da como resultado una estructura en donde LC está unida a HC con dos códigos de barras internos entre ellos, como se muestra en la Figura 20.

Mediante la adición de sitios de restricción apropiados y la adición de sitios de cebadores de secuenciación incorporados en los cebadores para la reacción de amplificación o más tarde mediante ligación, los productos de la PCR se pueden clonar en vectores adecuados para la secuenciación

Ejemplo 54: Estudios sobre diferentes métodos para agregar códigos de barras.

Se investigó una variedad de métodos a través de los cuales se podrían agregar secuencias de códigos de barras durante el curso de una reacción de transcripción o amplificación reversa utilizando un oligonucleótido que comprende la secuencia de códigos de barras. Se probó la adición de códigos de barras incorporándolos en cebadores específicos de genes (GSP) y en oligonucleótidos que contienen uno o más Gs que pueden agregarse al extremo 3' de los ADNc mediante cambio de plantilla. Con bases en la literatura y nuestro conocimiento científico, nuestra expectativa era que pudiéramos codificar de manera efectiva el ADNc utilizando oligonucleótidos con código de barras 5' o GSP con código de barras 3'.

Como se demuestra en la Figura 21, la RT se realizó con 1 µg de ARN de PBMC total y 0.5 µM de oligo de cambio de plantilla univ\_seq\_2 y 0.1 µM de GSP IgKC\_v3 (carriles 1-2) e GSP IgLC\_v5 (carriles 3-4) con un adicional de 5' secuencia de flanqueo, de la cual la primera parte es la secuencia Fixed\_PCR3, y el último AACAAATAC de 8 pb es el código de barras. La reacción de RT se limpió utilizando NucleoTraPCR (Macherey-Nagel) y se disolvió en un volumen final de 50 µl. Se utilizaron 2 µl de esta reacción en cada reacción de PCR subsiguiente, con un cebador interno 5' VK (carril 1) o VL (carril 3) o Univ\_seq\_2 (carriles 2 y 4) como cebador en 5', y Fixed\_PCR3 como el cebador Primer 3'. Téngase en cuenta que el cebador VK es específico para los genes kappa V 1 y 2, y el cebador VL es específico para el gen lambda V 2. Las secuencias están en la Tabla 33. Como se puede ver, los productos de la PCR en las calles 2 y 4 se ejecutaron como un frotis. En contraste, los cebadores en 5' internos produjeron bandas distintas (carriles 1 y 3), lo que muestra que los pares de cebadores funcionan, y el borrón que se muestra en los carriles 1 y 3 no se puede atribuir a cebadores mal diseñados. A medida que el oligonucleótido se agrega a todas las secuencias de ADNc de transcripción reversa de longitud completa, cuando se obtiene un frotis durante la amplificación por PCR con las GSP con código de barras univ\_seq\_2 y 3', esto sugiere que la transcripción reversa con GSP con códigos de barras da como resultado secuencias de ácido nucleico cebadas no específicamente en la reacción de RT. Nuestros resultados sugieren que el uso de adaptadores de secuencia universales 5' y cebadores de código de barras 3' no es una buena estrategia para el código de barras y la amplificación específica de inmunoglobulina u otros genes expresados por una célula B u otra célula.

En retrospectiva, varias propiedades biológicas del ADN y las reacciones moleculares utilizadas probablemente contribuyan a nuestras observaciones. La transcripción reversa se realiza generalmente a bajas temperaturas, como 42°C o 56°C. A diferencia de la PCR, donde el paso de fusionamiento se realiza generalmente a una temperatura ligeramente inferior a la Tm de los cebadores para promover la especificidad de cebado, esto no se puede hacer para la transcripción reversa, ya que las transcriptasas reversas se inactivan a altas temperaturas. Por lo tanto, los cebadores específicos de genes utilizados durante la RT generalmente no son muy específicos para el gen de interés porque la reacción se produce a una temperatura mucho más baja que la Tm del cebador. En tal situación, el cebador también puede unirse a secuencias de ARNm fuera del objetivo con algunos desajustes, y se produce un error de cebador. Si se agrega el código de barras en el GSP, el cebador también debe tener una secuencia fija 5' del código de barras para su uso en la PCR posterior. Esto hace que el cebador sea muy largo (~60 nt), lo que da como resultado un cebador con un grado de cebador mucho mayor. Sin embargo, la amplificación específica durante la PCR por lo general todavía se puede lograr utilizando un cebador directo altamente específico del gen; siempre que un miembro del par de cebadores sea específico, generalmente puede haber una amplificación específica, como se muestra en los carriles 1 y 3.

Si se utiliza una técnica de cambio de plantilla para agregar un adaptador, este adaptador se agrega a todos los ARNm que se sintetizan para formar el ADNc de primera cadena. Como se mencionó anteriormente, los GSP con código de barras tendrán cantidades significativas de cebador errónea, especialmente cuando la enzima RT Superscript III pierde su actividad de cambio de plantilla a 56°C, y la transcripción reversa se realiza a 42°C. Los cebadores en 5' o 3' anidados específicos no se pueden usar, ya que se perdería la capacidad de amplificar por PCR todos los genes de inmunoglobulina (por lo tanto, habrá que recurrir a la PCR multiplex con cebadores en 5' degenerados múltiples debido a los genes V variables) o perder el 3' código de barras.

Por lo tanto, los GSP con código de barras no son adecuados para el uso con adaptadores agregados de cambio de plantilla, o cualquier otro adaptador 5', ya que otros métodos para agregar adaptadores 5' como la cola TdT o la clonación de extremos romos también agregan adaptadores sin discriminación. Por lo tanto, también se requieren cebadores internos 3' o una estrategia de amplificación por PCR anidada o semianidada, y los GSP 3' con código de barras no permiten el uso de estas estrategias para la amplificación específica de genes de una célula B u otra célula. Con base en nuestros resultados, también se podría anticipar que los oligo con código de barras (dT) también tendrían un bajo rendimiento por muchas de las mismas razones por las que creemos que los GSP con códigos de barras tienen un bajo rendimiento. Estas razones incluyen pero no se limitan a una incapacidad para usar cebadores internos 3' o una estrategia de PCR anidada o semianidada para la amplificación específica de genes de una célula B u otra célula.

En contraste, nuestros resultados (véase otros ejemplos) demuestran la superioridad de la adición de la secuencia de códigos de barras durante el curso de una reacción de transcripción reversa usando un cebador que comprende la secuencia de códigos de barras y un adaptador que se acopla a la cola 3' de un ADNc generado durante una transcripción reversa reacción. En una realización de este tipo, la secuencia adaptadora puede comprender una secuencia de código de barras y usarse para marcar genes que codifican cadenas pesadas y livianas de anticuerpos. Por lo tanto, como se describe en este documento, el cambio de plantilla, o cualquier otro método para enlazar un ADNc agrega una secuencia que se puede usar para la amplificación por PCR sin el conocimiento previo de las secuencias 5' en sí, permitiendo una representación eficiente e imparcial del repertorio de anticuerpos. Además, esta metodología permite obtener el repertorio de otros genes coexpresados que codifican proteínas además de anticuerpos. Además, el enfoque del uso de adaptadores de cambio de plantilla tiene claras ventajas sobre los métodos descritos en la técnica que utilizan conjuntos de cebadores directos degenerados para amplificar múltiples genes V. Estos métodos tampoco logran capturar todo el repertorio de anticuerpos, porque los conjuntos de cebadores en 5' conocidos: a) cubren la mayoría, pero no todo el conjunto de repertorio; b) no son capaces de cubrir las variantes de los genes V (polimorfismos) todavía conocidas en la población humana; y c) puede que no sea capaz de amplificar eficazmente las secuencias de anticuerpos que han sufrido una hipermutación somática extensa (SHM). Consúltese, por ejemplo, Scheid et al., Scienceexpress, 14 de julio de 2011 para ver un ejemplo del efecto de SHM.

Por consiguiente, el uso de adaptadores de cambio de plantilla para la preparación de bibliotecas de genes expresados, por ejemplo, cadenas pesadas y livianas de anticuerpos, proporciona claras ventajas sobre otros métodos conocidos en la técnica al permitir una representación imparcial de familias de genes particulares y otros genes coexpresados. El uso de adaptadores de cambio de plantilla o adaptadores 5' agregados utilizando cualquier otro método como, entre otros, coleoTdT y ligación en extremo romo son también más compatibles con el uso de adaptadores 5' con código de barras que con GSPs 3' con código de barras u oligo con código de barras (dT)s por las razones expuestas anteriormente.

Ejemplo 55: Selección de los plasmoblastos por dispersión de avance y/o dispersión lateral en un citómetro de flujo y/o en combinación con otros marcadores de superficie celular

Los plasmoblastos son células B blásticas que se activan, han proliferado/están proliferando y se someten a una maduración de afinidad. Los plasmoblastos representan la respuesta inmune activa y al practicar los métodos y composiciones en el presente documento, se permite la construcción bioinformática de árboles evolutivos con familias clonales de anticuerpos que se unen a los antígenos de interés, ya sea una infección, una vacuna, un antígeno autoinmune o cáncer.

Los plasmoblastos son células B blásticas y son más grandes que las células B en reposo (Figura 40A-B). Por lo tanto, se pueden clasificar en un citómetro de flujo usando sus propiedades de dispersión hacia delante y hacia los lados. Como se muestra en la Figura 40c, los plasmoblastos tienen una mediana de FSC-A que es aproximadamente ~1.29-1.66x mayor que la mediana de FSC-A de otras células CD20<sup>+</sup> B, con una mediana de FSC-W que es 1.04-1.16x más grande que el resto de CD20<sup>+</sup> B Células. Los plasmoblastos también tienen una mediana de FSC-A que es de 0.85-0.98x la de los monocitos, y una mediana de FSC-W que es de 0.92-1.02x la de los monocitos según lo determinado por el intervalo de confianza del 95%. En este caso, FSC-A y FSC-H podrían ser intercambiables y equivalentes, ya que FSC-A y FSC-H se escalan en citómetros de flujo calibrados para obtener los mismos valores. De manera similar para SSC-A (y SSC-H, debido al escalado) y SSC-W, los plasmoblastos tienen una SSC-A mediana que es de 0.74-2.56x que la de las células B CD20<sup>+</sup> y 0.21-0.84x que la de los monocitos, y una mediana SSC-W que es 1.01-1.20x que las células B CD20<sup>+</sup> y 0.82-1.03x que los monocitos. La proporción de plasmoblastos a células B es representativa de la de linfocitos, ya que los linfocitos en reposo son similares en tamaño.

Un enfoque alternativo para identificar plasmoblastos es usar el percentil 20 FSC o SSC de plasmoblastos para la mediana FSC o SSC de células B CD20<sup>+</sup> B o monocitos (Figura 44D), que es 1.04-1.50x (1.02-1.11x) que de la mediana FSC-A (FSC-W) para células B CD20<sup>+</sup> y 0.70-0.88x (0.88-1.00x) para monocitos. Los plasmoblastos tienen un percentil 20 SSC-A (SSC-W) que es 0.67-1.89x (0.99-1.11x) del SSC-A mediano (SSC-W) para células B CD20<sup>+</sup> y 0.20-0.62x (0.77-0.99x) para los monocitos. Estos números permiten que un corte de activación incluya el 80% de los plasmoblastos y excluya otros linfocitos. Esto permite el uso de FSC (y/o SSC) en combinación con tinciones de color simple o doble para los plasmoblastos en plasmoblastos de selección de células individuales. Estas combinaciones pueden incluir FSC<sup>hi</sup>CD19<sup>lo</sup> (Figura 39b), CD19<sup>+</sup>FSC<sup>hi</sup> (Figura 39c) CD19<sup>+</sup>FSC<sup>hi</sup>CD20<sup>-</sup> (Figura 39d), CD19<sup>+</sup>FSC<sup>hi</sup>CD38<sup>hi</sup> (Figura 39e) y CD19<sup>+</sup>FSC<sup>hi</sup>CD27<sup>+</sup> (Figura 39f). Las células clasificadas pueden someterse a RT, PCR para el código de barras como se lleva a cabo como en la PCR "sin toma de contacto" en los materiales y métodos. La preparación posterior para la secuenciación, la clonación y la expresión son las siguientes en los ejemplos 6 y 8. Tenga en cuenta que las relaciones proporcionadas son las del intervalo de confianza del 95%, o donde el 95% de las relaciones deben estar dentro de este rango.

Ejemplo 56: Selección de los plasmoblastos por tamaño en cualquier dispositivo de cribado, como un dispositivo de microfluidos

Los plasmoblastos son células B blásticas que se activan, han proliferado/están proliferando y se someten a una maduración de afinidad. Los plasmoblastos representan la respuesta inmune activa y al practicar los métodos y composiciones en el presente documento, se permite la construcción bioinformática de árboles evolutivos con familias clonales de anticuerpos que se unen a los antígenos de interés, ya sea una infección, una vacuna, un antígeno autoinmune o cáncer.

Los plasmoblastos son células B blásticas y son más grandes que las células B en reposo. Los plasmoblastos y las células B CD20<sup>+</sup> se clasificaron mediante FACS y se tiñeron con azul de tripano para excluir las células muertas y se generaron imágenes a 200 aumentos. Se generaron imágenes de 52 plasmoblastos y 51 células B CD20<sup>+</sup> y se midió el área celular con ImageJ. Los plasmoblastos tomados fueron de entre 7.8-13µm de diámetro, y entre 48-121µm<sup>2</sup> de área, y entre 251-996µm<sup>3</sup> de volumen. Las células CD20<sup>+</sup> B no son blásticas y son más pequeñas, la mayoría tienen entre 6-8µm de diámetro, o menos de 50µm<sup>2</sup>, o más pequeñas que 268µm<sup>3</sup>, con solo 4 células de 51 más que eso (Figura 41). Cualquier dispositivo de cribado que sea capaz de separar células mayores o menores a 8 µm de diámetro o 50µm<sup>2</sup> en área o 268µm<sup>3</sup> en volumen es capaz de separar plasmoblastos de células B en reposo CD20<sup>+</sup>, con el 96% de los plasmoblastos capturados, y cribar el 92% de las células B en reposo. Un dispositivo de este tipo puede ser un tamiz fino con orificios de 8 µm de diámetro, o un dispositivo de microfluidos con canales que solo permite o evita que pasen las células de más de 8µm de diámetro o 50µm<sup>2</sup> en el área o 268µm<sup>3</sup> de volumen. Estas células se pueden seleccionar por accionadores/bombas en el dispositivo de microfluidos en pozos de manera que solo haya 1 o unas pocas células/pozo y RT, la PCR para el código de barras se puede realizar como en la PCR "sin toma de contacto" en los materiales y métodos utilizando las mismas concentraciones de reactivos. La preparación descendente para secuenciación, clonación y expresión es la siguiente en los ejemplos 6 y 8.

Ejemplo 57: los anticuerpos Anti-*Staphylococcus aureus* mejoran la fagocitosis de *S. aureus* por una línea celular de neutrófilos

Seres humanos con infecciones por *S. aureus* que montan respuestas inmunitarias efectivas contra su infección por *S. aureus*, por ejemplo los humanos que eliminan *S. aureus* sin la necesidad de terapia con antibióticos, se utilizan como fuentes de sangre periférica a partir de la cual se tiñen y clasifican los plasmoblastos de sangre periférica. Los plasmoblastos se clasifican en una sola célula y están codificados por barras, como se detalla en "PCR no Touchdown" en Materiales y métodos, y se preparan para la secuenciación como se detalla en "Preparación para la secuenciación 454 de XLR70" en Materiales y métodos. Los árboles evolutivos se construyen bioinformáticamente y se seleccionan unos pocos representantes selectos de cada familia clonal y se clonan para la expresión como anticuerpos recombinantes, como en el ejemplo 8. La cepa de *S. aureus* Wood, que tiene un ~5% de proteína A positiva, se siembra en agar soja tripticasa 5% (TCA), agar sangre y una colonia se cultiva y mantiene a 4°C como reserva. Esta reserva se actualiza semanalmente al elegir otra colonia. Se utiliza 1 mL de esta reserva para inocular el crecimiento de *S. aureus* hasta OD550 = 0.5, que es aproximadamente la fase de crecimiento media del tronco. *S. aureus* se fija ligeramente en paraformaldehído al 4% (PFA) durante 15 minutos a temperatura ambiente y se lava una vez con solución salina equilibrada de Hanks (HBSS), antes de teñir con CFSE 1µM durante 15 minutos a temperatura ambiente. Las bacterias fijas se lavan y se incuban con 10 µg/ml del anti-S recombinante expresado. anticuerpos de aureus, o 10 µg/ml de anticuerpos contra el virus de la influenza expresados como control negativo. Las bacterias se lavan dos veces. HL-60, una línea celular de neutrófilos, se activa durante 96 horas con ácido retinoico 25 µM, y se incuban con bacterias fijas marcadas de 1:1 a 1:100 durante 45 minutos a 37°C, agitando suavemente a 300 rpm en placas de 96 pozos. El HL-60 luego se lava dos veces y se analiza en un citómetro de flujo. La cantidad de marcación CFSE en HL-60s es indicativa de la cantidad de *S. aureus* fagocitado. Algunos anticuerpos anti-*S. aureus* expresados se unen a las proteínas de la superficie de las células estafilocócicas y opsonizan las bacterias, lo que lleva a un aumento de la fagocitosis.

Ejemplo 58: Los anticuerpos Anti-*Staphylococcus aureus* mejoran la destrucción de *S. aureus* mediada por neutrófilos

Seres humanos que pudieron controlar y/o eliminar de manera efectiva sus infecciones por *S. aureus* se seleccionaron como en el ejemplo relevante anterior, y se aislaron los plasmoblastos y se clasificaron las células individuales para secuenciación y clonación y expresión como en el ejemplo relevante anterior. Se aisló un aislado clínico de *S. aureus* en agar con sangre al 5% de TCA y se cultivó una colonia y se mantuvo a 4°C como reserva. Esta reserva se actualizó  
 5 semanalmente al elegir otra colonia. Se usó 1 ml de esta reserva para inocular el crecimiento de *S. aureus* hasta OD550 = 0.5, que es aproximadamente la fase de crecimiento media del tronco. *S. aureus* luego se incubó con 2 µg/ml de anticuerpos anti-*S. aureus* expresados durante 30 minutos a 4°C antes de lavar dos veces. La célula de neutrófilos HL-60 se activó durante 96 horas con 25µM de ácido retinoico, y se incubó con complemento de conejo bebé y *S. aureus* en una proporción de 1:1 a 1:100 durante 45 minutos a 37°C, agitando a 300 rpm en placas de 96 pozos.  
 10 Luego, las células HL-60 se pusieron rápidamente en hielo y se lavaron 3x para eliminar *S. aureus* unido de forma suelta. El *S. aureus* extracelular se diluyó en serie y se sembró en agar con sangre al 5% TSA y se cultivó durante la noche a 37°C. Las colonias se contaron al día siguiente para determinar el número de unidades formadoras de colonias (UFC). Una disminución en las UFC por anticuerpos recombinantes anti-*S. aureus* específicos (después de la incubación con *S. aureus*) demuestran que esos anticuerpos fueron efectivos para mediar la fagocitosis mejorada y matar, o reducir el crecimiento, de *S. aureus* por las células HL-60 (Figura 46).  
 15

Ejemplo 59: Tratamiento de ratones infectados con *Staphylococcus aureus* usando anticuerpos anti-*S. aureus* expresados en modelo de ratón

Anticuerpos anti-*S. aureus* que demuestran matanza *in vitro*, el crecimiento reducido o la actividad de unión como en el ejemplo 58 también pueden tener actividad *in vivo*. Los anticuerpos anti-*S. aureus* con actividad destructora se  
 20 aíslan de humanos infectados con *S. aureus* que pueden controlar su infección por estafilococos como en los ejemplos 55-58. A los ratones se les da una dosis letal de *S. aureus* y luego se tratan con un anticuerpo de control o un anticuerpo(s) anti-*S. aureus* recombinante con destrucción *in vitro* demostrada, crecimiento reducido o actividad de unión. Se considera que los ratones están protegidos si tienen una supervivencia más larga o una gravedad reducida de la infección según lo determinado por la prueba de supervivencia de Kaplan-Meier. Los anticuerpos anti-*S. aureus*  
 25 derivados de humanos que controlan o reducen la gravedad de sus infecciones por *S. aureus* se evalúan por su capacidad de conferir protección pasiva contra *S. aureus*.

Ejemplo 60: Uso de las dianas antigénicas de respuestas inmunes *anti-Staphylococcus aureus* eficaces para desarrollar vacunas.

Los antígenos de *S. aureus* que son el blanco de los anticuerpos anti-*S. aureus* que muestran una destrucción, crecimiento reducido o actividad de unión son buenos candidatos para una vacuna de *S. aureus*. Los vacunados que  
 30 desarrollan una respuesta fuerte contra esos antígenos específicos pueden protegerse o exhibir una gravedad reducida de la infección por *S. aureus*. Los anticuerpos anti-*S. aureus* con muerte, crecimiento reducido o actividad de unión se aíslan de los humanos infectados con *S. aureus* que son capaces de controlar su infección por *S. aureus*, y sus antígenos diana se identifican usando espectrometría de masas como en los ejemplos 55-58. Los ratones se  
 35 vacunan con un vector simulado o se vacunan con antígenos candidatos de *S. aureus* y luego se refuerzan dos veces durante un período de dos meses. Los ratones pueden inmunizarse con antígenos candidatos de *S. aureus* individualmente o en combinación. El título de anticuerpos del antígeno anti-*S. aureus* se confirma por ELISA. Los ratones luego son desafiados con una dosis letal de *S. aureus*. Se considera que los ratones están protegidos contra  
 40 *S. aureus* si tienen una supervivencia más larga según lo determinado por la prueba de supervivencia de Kaplan-Meier. La inmunización contra estos antígenos de *S. aureus* seleccionados por lo tanto confiere protección o reduce la gravedad de la infección, lo que demuestra que las composiciones y los métodos en este documento pueden ayudar en el diseño de la vacuna.

Ejemplo 61: Tratamiento de humanos infectados con *Staphylococcus aureus* usando anticuerpos recombinantes anti-*S. aureus* con destrucción *in vivo*, crecimiento reducido o actividad de unión en humanos

Los anticuerpos derivados de humanos que controlan su infección por *S. aureus* y exhiben destrucción *in vitro* e *in vivo*, crecimiento reducido o actividad de unión como en los ejemplos 55-59 se pueden usar para tratar pacientes infectados por *S. aureus*. Los anticuerpos se obtienen y se analizan para determinar su actividad *in vitro* e *in vivo* como  
 45 en los ejemplos 55-59. Los anticuerpos monoclonales anti-*S. aureus* fabricados por Good Manufacturing Practice (GMP) pueden administrarse por vía intravenosa o subcutánea a humanos infectados con *S. aureus*, especialmente a pacientes infectados con *S. aureus* resistente a la metililina (SARM) u otras cepas de *S. aureus* resistente a los medicamentos, y se compararon a los antibióticos a lo largo de la eficacia. Se considera que los anticuerpos anti-*S. aureus* tienen utilidad terapéutica si los pacientes están protegidos contra infecciones invasivas por *S. aureus*, tienen infecciones menos graves por *S. aureus* y/o se recuperan más rápidamente que los pacientes que recibieron  
 50 antibióticos solos o que no recibieron antibióticos. Los anticuerpos anti-*S. aureus* recombinantes pueden administrarse terapéuticamente a pacientes con infecciones por *S. aureus* activo para reducir la gravedad de la infección y/o para mejorar la eliminación de la infección, así como profilácticamente a poblaciones de pacientes de alto riesgo, como los pacientes en hemodiálisis por insuficiencia renal. fracaso, pacientes ingresados en el hospital o pacientes con un examen de detección positivo para *S. aureus* o MRSA.  
 55

Ejemplo 62: Vacunación contra estafilococo aureus utilizando antígenos de *S. aureus* identificados para conferir inmunidad protectora en humanos

5 Los antígenos de *S. aureus* que se dirigen a los anticuerpos anti-*S. aureus* que exhiben actividad de destrucción o unión in vivo, y cuando se vacunan en un modelo de ratón confieren protección contra el desafío de *S. aureus*, pueden ser buenos candidatos para una vacuna profiláctica en humanos. Los anticuerpos anti-*S. aureus* se derivan de seres humanos que controlan sus infecciones por *S. aureus* y se clonan, expresan y analizan para determinar la destrucción *in vivo*, el crecimiento reducido o la actividad de unión y como candidatos de vacuna en ratones como en los ejemplos 55-58 y 59. Los seres humanos reciben entonces la vacuna de *S. aureus* que contiene antígenos *S. aureus* que son los objetivos de los anticuerpos anti-*S. aureus* con muerte, crecimiento reducido o actividad de unión, con un placebo como control. Se realiza un seguimiento de las cohortes por su incidencia o gravedad de las infecciones por *S. aureus*. La vacuna se considera exitosa si disminuye la incidencia de la cohorte vacunada o la gravedad de las infecciones por *S. aureus* en comparación con la cohorte de placebo.

Ejemplo 63: Monitorización de las respuestas inmunes inducidas por vacunas candidatas a *Staphylococcus aureus* como correlato de protección.

15 Después de la inmunización de humanos con una vacuna de *S. aureus* candidata como en el ejemplo 62, la respuesta de la vacuna puede monitorizarse determinando si se obtienen familias clonales sólidas contra los antígenos de interés de *S. aureus* objetivo. La sangre se extrae entre los 7 y 14 días posteriores a la vacunación y los plasmoblastos se seleccionan una sola célula, se codifican por barras y se secuencian por 454 como se detalla en "PCR no Touchdown" y "preparación para la secuenciación 454 de XLR70" en los materiales y métodos. Los árboles evolutivos se dibujan y luego se clonan 2-3 miembros de cada familia clonal y se expresan como en el ejemplo 8 y se analizan para determinar su unión a los antígenos de estafilococos de interés en un ELISA. Se espera que los humanos que tienen una fuerte respuesta inmune anti-*S. aureus* inducida por la vacuna exhiban grandes familias clonales contra los antígenos de *S. aureus* dirigidos a respuestas inmunes humanas eficaces. Tal enfoque tiene el potencial de proporcionar un correlato de protección para una vacuna contra *S. aureus* y, al hacerlo, permite que los ensayos clínicos y el desarrollo se optimicen. Esta monitorización de repertorio inmune de anticuerpos y/o TCR permitiría una evaluación rápida de la probabilidad de que una vacuna candidata proporcione eficacia.

Ejemplo 64: Tratamiento de ratones con adenocarcinoma de pulmón usando anticuerpos de adenocarcinoma anti-pulmón recombinantes.

30 El anticuerpo anti-adenocarcinoma de pulmón que se une a una proteína de la superficie celular u otras proteínas de adenocarcinoma de pulmón puede ser útil como portador para dirigir toxinas a células de adenocarcinoma de pulmón u otras moléculas expresadas por células de adenocarcinoma de pulmón. Anticuerpos anti-adenocarcinoma de pulmón con actividad de unión a la superficie celular u otros antígenos de adenocarcinoma de pulmón se aíslan de un paciente de cáncer de adenocarcinoma de pulmón no progresivo a largo plazo como en el ejemplo 11. A los ratones desnudos se les administra una inyección subcutánea de la línea celular H1650 de adenocarcinoma de pulmón y el tumor dejó crecer durante una semana. El anticuerpo anti-pulmón adenocarcinoma luego se conjuga con una toxina, como la toxina de la difteria que carece del dominio-R, que es el dominio de unión celular y permite la toxina de la difteria en la célula. La toxina diftérica que carece del dominio-R, por lo tanto, es letal solo para las células de adenocarcinoma de pulmón a las que el anticuerpo se une y administra la carga útil de la toxina de la difteria. Se usa como control el anticuerpo de control conjugado con la toxina de la difteria sin el dominio R. Se considera que el anticuerpo pulmón-adenocarcinoma ha entregado con éxito su carga útil para matar las células de adenocarcinoma si la carga tumoral disminuye más que en el control. Alternativamente, en ciertos casos, el propio anticuerpo recombinante puede ser capaz de mediar la muerte de células tumorales o prevenir el crecimiento de células tumorales (en ausencia de una toxina conjugada).

45 Ejemplo 65: Tratamiento de pacientes con adenocarcinoma de pulmón usando el anticuerpo expresado anti-adenocarcinoma de pulmón.

50 El anticuerpo anti-adenocarcinoma pulmonar que se une al antígeno de la superficie celular puede ser útil como portador para dirigir las toxinas a las células del adenocarcinoma pulmonar. Los anticuerpos contra el adenocarcinoma pulmonar con actividad de unión a la superficie celular se aíslan de un paciente con cáncer adenocarcinoma pulmonar no progresivo a largo plazo, como en el ejemplo 11. El anticuerpo monoclonal GMP u otro adenocarcinoma pulmonar anti-pulmonar pueden administrarse por vía intravenosa o subcutánea a pacientes con adenocarcinoma pulmonar, especialmente a pacientes cuyas células de adenocarcinoma biopsiadas expresaron niveles altos de antígeno de superficie celular dirigido por el anticuerpo monoclonal. El (los) antígeno(s) de adenocarcinoma pulmonar(s) monoclonal (es) u otros miembros de las familias clonales de las que se derivan, se pueden usar para teñir inmunohistoquímicamente una muestra de biopsia del adenocarcinoma pulmonar de un paciente individual para obtener información sobre los niveles de expresión de antígeno tumoral y esta información puede usarse para determinar si un paciente individual es probable que responda a la terapia con este anticuerpo monoclonal. Los anticuerpos de adenocarcinoma anti-pulmón pueden conjugarse con una toxina, como la toxina de la difteria que carece del dominio-R, que es el dominio de unión celular y permite la toxina de la difteria en la célula. La toxina diftérica que carece del dominio-R, por lo tanto, es letal solo para las células de adenocarcinoma de pulmón a las que el

anticuerpo se une y administra la carga útil de la toxina de la difteria. La quimioterapia estándar se usa para el tratamiento del grupo de comparación. Se considera que los anticuerpos anti-adenocarcinoma han entregado su carga útil y tienen utilidad terapéutica si los pacientes sobreviven más tiempo o exhiben tiempos más largos antes de la recaída o la progresión. Alternativamente, en ciertos casos, el propio anticuerpo recombinante, contra los antígenos del adenocarcinoma de pulmón, puede ser capaz de mediar la muerte de las células tumorales o prevenir el crecimiento de las células tumorales (en ausencia de una toxina conjugada).

Ejemplo 66: Vacunación terapéutica contra el adenocarcinoma de pulmón utilizando antígenos identificados en humanos.

Los antígenos de la superficie celular unidos por los anticuerpos anti-adenocarcinoma de pulmón pueden usarse en una vacuna terapéutica para tratar el adenocarcinoma de pulmón establecido o para protegerlo contra el desarrollo del mismo. Anticuerpos anti-pulmón adenocarcinoma con actividad de unión a la superficie celular se aíslan a largo plazo de paciente(s) con cáncer de adenocarcinoma no progresivo de pulmón como en el ejemplo 11, y el antígeno diana se identifica mediante inmunoprecipitación o inmunotransferencia y espectrometría de masas. A los vacunados se les administra una vacuna que contiene el (los) antígeno(s) de adenocarcinoma de pulmón de interés o una vacuna de control. Luego se hace seguimiento a las cohortes y se realiza un seguimiento de su incidencia o progresión del adenocarcinoma de pulmón. Se considera que la vacuna es exitosa si los humanos vacunados con el antígeno objetivo tienen una supervivencia prolongada o un tiempo prolongado hasta la recaída en comparación con el grupo de comparación de estándar de atención.

Ejemplo 67: Monitorización inmune de la vacunación de adenocarcinoma de pulmón en cuanto a la eficacia de la respuesta.

Después de la inmunización de humanos con la vacuna de adenocarcinoma de pulmón como en el ejemplo 66, la respuesta de la vacuna puede monitorizarse determinando si se obtienen familias clonales robustas contra antígeno(s) de adenocarcinoma de interés en la vacuna. La sangre se extrae entre los 7 y 14 días posteriores a la vacunación y los plasmoblastos se clasifican en una sola célula y están codificados con códigos de barras, y la secuenciación 454 se realiza como se detalla en "PCR no Touchdown" y "preparación para la secuenciación 454 de XLR70" en los materiales y métodos. Los árboles evolutivos se dibujan y luego se clonan 2-3 miembros de cada familia clonal y se expresan como en el ejemplo 8 y se analizan para determinar su unión a los antígenos estafilococos de interés en un ELISA. Se espera que los humanos que tienen una fuerte respuesta inmune tengan grandes familias clonales contra los antígenos de interés del adenocarcinoma de pulmón y/o muchas familias clonales contra los antígenos de interés del adenocarcinoma. Esta monitorización inmunológica nos permite predecir rápidamente la eficacia de una vacuna candidata.

Ejemplo 68. Uso de Superscript III para el cambio de plantilla durante la transcripción reversa.

En nuestros métodos, la región de identificación de la muestra y las regiones del adaptador se agregan durante la transcriptasa reversa. Estos utiliza la actividad de cola de 3' y la actividad de cambio de plantilla de las transcriptasas reversas de la ARNasa H<sup>-</sup>. Con mayor frecuencia, se utiliza una transcriptasa reversa como Superscript II (Invitrogen) a su temperatura de trabajo de 42°C. Se ha informado que las transcriptasas reversas MMLV H<sup>-</sup> que también se han diseñado para la estabilidad térmica, como Superscript III, con una temperatura de trabajo recomendada de 50°C, no tienen esta actividad de cola de 3' y, por lo tanto, no tienen capacidad de cambio de plantilla ([http://tools.invitrogen.com/content/sfs/ProductNotes/F\\_Superscript%20III%20Enzyme%20RD-MKT-TL-HL0506021.pdf](http://tools.invitrogen.com/content/sfs/ProductNotes/F_Superscript%20III%20Enzyme%20RD-MKT-TL-HL0506021.pdf)?ProductNoteId=36). Sin embargo, en la Figura 42, se muestra que el Superscript III tiene actividad de cambio de plantilla y cola de 3'. Esta propiedad es débil a 50°C, la temperatura de transcripción reversa recomendada para Superscript III, y puede explicar por qué la actividad de relaves en 3' de Superscript III no se ha informado anteriormente. Sin embargo, la actividad de la cola 3' y el cambio de plantilla aumentan significativamente a medida que la temperatura de RT se redujo de 50°C a 45.5°C a 42°C. Se espera que todas las transcriptasas reversas de la ARNasa H MMLV que se han diseñado para la estabilidad térmica también tengan actividad de relaves 3' a temperaturas de trabajo más bajas, es decir, entre 42°C y 50°C.

Ejemplo 69. Análisis del gen coexpresado para identificar los anticuerpos asociados con las células de memoria B y las respuestas de las células plasmáticas, así como la orientación a tejidos específicos.

El código de barras de todo el ADNc producido por las células B, las células T u otras células clasificadas en pozos individuales como se describe en los métodos y composiciones de este documento permite la caracterización de la coexpresión de genes en plasmoblastos, otras células B, células T y otras células a nivel de una sola célula. Esto permite el uso de genes coexpresados para identificar los anticuerpos específicos y TCR expresados por las células B y T que han sido inducidos para diferenciarse en células B de memoria, células plasmáticas, células T de memoria, tipos específicos de células T efectoras (por ejemplo, células T reguladoras Th1, Th2, Th17 o T) o inducidas para que se alojen en un tejido o sitio específico (por ejemplo, el tracto gastrointestinal, la piel o el cerebro). El código de barras de todo el ADNc producido por la célula individual o la colección de células en una muestra específica permite el uso de cebadores de PCR 3' adicionales para la 1.<sup>a</sup> y 2.<sup>a</sup> PCR para caracterizar la coexpresión de dichos genes específicos. Los cebadores en 5' siguen siendo los mismos que los utilizados para amplificar los genes de las regiones variables.

Además, el análisis de genes coexpresados permite el análisis bioinformático de las relaciones entre la maduración de afinidad de las familias clonales y la coexpresión de genes asociados con la diferenciación de células B a células B de memoria, plasmoblastos de vida corta y células plasmáticas de larga vida (Tabla 34), la diferenciación de células T ingenuas o de memoria a Tregs o células Th1, Th2, Th17 (Tabla 35), o la localización de células B o T en sitios específicos. Dicho análisis puede precisar aún más los anticuerpos críticos o TCR que median una respuesta inmune efectiva.

Por ejemplo, los PMBC derivados de individuos con una respuesta inmune se utilizan para plasmoblastos de selección de células individuales. Los métodos y composiciones del presente documento se utilizan para analizar la coexpresión de genes asociados con la fusión de plasmoblastos en diferentes tejidos (véase Tabla 36). El análisis bioinformático de los conjuntos de datos identifica los anticuerpos asociados con la secreción en diferentes ubicaciones corporales. Estos genes de anticuerpos se expresan de forma recombinante para caracterización en ensayos de selección in vitro como en el Ejemplo 8.

Los expertos en la materia entenderán que se pueden realizar diversos cambios en la forma y los detalles sin apartarse del espíritu y alcance de los diversos aspectos de la invención.

Debe observarse que, tal como se utiliza en la especificación y en las reivindicaciones adjuntas, las formas singulares "una", "una" y "el/la" incluyen referentes plurales, a menos que el contexto indique claramente lo contrario.

Todas las referencias, patentes concedidas y solicitudes de patente citadas en el cuerpo de la presente especificación se incorporan aquí como referencia en su totalidad, para todos los fines.

TABLAS

Tabla 1 - Cebadores y moléculas adaptadoras		
seq id n.º	DESCRIPCIÓN	Secuencia
	Adaptador de ID de muestra	CACGACCGGTGCTCGATTTAG (ID de muestra) GGG
	Cebador FW largo 1	GAGAGACTGACAGCGTATCGCCTCCCTCGCGCCATCAG (placa-ID) CACGACCGGTGCTCGATTTAG
	Cebador FW corto 1	GAGAGACTGACAGCGTATCGCCTC
	kappa GSP1	CGATTGGAGGGCGTTATCCAC
	lambda GSP1	TYTGTGGGACTTCCACTGCTC
	gamma GSP1	TCTTGTCCACCTTGGTGTTGCTG
	Cebador de avance 2	CGTATCGCCTCCCTCGCG
	Cebador kappa GSP largo 2	CTATGCGCCTTGCCAGCCCGCTCAG (placa-ID) TCAGATGGC GGGAAGATGAAGAC
	Cebador lambda GSP largo 2	CTATGCGCCTTGCCAGCCCGCTCAG (placa-ID) GAGGAGGGY GGGAACAGAGTGAC
	gamma GSP largo 2	CTATGCGCCTTGCCAGCCCGCTCAG (placa-ID) GGAAGTAGT CCTTGACCAGGCAG

ES 2 716 013 T3

Tabla 1 - Cebadores y moléculas adaptadoras		
seq id n.º	DESCRIPCIÓN	Secuencia
	kappa GSP 2s	CTATGCGCCTTGCCAGCCCGCTCAGTCAGATGGCGGGAA GATGAAGAC
	lambda GSP 2s	CTATGCGCCTTGCCAGCCCGCTCAGGAGGAGGGY GGGAACAGAGTGAC
	gamma GSP 2s	CTATGCGCCTTGCCAGCCCGCTCAGGGAAGTAGTCCTTG ACCAGGCAG
	Cebador RV 2	CTATGCGCCTTGCCAGCCC
<p>*Para las secuencias de ID de muestra, véase la Tabla 2. Para las secuencias de ID de placa, véase la Tabla 3. Kappa GSP 2s, lambda GSP 2s y gamma GSP 2s son idénticos a los cebadores kappa, lambda y gamma GSP largos 2, excepto que no tienen una secuencia de placa de identificación. La secuencia de ID de la placa no es necesaria cuando se realizan ejecuciones XL+, o cuando solo se desean lecturas de avance cuando se realizan ejecuciones XLR70 con la química LibA de Titanio.</p> <p>** las secuencias de los cebadores se diseñaron para poder amplificar todas las diferentes variantes de genes constantes que se encuentran en la base de datos IMGT (<a href="http://imgt.cines.fr/">http://imgt.cines.fr/</a>).</p>		

Tabla 2 - ID de muestra			
Pozo	ID muestra	Secuencia de ID de muestra	SEQ ID N°
A1	1	ACGTCTCATCA	
A2	2	ACTCATCTACA	
A3	3	AGAGCGTCACA	
A4	4	AGTAGTGATCA	
A5	5	ATAGATAGACA	
A6	6	ATCTACTGACA	
A7	7	CACGTGTCGCA	
A8	8	CATACTCTACA	
A9	9	CGAGACGCGCA	



ES 2 716 013 T3

Tabla 2 - ID de muestra			
Pozo	ID muestra	Secuencia de ID de muestra	SEQ ID N°
A10	10	CGTCGATCTCA	
A11	11	CTACGACTGCA	
B1	12	TAGTGTAGATA	
B2	13	TCTAGCGACTA	
B3	14	TGTGAGTAGTA	
B4	15	ACAGTATATAA	
B5	16	AGCTCACGTAA	
B6	17	TCGATAGTGAA	
B7	18	TCGCTGCGTAA	
B8	19	TGAGTCAGTAA	
B9	20	TGTAGTGTGAA	
B10	21	TGTCGTCGCAA	
B11	22	ACGACAGCTCA	
C1	23	TACACGTGATTAGGGATT	
C2	24	TACAGATCGTTAGGGAAA	
C3	25	TAGTGTAGATTTGGGTTT	
C4	26	TCT AGCGACTTTGGGTTT	
C5	27	ACGCGATCGAAGGGTTT	
C6	28	AGCTCACGTATTGGGTTT	
C7	29	AGTGCTACGAAGGGAAA	
C8	30	TCTGACGTCAAAGGGAAA	

Tabla 2 - ID de muestra			
Pozo	ID muestra	Secuencia de ID de muestra	SEQ ID N°
C9	31	ACGTCTCATCAAGGAAGGAA	
C10	32	TATAGACATCAACACACAAA	
C11	33	AGTGAGCTCGTTGGGTTT	
D1	34	ATCGTCTGTGTTGGGTTT	
D2	35	CACACGATAGTTGGGTTT	
D3	36	CTGCGTCACGATCTCTCTT	
D4	37	TAGCATACTGTTGGGTTT	
D5	38	TATCTGATAGTTGGGTTT	
D6	39	TGACATCTCGCAGTTCTTT	
D7	40	TGATAGAGCGTAACAACAGA	
D8	41	TCACGCGAGAAAGGGAAA	
D9	42	ACACATACGCAGGGAAA	
D10	43	ACTAGCAGTAA	
D11	44	CGCAGTACGAA	
E1	45	TGAGTCAGTAAGGGAAA	
E2	46	ACTCATCTACAGGGAAA	
E3	47	ACTCGCGCACAGAGAGAA	
E4	48	AGAGCGTCACAGAGAGAA	
E5	49	AGCGACTAGCAACACACAAA	
E6	50	ATCTACTGACAACACACAA	
E7	51	CATACTCTACAACACACAAA	

ES 2 716 013 T3

Tabla 2 - ID de muestra			
Pozo	ID muestra	Secuencia de ID de muestra	SEQ ID N°
E8	52	TCGAGCTCTCAGAGAGAAA	
E9	53	AGAGAGTGTGTTGGGTTT	
E10	54	ATCGTAGCAGAACACACAAA	
E11	55	CACTCGCACGTTGGGTTT	
F1	56	CAGACGTCTGAACACACAAA	
F2	57	CTCGATATAGTTGGGTTT	
F3	58	TCTGATCGAGAAGGGAAA	
F4	59	TACACACACTTAGGGATT	
F5	60	TACGTCATCAGGGAAA	
F6	61	CTACGCTCTAAGGGAAA	
F7	62	TAGTCGCATAAAGGGAAA	
F8	63	CGATCGTATAA	
F9	64	CGCGTATACAA	
F10	65	CTACGCTCTAA	
F11	66	TCACGCGAGAA	
G1	67	AGTATACATATTGGGTTT	
G2	68	TCGATAGTGAAGGGAAA	
G3	69	TCGCTGCGTAAAGGAGAAA	
G4	70	TGTAGTGTGAAGGGAAA	
G5	71	TGTCGTCGCAAGAGAGAG	
G6	72	CTACGACTGCAAGGGAAA	

ES 2 716 013 T3

Tabla 2 - ID de muestra			
Pozo	ID muestra	Secuencia de ID de muestra	SEQ ID N°
G7	73	CTCTACGCTCA	
G8	74	TAGCTCTATCA	
G9	75	TATAGACATCA	
G10	76	TCACTCATACA	
G11	77	CTAGTCACTCAAGGGAAA	
H1	78	TGTGAGTAGTTTGGGTTT	
H2	79	TGTCACACGAAGGGAAA	
H3	80	CTGTGCGTCGAAGGGAAA	
H4	81	TAGTGTAGATTCGC	
H5	82	TCGAGCTCTCTCGC	
H6	83	ATCACGTGCGTCGC	
H7	84	CAGACGTCTGTTCGC	
H8	85	TATCACTCAGTCGC	
H9	86	TGCTATAGACTTGGGTTT	
H10	87	CAGTACTGCGTTGGGTTT	
H11	88	CGACAGCGAGAACACACAAA	
A12 - H12	89	TATGCTAGTAA (control negativo)	

Tabla 3 - Placa-ID		
Placa	Secuencia de placa-ID	SEQ ID N°
1	ACGAGTGCCT	

Tabla 3 - Placa-ID		
Placa	Secuencia de placa-ID	SEQ ID N°
2	ACGCTCGACA	
3	AGACGCACTC	
4	AGCACTGTAG	
5	ATCAGACACG	
6	ATATCGCGAG	
7	CGTGTCTCTA	
8	CTCGCGTGTC	
9	TGATACGTCT	
10	CATAGTAGTG	
11	CGAGAGATAC	
12	ATACGACGTA	
13	TCACGTACTA	
14	CGTCTAGTAC	
15	TCTACGTAGC	
16	TGTACTACTC	
17	CGTAGACTAG	
18	TACGAGTATG	
19	TACTCTCGTG	
20	TAGAGACGAG	
21	TCGTCGCTCG	
22	ACATACGCGT	
23	ACGCGAGTAT	

Tabla 3 - Placa-ID		
Placa	Secuencia de placa-ID	SEQ ID N°
24	ACTACTATGT	
25	ACTGTACAGT	
26	AGACTATACT	
27	AGCGTCGTCT	
28	AGTACGCTAT	
29	ATAGAGTACT	
30	CACGCTACGT	
31	CAGTAGACGT	
32	CGACGTGACT	
33	TACACACACT	
34	TACACGTGAT	
35	TACAGATCGT	
36	TACGCTGTCT	
37	TAGTGATAGT	
38	TCGATCACGT	
39	TCGCACTAGT	
40	TCT AGCGACT	
41	TCTATACTAT	
42	TGACGTATGT	
43	TGTGAGTAGT	
44	ACAGTATATA	
45	ACGCGATCGA	

Tabla 3 - Placa-ID		
Placa	Secuencia de placa-ID	SEQ ID N°
46	ACTAGCAGTA	
47	AGCTCACGTA	
48	AGTATACATA	
49	AGTCGAGAGA	
50	AGTGCTACGA	
51	CGATCGTATA	
52	CGCAGTACGA	
53	CGCGTATACA	
54	CGTACAGTCA	
55	CGTACTCAGA	
56	CTACGCTCTA	
57	CTATAGCGTA	
58	TACGTCATCA	
59	TAGTCGCATA	
60	TATATATACA	

Tabla 4 - Cebadores de clonación		
seq id n.º	DESCRIPCIÓN	Secuencia
	Clon_Pacl	ACTGTTAATTAA (ID de muestra)
	Clon Ascl	ATTAGGCGCGCC (ID de muestra)
	Clon_Fsel	ATTAGGCCGCGCC (ID de muestra)

Tabla 4 - Cebadores de clonación		
seq id n.º	DESCRIPCIÓN	Secuencia
	Clon_AsiSI	ATTAGCGATCGC (ID de muestra)
	K_Nhela_DHFR	ACGTGCTAGCAGTTCCAGATTTCAACTGCTCATCAGA
	K_Xhold_DHFR	ACGTCTCGAGGATAGAAGTTATTCAGCAGGCACACAACA
	L_Xhol_PspXI_DHFR	ACTTGCTCGAGTCTGCYTTCCARGCMACTGT
	L_Nhel_DHFR	AGTCGCTAGCCGCRCTACTTGTGTTGCTYTGTTTG
	G_EcoRI_DHFR	AGTCGAATTCCACGACACCGTCACCGTT
	G_SacII_DHFR	ATTACCGCGGGGAAGGTGTGCACGCCG
	G_Xhol_PspXI_Lonza	ACGTCTCGAGGGTGCCAGGGGAAGACCGATG
	G_AgeI_Lonza	ACTGACCGGTTCCGGGAAGTAGTCCTTGACCAGGCA
	G_EcoRI_Lonza	TGCAGAATTCCACGACACCGTCACCG
	G_ApaI_Lonza	TGTAGGGCCCTGAGTTCCACGACACCGTC
	L_XmaI_Lonza	TGATCCCGGGATAGAAGTCACTKATSAGRCACACYAGTGTGG
	L_BstEII_Lonza	TGCAGGTCACCGCTCCCGGGTAGAAGTCACTKATSAGR
	L_Xhol_PspXI_v2_Lonza	TGATGCTCGAGTCTGCYTTCCARGCMACTGTC
	K_XmaI_Lonza	TAGTCCCGGGGATAGAAGTTATTCAGCAGGCACAC

\*La clonación directa de la clonación comienza con un sitio de restricción de flanqueo 5' y termina con las secuencias de Identificación de muestra en el extremo 3'. Esto permite que los cebadores de clonación discriminen entre secuencias con diferentes orígenes de pozos y amplifiquen de forma selectiva los amplicones con secuencias de ID de muestra específicas. Por lo tanto, hay varios cebadores de clonación de avance, cada uno específico para ID de muestra en particular. Las secuencias 3' del cebador directo de clonación son complementarias de la ID del pozo y se proporcionan en la Tabla 5. Los cebadores con nombres que comienzan con "Clon" son los cebadores directos. Los cebadores con nombres que comienzan con "K", "L" o "G" son los cebadores reversos que son regiones constantes específicas para las cadenas kappa, lambda y gamma, respectivamente. El nombre de los cebadores reversos también denota el sitio de restricción que incorporará el cebador. Finalmente, "DHFR" o "Lonza" denota si los cebadores de región constante son para el conjunto de vectores pADNc3.3 y pOptivec o vectores de Lonza pEE12.4 y pEE6.4 respectivamente, con inserciones de región constante agregadas.



ES 2 716 013 T3

Tabla 5 - Clonación de Cebadores Secuencia Especifica de Pozo		
Pozo	Secuencia	SEQ ID NO
A1	GGTGCTCGATTTAGACGTCTCATCAG	
A2	CGGTGCTCGATTTAGACTCATCTACAG	
A3	GTGCTCGATTTAGAGAGCGTCACAG	
A4	CGGTGCTCGATTTAGAGTAGTGATCA	
A5	ACCGGTGCTCGATTTAGATAGATAGACA	
A6	CGGTGCTCGATTTAGATCTACTGACAG	
A7	CTCGATTTAGCACGTGTCGCA	
A8	CGGTGCTCGATTTAGCATACTCTACA	
A9	CGATTTAGCGAGACGCGCA	
A10	TGCTCGATTTAGCGTCGATCTCA	
A11	GTGCTCGATTTAGCTACGACTGCA	
B1	GACCGGTGCTCGATTTAGTAGGTAGATAG	
B2	CGGTGCTCGATTTAGTCTAGCGACTAG	
B3	ACCGGTGCTCGATTTAGTGTGAGTAGTAG	
B4	CGACCGGTGCTCGATTTAGACAGTATATAA	
B5	GGTGCTCGATTTAGAGCTCACGTAAG	
B6	CGGTGCTCGATTTAGTCGATAGTGAA	
B7	TGCTCGATTTAGTCGCTGCGTAAG	
B8	CGGTGCTCGATTTAGTGAGTCAGTAA	
B9	CGGTGCTCGATTTAGTGTAGGTGAA	
B10	GCTCGATTTAGTGTGTCGCGCAA	
B11	GTGCTCGATTTAGACGACAGCTCA	

ES 2 716 013 T3

Tabla 5 - Clonación de Cebadores Secuencia Especifica de Pozo		
Pozo	Secuencia	SEQ ID NO
C1	CGGTGCTCGATTTAGTACACGTGATT	
C2	CGGTGCTCGATTTAGTACAGATCGTT	
C3	GACCGGTGCTCGATTTAGTAGTGTAGATTT	
C4	CGGTGCTCGATTTAGTCTAGCGACTTT	
C5	GCTCGATTTAGACGCGATCGAA	
C6	GGTGCTCGATTTAGAGCTCACGTATT	
C7	GGTGCTCGATTTAGAGTGCTACGAA	
C8	GGTGCTCGATTTAGTCTGACGTCAA	
C9	GGTGCTCGATTTAGACGTCTCATCAA	
C10	ACCGGTGCTCGATTTAGTATAGACATCAA	
C11	GGTGCTCGATTTAGAGTGAGCTCGT	
D1	GGTGCTCGATTTAGATCGTCTGTGT	
D2	GGTGCTCGATTTAGCACACGATAGT	
D3	GCTCGATTTAGCTGCGTCACGA	
D4	CCGGTGCTCGATTTAGTAGCATACTGT	
D5	GACCGGTGCTCGATTTAGTATCTGATAGT	
D6	GTGCTCGATTTAGTGACATCTCGC	
D7	CGGTGCTCGATTTAGTGATAGAGCGT	
D8	GCTCGATTTAGTCACGCGAGAAA	
D9	GGTGCTCGATTTAGACACATACGCA	
D10	CCGGTGCTCGATTTAGACTAGCAGTAA	
D11	TGCTCGATTTAGCGCAGTACGAA	

ES 2 716 013 T3

Tabla 5 - Clonación de Cebadores Secuencia Especifica de Pozo		
Pozo	Secuencia	SEQ ID NO
E1	CGGTGCTCGATTTAGTGAGTCAGTAA	
E2	CGGTGCTCGATTTAGACTCATCTACAG	
E3	GCTCGATTTAGACTCGCGCACA	
E4	GTGCTCGATTTAGAGAGCGTCACAG	
E5	GGTGCTCGATTTAGAGCGACTAGCA	
E6	CGGTGCTCGATTTAGATCTACTGACAA	
E7	CGGTGCTCGATTTAGCATACTCTACA	
E8	GTGCTCGATTTAGTCGAGCTCTCAG	
E9	CGGTGCTCGATTTAGAGAGAGTGTGT	
E10	GGTGCTCGATTTAGATCGTAGCAGA	
E11	GCTCGATTTAGCACTCGCACGT	
F1	TGCTCGATTTAGCAGACGTCTGAA	
F2	CGGTGCTCGATTTAGCTCGATATAGT	
F3	GGTGCTCGATTTAGTCTGATCGAGA	
F4	CGGTGCTCGATTTAGTACACACTT	
F5	CGGTGCTCGATTTAGTACGTCATCA	
F6	CGGTGCTCGATTTAGCTACGCTCTAA	
F7	CGGTGCTCGATTTAGTAGTCGCATAA	
F8	GGTGCTCGATTTAGCGATCGTATAA	
F9	GGTGCTCGATTTAGCGCGTATACAA	
F10	CGGTGCTCGATTTAGCTACGCTCTAA	
F11	GCTCGATTTAGTCACGCGAGAAG	

ES 2 716 013 T3

Tabla 5 - Clonación de Cebadores Secuencia Especifica de Pozo		
Pozo	Secuencia	SEQ ID NO
G1	ACGACCGGTGCTCGATTTAGAGTATACATAT	
G2	CGGTGCTCGATTTAGTCGATAGTGAA	
G3	GCTCGATTTAGTCGCTGCGTAAA	
G4	CGGTGCTCGATTTAGTGTAGTGTGAA	
G5	GCTCGATTTAGTGTGTCGTCGAA	
G6	GTGCTCGATTTAGCTACGACTGCA	
G7	GGTGCTCGATTTAGCTCTACGCTCA	
G8	CCGGTGCTCGATTTAGTAGCTCTATCA	
G9	ACCGGTGCTCGATTTAGTATAGACATCAG	
G10	CGGTGCTCGATTTAGTCACTCATACA	
G11	CGGTGCTCGATTTAGCTAGTCACTCA	
H1	CGGTGCTCGATTTAGTGTGAGTAGTTT	
H2	GGTGCTCGATTTAGTGTACACGAA	
H3	GCTCGATTTAGCTGTGCGTCGA	
H4	GACCGGTGCTCGATTTAGTAGTGTAGATTC	
H5	GGTGCTCGATTTAGTCGAGCTCTCTC	
H6	TGCTCGATTTAGATCACGTGCGT	
H7	GTGCTCGATTTAGCAGACGTCTGTC	
H8	CCGGTGCTCGATTTAGTATCACTCAGT	
H9	ACCGGTGCTCGATTTAGTGCTATAGACT	
H10	GGTGCTCGATTTAGCAGTACTGCGT	
H11	GCTCGATTTAGCGACAGCGAGA	

ES 2 716 013 T3

TABLA 6		
SEQ ID N°	DESCRIPCIÓN	SECUENCIA
	Para kappa	
	IgKC_v3	CAGATGGCGGGAAGATGAAGAC
	Para Lambda	
	IgLC_v5	CTCCCGGGTAGAAGTCAC
	IgLC_v6	TCCCGGGTAGAAGTCAC
	IgLC_v7	GCTCCCGGGTAGAAGTC
	IgLC_v8	AGTGTGGCCTTGTGG
	Para gamma	
	IgHGC_v10	GCCAGGGGGAAGACCGATG
	IgHGC_v11	CAGGGGGAAGACCGATG
	IgHGC_v13	AAGTAGTCCTTGACCAGGC
	IgHGC_v15	GAAGACCGATGGGCCCTTGG
	IgHGC_v16	AAGACCGATGGGCCCTTG
	<b>Adaptador_v1</b>	<b>G</b>
	<b>Adaptador_v2</b>	<b>GGGGG</b>
	<b>Adaptador_v3</b>	<b>rGrGrG</b>
	Univ_ seq_2	AACGCGTGACGAGAGACTGACAG
	Univ_ seq_4	TTGTTGCGTTCCTAGCCGCTATAG
	Univ_ seq_5	CTCTACGACCGGTGCTCGATTTAG
	Univ_ seq_e	CCGTCGGTCGGCAGTG

ES 2 716 013 T3

<b>Cadena liviana kappa especifica</b>	
IgKC_v4	ATGGCGGGAAGATGAAGAC
K_GSP1	GTGCTGTCCCTTGCTGTCCTGCT
K_GSP1c	AGCAGGCACACAACAGAG
K_GSP1e	TTGTGTTTCTCGTAGTCTGCTTTGC
K_GSP1f	TCTCCCCTGTTGAAGCTCTTTGTG
IgLC_v5	CTCCCGGGTAGAAGTCAC
L_GSP1	ATCTGCCTTCCAGGCCACTGTC
L_GSP1c	CTCCCGGGTAGAAGTCAC
L_GSP1d	ACRGCTCCCGGGTAGAAGTCAC
L_GSP1f	TCCACGGTGCTCCCTTCAT
L_GSP1g	GGCCGCRTACTTGTTGTTGC
L_GSP1h	GCCTTCCAGGCCACTGTAC
L_GSP1i	CTGCCTTCCAGGCCACTGTC
L_GSP1j	CTCCACGGTGCTCCCTTCA
L_GSP1k	GCTCCCTTCATGCGTGACC
L_GSP1l	TCTGTGGGACTTCCACTGCTC
L_GSP1m	GGGGCCACTGTCTTCTCCA
L_GSP1n	CTTCTGTGGGACTTCCACTGCT
L_GSP1o	ATCTGCCTTCCAGGCCACTGT
L_GSP1x	CTTYTGTGGGACTTCCACTGCTC
L_GSP1y	GCTTYTGTGGGACTTCCACTGCTC
IgHGC_v13	AAGTAGTCCTTGACCAGGC
G_GSP1c	TTCCACGACACCGTCAC

ES 2 716 013 T3

<b>Cadena liviana kappa especifica</b>		
G_GSP1d		CACGCCGCTGGTCAG
G_GSP1g		GCTGCTGAGGGAGTAGAGTCCTGA
G_GSP1h		TCTTGTCCACCTTGGTGTTGCT
G_GSP1k		GCTGGAGGGCACGGTCAC
G_GSP1		TCTTGTCCACCTTGGTGTTGCTG
G_GSP1m		TCTTGTCCACCTTGGTGTTGCT
G_GSP1n		GACTGTAGGACAGCCGGGAAGG
G_GSP1o		ACCACGCTGCTGAGGGAGTAG
G_GSP1p		TTGTCCACCTTGGTGTTGCTG
G_GSP1q		TGAGTTCCACGACACCGTCAC
G_GSP1t		GAGTTCCACGACACCGTCACC
Especifica de Kappa		
K_GSP2		ATGGCGGGAAGATGAAGAC
K_GSP2v2a		ATGGCGGGAAGATGAAGAC
K_GSP2v2b		TGGCGGGAAGATGAAGAC
K_GSP2v2d		CGGAAGATGAAGACAGATGGT
K_GSP2v2e		GCAGTTCCAGATTTCAACTG
K_GSP2v2f		ATGGTGCAGCCACAGTT
K_GSP2v2c		CAGATTTCAACTGCTCATCAGAT
K_GSP2v2g		TCAGATGGCGGGAAGATGAAGAC
Especifica de Lambda		
L_GSP2		CTCCCGGGTAGAAGTCAC

ES 2 716 013 T3

Especifica de Lambda	
L_GSP2v2c	AGGGYGGGAACAGAGTGAC
L_GSP2v2	CTCCCGGGTAGAAGTCAC
L_GSP2v2d	GAGGAGGGYGGGAACAGAGTGAC
Específica de Gamma	
G_GSP2v2c1	GCCAGGGGGAAGACCG
G_GSP2v2c2	GGAAGTAGTCCTTGACCAGG
G_GSP2b	GGAAGTAGTCCTTGACCAGGCAG
G_GSP2	AAGTAGTCCTTGACCAGGC

Tabla 7	
ID-Placa	SEQ ID NO
TATGCTAGTA	
TCACGCGAGA	
TCGATAGTGA	
TCGCTGCGTA	
TCTGACGTCA	
TGAGTCAGTA	
TGTAGTGTGA	
TGTCACACGA	
TGTCGTCGCA	
ACACATACGC	
ACAGTCGTGC	
ACATGACGAC	



ES 2 716 013 T3

ACGACAGCTC	
ACGTCTCATC	
ACTCATCTAC	
ACTCGCGCAC	
AGAGCGTCAC	
AGCGACTAGC	
AGTAGTGATC	
AGTGACACAC	
AGTGTATGTC	
ATAGATAGAC	
ATATAGTCGC	
ATCTACTGAC	
CACGTAGATC	
CACGTGTCGC	
CATACTCTAC	
CGACACTATC	
CGAGACGCGC	
CGTATGCGAC	
CGTCGATCTC	
CTACGACTGC	
CTAGTCACTC	
CTCTACGCTC	
CTGTACATAC	
TAGACTGCAC	

ES 2 716 013 T3

TAGCGCGCGC	
TAGCTCTATC	
TATAGACATC	
TATGATACGC	
TCACTCATAC	
TCATCGAGTC	
TCGAGCTCTC	
TCGCAGACAC	
TCTGTCTCGC	
TGAGTGACGC	
TGATGTGTAC	
TGCTATAGAC	
TGCTCGCTAC	
ACGTGCAGCG	
ACTCACAGAG	
AGACTCAGCG	
AGAGAGTGTG	
AGCTATCGCG	
AGTCTGACTG	
AGTGAGCTCG	
ATAGCTCTCG	
ATCACGTGCG	
ATCGTAGCAG	
ATCGTCTGTG	

ES 2 716 013 T3

ATGTACGATG	
ATGTGTCTAG	
CACACGATAG	
CACTCGCACG	
CAGACGTCTG	
CAGTACTGCG	
CGACAGCGAG	
CGATCTGTCTG	
CGCGTGCTAG	
CGCTCGAGTG	
CGTGATGACG	
CTATGTACAG	
CTCGATATAG	
CTCGCACGCG	
CTGCGTCACG	
CTGTGCGTCTG	
TAGCATACTG	
TATACATGTG	
TATCACTCAG	
TATCTGATAG	
TCGTGACATG	
TCTGATCGAG	
TGACATCTCG	
TGAGCTAGAG	

ES 2 716 013 T3

TGATAGAGCG	
TGCGTGTGCG	
TGCTAGTCAG	
TGTATCACAG	
TGTGCGCGTG	

Tabla 8	
Muestra-ID	SEQ ID NO
ACGAGTGCGT	
TAGACTGCAC	
TAGCGCGCGC	
TCATCGAGTC	
TCGCAGACAC	
TCTGTCTCGC	
TGATACGTCT	
TGAGTGACGC	
TGCTCGCTAC	
ACGTGCAGCG	
ACTCACAGAG	
AGACTCAGCG	
AGCTATCGCG	
AGTCTGACTG	
ATAGCTCTCG	
CATAGTAGTG	

ES 2 716 013 T3

Tabla 8	
Muestra-ID	SEQ ID NO
CGATCTGTCG	
CGCGTGCTAG	
CGCTCGAGTG	
CGAGAGATAC	
TGAGCTAGAG	
ATACGACGTA	
TGCGTGTGCG	
TGCTAGTCAG	
TGTATCACAG	
TGTGCGCGTG	
TCACGTAATA	
CGTCTAGTAC	
TCTACGTAGC	
TGTAATACTC	
ACGCTCGACA	
ACGACTACAG	
CGTAGACTAG	
TACTCTCGTG	
TAGAGACGAG	
TCGTCGCTCG	
ACATACGCGT	
ACGCGAGTAT	

ES 2 716 013 T3

Tabla 8	
Muestra-ID	SEQ ID NO
ACTGTACAGT	
AGACGCACTC	
AGACTATACT	
AGCGTCGTCT	
AGTACGCTAT	
ATAGAGTACT	
CACGCTACGT	
CAGTAGACGT	
CGACGTGACT	
AGCACTGTAG	
TACGCTGTCT	
TCGATCACGT	
TCGCACTAGT	
TCTATACTAT	
ATCAGACACG	
AGTCGAGAGA	
CGTACAGTCA	
CGT ACTCAGA	
ATATCGCGAG	
CTATAGCGTA	
TATATATACA	
CGTGTCTCTA	

ES 2 716 013 T3

Tabla 8	
Muestra-ID	SEQ ID NO
ACAGTCGTGC	
CTCGCGTGTC	
AGTGACACAC	
ATATAGTCGC	
CACGTAGATC	
CGACACTATC	
CTGTACATAC	

Tabla 9 - Cebadores para ligar en adaptadores para secuenciación XL+

seq id n.º	DESCRIPCIÓN	Secuencia
	5LIB-LA	CCATCTCATCCCTGCGTGTCTCCGACTCAGCGTATCGCCTCCCTCGCGCCAT
	5LIB-LB	CCTATCCCCTGTGTGCCTTGGCAGTCTCAGCGTATCGCCTCCCTCGCGCCAT
	3LIB-LA	CCATCTCATCCCTGCGTGTCTCCGACTCAGCTATGCGCCTTGCCAGCCCCTCA
	3LIB-LB	CCTATCCCCTGTGTGCCTTGGCAGTCTCAGCTATGCGCCTTGCCAGCCCCTCA

Tabla 10 - Cebadores 3' para otros genes humanos

seq id n.º	DESCRIPCIÓN	Secuencia
	región constante específica mu	
	mu GSP1	CTCTCAGGACTGATGGGAAGCC
	mu GSP2	CTATGCGCCTTGCCAGCCCGCTCAGGGGAATTCTCACAGGAGACGAGG
	región constante específica alfa	
	alfa GSP1	ATTCGTGTAGTGCTTCACGTGGC

ES 2 716 013 T3

Tabla 10 - Cebadores 3' para otros genes humanos		
seq id n.º	DESCRIPCIÓN	Secuencia
	alfa GSP2	CTATGCGCCTTGCCAGCCCGCTCAGCTCAGCGGGAAGACCTTGGG
	constante región específica TCR alfa	
	TR alfa GSP1a	cgttgacatgcaaagtcagatt
	TR alfa GSP2b	CTATGCGCCTTGCCAGCCCGCTCAGTcgggtaataggcagacagacttg
	constante región específica TRC beta	
	TR beta GSP1	CCTATCCTGGGTCCACTCGTCA
	TR beta GSP2	CTATGCGCCTTGCCAGCCCGCTCAGCTGCTTCTGATGGCTCAAACACA
Tabla 11 - Cebadores 3' para genes de ratón		
seq id n.º	DESCRIPCIÓN	Secuencia
	región constante específica mu	
	ratón_mu_GSP1	CTGAACCTTCAAGGATGCTCTTGG
	ratón_mu_GSP2	CTATGCGCCTTGCCAGCCCGCTCAGGGAAGACATTTGGAAGGACTGACTC
	región constante específica alfa	
	ratón_alfa_GSP1	TTCCTTCTGGGCACTCGACAG
	ratón_alfa_GSP2	CTATGCGCCTTGCCAGCCCGCTCAGGGGAGTGTCAGTGGGTAGATGGTG
	constante región específica gamma	
	mo_g12b_GSP1d	AGGGGACAGTCACTGAGCTGCT



ES 2 716 013 T3

Tabla 11 - Cebadores 3' para genes de ratón		
seq id n.º	DESCRIPCIÓN	Secuencia
	mo_g2ac_GSP1d	TCGAGGTTACAGTCACTGAGCTGCT
	mo_g3_GSP1d	TGGAGGGTACAGTCACCAAGCTGCT
	mo_g12_GSP2d	CTATGCGCCTTGCCAGCCCGCTCAGGGGCCAGTGGATAGACHGATGG
	mo_g3_GSP2d	CTATGCGCCTTGCCAGCCCGCTCAGGGGACCAAGGGATAGACAGATGG
	mo_g12_GSP2e	CTATGCGCCTTGCCAGCCCGCTCAGCTGGACAGGGATCCAGAGTTCC
	mo_g3_GSP2e	CTATGCGCCTTGCCAGCCCGCTCAGCTGGACAGGGCTCCATAGTTCC
	constante región específica kappa	
	ratón_kappa_GSP1	GAAGTTGATGTCTTGTGAGTGGCCT
	ratón_kappa_GSP2	CTATGCGCCTTGCCAGCCCGCTCAGTGCTCACTGGATGGTGGAA
	constante región específica lambda	
	ratón_lambda_GSP1a	ACTCTTCTCCACAGTGTCCCCTTCATG
	ratón_lambda_GSP1b	ACTCTTCTCCACAGTGTGACCTTCATG
	ratón_lambda_GSP2a	CTATGCGCCTTGCCAGCCCGCTCAGAGAGGAAGGTGGAAACASGGTGA
	ratón_lambda_GSP2b	CTATGCGCCTTGCCAGCCCGCTCAGAGGGGAAGGTGGAAACATGGTGA
	región constante específica TCR alfa	
	mo_TRA_GSP1b	TTGAAGATATCTTGGCAGGTGAAGCTT
	ratón_TRA_GSP2	
	constante región específica TRC beta	
	ratón_TRB_GSP1	GAAAGCCCATGGAAGTGCACCTTG

ES 2 716 013 T3

Tabla 11 - Cebadores 3' para genes de ratón		
seq id n.º	DESCRIPCIÓN	Secuencia
		CTATGCGCCTTGCCAGCCCGCTCAGGGGTGGAGTCACATTTCTCAGATCC
<p>* lambda GSP1a y GSP1b se mezclarán 50:50, lambda GSP2a y lambda GSP2b también se mezclarán 50:50 para amplificar todos los alelos de región constante lambda que se encuentran en IMGT. Todos los GSP1ds gamma también se deben mezclar por igual para amplificar todos los alelos gamma 1, 2a, 2b, 2c y 3 de región constante. gamma GSP2ds también se mezclarán 50:50, y gamma GSP2es también se mezclarán 50:50 para amplificar todos los alelos gamma 1, 2a, 2b, 2c y 3 regiones constantes que se encuentran en la base de datos IMGT.</p>		

Tabla 12 - Referencia de placa a paciente.	
ID-placa	Paciente
2	B Staph
3	B Staph
4	B Staph
5	B Staph
7	Adeno Pulmón
8	Adeno Pulmón
9	Adeno Pulmón
10	Adeno Pulmón
11	360 CCP+RA
12	360 CCP+RA
13	360 CCP+RA
14	361 Staph
15	361 Staph
16	361 Staph
17	361 Staph

Tabla 12 - Referencia de placa a paciente.	
ID-placa	Paciente
18	368 CCP+RF+RA
19	368 CCP+RF+RA
21	368 CCP+RF+RA
22	368 CCP+RF+RA
26	372 CCP+RF+RA
27	372 CCP+RF+RA
40	375 CCP+RF+
41	375 CCP+RF+
43	375 CCP+RF+
46	372 CCP+RF+RA
47	372 CCP+RF+RA
48	375 CCP+RF+
49	Gripe
51	Gripe
52	Gripe
53	Gripe

Tabla 13	
ID placa	Expresión Regular de la Región de Identificación de Placa
1	ACGAGTGCCT
2	ACGCTCGACA
3	AGACGCACTC
4	AGCACTGTAG

ES 2 716 013 T3

Tabla 13	
ID placa	Expresión Regular de la Región de Identificación de Placa
5	ATCAGACACG
6	ATATCGCGAG
7	CGTGTCTCTA
8	CTCGCGTGTC
9	TGATACGTCT
10	CATAGTAGTG
11	CGAGAGATAC
12	ATACGACGTA
13	TCACGTAATA
14	CGTCTAGTAC
15	TCTACGTAGC
16	TGTACTACTC
17	CGTAGACTAG
18	TACGAGTATG
19	TACTCTCGTG
20	TAGAGACGAG
21	TCGTCGCTCG
22	ACATACGCGT
23	ACGCGAGTAT
24	ACTACTATGT
25	ACTGTACAGT
26	AGACTATACT

ES 2 716 013 T3

Tabla 13	
ID placa	Expresión Regular de la Región de Identificación de Placa
27	AGCGTCGTCT
28	AGTACGCTAT
29	ATAGAGTACT
30	CACGCTACGT
31	CAGTAGACGT
32	CGACGTGACT
33	TACACACACT
34	TACACGTGAT
35	TACAGATCGT
36	TACGCTGTCT
37	TAGTGTAGAT
38	TCGATCACGT
39	TCGCACTAGT
40	TCTAGCGACT
41	TCTATACTAT
42	TGACGTATGT
43	TGTGAGTAGT
44	ACAGTATATA
45	ACGCGATCGA
46	ACTAGCAGTA
47	AGCTCACGTA
48	AGTATACATA

ES 2 716 013 T3

Tabla 13	
ID placa	Expresión Regular de la Región de Identificación de Placa
49	AGTCGAGAGA
50	AGTGCTACGA
51	CGATCGTATA
52	CGCAGTACGA
53	CGCGTATACA
54	CGTACAGTCA
55	CGTACTCAGA
56	CTACGCTCTA
57	CTATAGCGTA
58	TACGTCATCA
59	TAGTCGCATA
60	TATATATACA

Tabla 14	
ID Muestra	Expresión Regular de la Región de Identificación de Muestra
1,31	ACGTCTCATC
2,46	ACTCATCTAC
3,48	AGAGCGTCAC
4	AGTAGTGATC
5	ATAGATAGAC
6,50	ATCTACTGAC
7	CACGTGTCGC

ES 2 716 013 T3

Tabla 14	
ID Muestra	Expresión Regular de la Región de Identificación de Muestra
8,51	CATACTCTAC
9	CGAGACGCGC
10	CGTCGATCTC
11,72	CTACGACTGC
12,81,25	TAGTG TAGAT
13,26	TCTAGCGACT
14,78	TGTGAGTAGT
15	ACAGTATATA
16,28	AGCTCACGTA
17,68	TCGATAGTGA
18,69	TCGCTGCGTA
19,45	TGAGTCAGTA
20,70	TGTAGTGTGA
21,71	TGTCGTCGCA
22	ACGACAGCTC
23	TACACGTGAT
24	TACAGATCGT
27	ACGCGATCGA
29	AGTGCTACGA
30	TCTGACGTCA
75,32	TATAGACATC
33	AGTGAGCTCG

ES 2 716 013 T3

Tabla 14	
ID Muestra	Expresión Regular de la Región de Identificación de Muestra
34	ATCGTCTGTG
35	CACACGATAG
36	CTGCGTCACG
37	TAGCATACTG
38	TATCTGATAG
39	TGACATCTCG
40	TGATAGAGCG
66,41	TCACGCGAGA
42	ACACATACGC
43	ACTAGCAGTA
44	CGCAGTACGA
47	ACTCGCGCAC
49	AGCGACTAGC
52,82	TCGAGCTCTC
53	AGAGAGTGTG
54	ATCGTAGCAG
55	CACTCGCACG
56,84	CAGACGTCTG
57	CTCGATATAG
58	TCTGATCGAG
59	TACACACACT
60	TACGTCATCA



ES 2 716 013 T3

Tabla 14	
ID Muestra	Expresión Regular de la Región de Identificación de Muestra
65,61	CTACGCTCTA
62	TAGTCGCATA
63	CGATCGTATA
64	CGCGTATACA
67	AGTATACATA
73	CTCTACGCTC
74	TAGCTCTATC
76	TCACTCATAC
77	CTAGTCACTC
79	TGTCACACGA
80	CTGTGCGTCG
83	ATCACGTGCG
85	TATCACTCAG
86	TGCTATAGAC
87	CAGTACTGCG
88	CGACAGCGAG
89	TATGCTAGTA

Tabla 15	
ID Muestra	Expresión Regular de la Región de Identificación de Muestra
1,31	ACGTCTCATCAAGGAAGGAAGG+
2,46	ACTCATCTACAGG+AA+GG+
3,48	AGAGCGTCACAGAGAGAAGG+

ES 2 716 013 T3

Tabla 15	
ID Muestra	Expresión Regular de la Región de Identificación de Muestra
6,50	ATCTACTGACAACACACAAGG+
8,51	CATACTCTACAACACACAA+GG+
11,72	CTACGACTGCAAGG+AA+GG+
12,81,25	TAGTG TAGATT+GG+TT+GG+
12,81	TAGTG TAGATTGCGG+
13,26	TCTAGCGACTT+GG+TT+GG+
14,78	TGTGAGTAGTT+GG+TT+GG+
16,28	AGCTCACGTATTGG+TT+GG+
17,68	TCGATAGTGAAGG+AA+GG+
18,69	TCGCTGCGTAA+GGAGAA+GG+
19,45	TGAGTCAGTAAGG+AA+GG+
20,70	TGTAGTGTGAAGG+AA+GG+
21,71	TGTCGTGCAAGAGAGAGG+
75,32	TATAGACATCAACACACAA+GG+
66,41	TCACGCGAGAAAGG+AA+GG+
52,82	TCGAGCTCTCTGCGG+
56,84	CAGACGTCTGTGCGG+
65,61	CTACGCTCTAAGG+AA+GG+

Tabla 16. Secuencias de inserción de región constante para los vectores DHFR pADNc3.3 y pOptivec		
Seq ID no	Descripción	Secuencia
	IGHG1-G1m3 secuencia de inserción de región constante (sitios de restricción BstBI, EcoRI, SacII, BbvCI para empalmar en la región variable introducida mediante mutaciones silenciosas)	TAAGCTTACTGATAAGGCGCGCCGCGATCGCGCTCCACCAAGGACCCCTC TGTCTTCCCCCTGGCACCTCTCGAAGAGCACCTCTGGGGACACAGCAGC CCTGGGCTGCCTGGTCAAGGACTACTTCCCCGAACCAGTGACAGTGAGTT GGAATTCAGGCGCCCTGACCAGCGCGTGCACACCTCCCcGcGGTCTTC AGTCTTCAGGACTCTACTCCCTCAGCAGCGTGGTACCCTGCCCTCCAGCA GCTTGACACCCAGACCTACATCTGCAACGTGAATCACAAGCCCAGCAACA CCAAGGTGGACAAGAGAGTTGAGCCAAATCTTGTGACAAAACCTCACACA TGCCACCGTGCCAGCACCTGAACTCCTGGAGGGACCGTCAGTCTTCCTC TTCCTCCCAAACCAAGGACACCTCATGATCTCCCGGACCTAGAGGCC ACATGCGTGGTGGTGGACGTGAGCCACGAAGACCCCGAGGTCAAGTTCAA CTGGTACGTGGACGGCGTGGAGGTGCATAATGCCAAGACAAAGCCTCGG GAGGAGCAGTACAACAGCACTTACCGTGTGGTCAGCGTCTCACCGTCT GCACCAGGACTGGCTGAATGGCAAGGAGTACAAGTGAAGGTCTCCAACA AAGCCCTCCCAGCCCCATCGAGAAAACCATCTCCAAAGCCAAAGGGCAG CCCCGAGAACCACAGGTCTACACCCTGCCCCATCCCGGGAGGAGATGAC CAAGAACCAGGTCAGCCTGACCTGCCTGGTCAAAGGCTTCTATCCCAGCGA CATCGCCGTGGAGTGGGAGAGCAATGGGCAGCCGGAGAACAACACTACAAG ACCACGCCTCCCGTGCTGGACTCCGACGGCTCCTTCTTCTCTATAGCAAGC TCACCGTGGACAAGAGCAGGTGGCAGCAGGGGAACGTCTTTTCATGCTCC GTGATGCATGAGGCTCTGCACAACCACTACACACAGAAGAGCCTCTCCCTG TCCCAGGTAATGATAATGTACAACCTGACTGAGGATCCT
	IGKC-Km3 secuencia de inserción de región constante (sitios de restricción NheI, XhoI para empalmar en la región variable introducida mediante mutaciones silenciosas)	TAAGCTTACTGACTAGGCCGGCCTTAATTAACGTACGGTGGCTGCACCATC TGTCTTCATCTTCCCTCCATCTGATGAGCAGTTGAAATCTGGAAGTCTAGC GTCGTGTGCCTGCTGAATAACTTTTATCCTCGAGAGGCCAAAGTGCAGTGG AAGGTGGATAACGCCCTCCAATCCGGTAACTCCCAGGAGTCCGTACAGA GCAGGACAGCAAGGACAGCACCTACAGCCTGAGCAGCACCTGACACTGA GCAAAGCAGACTACGAGAAACACAAAGTCTACGCCTGCGAAGTCACCCAT CAGGGCTGAGCTCCCCGTCAAAAGAGCTTCAACAGGGGAGAGTGTTG ATAAATCGATACTGACTGAGGATCCT
	IGLC-Mcg-Ke-Oz- secuencia de inserción de la región constante (sitios de restricción Bsu36I, XhoI, PspXI, NheI para el empalme en la región variable introducida mediante mutaciones silenciosas)	TAAGCTTACTGACTAGGCCGGCCTTAATTAAGGTCAGCCTAAGGCTGCCCC CAGCGTCACTCTGTTCCCTCCCTCCTCTGAGGAGCTTCAAGCCAACAAGGC CACACTGGTGTGTCTCATCAGTGACTTCTACCCCGGAGCCGTGACAGTGGC TTGGAAAGCAGACTCGAGCCCCGTCAAGGCTGGAGTGGAGACCACCACAC CTTCCAAACAAGCAACAACAAGTACGCAGCTAGCAGCTACCGCAGCCTG ACCCCTGAGCAGTGGAAAGTCCCACAGAAGCTACTCCTGCCAGGTCACACAT GAAGGGAGCACCGTGGAGAAGACAGTGGCCCCCACAGAATGTTTCATGAT AAATCGATACTGACTGAGGATCCT

Tabla17. Secuencias de inserción de región constante para vectores Lonza		
seq id n.º	DESCRIPCIÓN	Secuencia
	IGHG1-G1m3 secuencia de inserción de región constante (sitios de restricción EcoRI, ApaI, AgeI,	AAGCTTGGCGCGCCTTAATTAAGCCAGCACAAAAGGCCCCAGTG TGTTTCCCTTGGCACCCCTCGAGCAAGAGTACATCTGGAGGTACA GCTGCCTTGGGCTGTTTGGTGAAAGACTATTTCCCCGAACCGGT TACTGTCTCTTGAATTCCGGGGCCCTCACCAGTGGTGTCCATA
	KpnI, XhoI para empalmar en una región variable introducida a través de mutaciones silenciosas)	CCTTTCCCGCGGTGCTTCAGAGTTCCGGTTTGTATTCCCTGTCA AGTGTCTGTGACGGTACCAAGTTCAAGTCTAGGCACCCAGACATA TATCTGTAACGTCAACCACAAGCCAAGCAACACCAAGGTTGACA AGCGGGTTGAACCTAAGTCTGTGACAAGACCCATACCTGCCCC CCATGCCCCGCACCCGAGCTCCTCGGAGGGCCTTCCGTCTTTCT TTTCCCTCCCAAACCCAAGGACACTTTGATGATCTCAAGAACAC CAGAAGTCACTTGCCTCGTGGTTGACGTGTCTCACGAAGATCCC GAAGTGAAGTTCAACTGGTACGTGGATGGGGTAGAGGTTTCATAA CGCCAAGACCAAACCCCGAGAGGAACAGTATAACTCCACCTATA GGGTAGTGTCCGTGCTCACCGTGCTCCACCAAGACTGGCTGAAT GGCAAGGAATACAAGTGCAAGGTGAGTAATAAGGCACTGCCTGC ACCCATTGAGAAGACAATATCTAAAGCAAAGGGACAGCCCAGAG AGCCCCAGGTTTATACTCTGCCACCTAGCAGAGAGGAAATGACT AAAACCAGGTCAGCCTTACTTGTCTCGTAAAAGGCTTTTATCC AAGCGACATCGCTGTGGAGTGGGAATCAAATGGCCAACCTGAGA ATAATTATAAGACTACACCTCCCGTCTTGACTCAGACGGTTCC TTCTTCTGTATAGCAAGCTCACCGTCGATAAAAAGTCGGTGGCA ACAGGGAAACGTGTTCTCATGCAGCGTCATGCACGAGGCCTTGC ACAATCATTACACCCAGAAGTCTCTGTCCCTGAGCCCTGGAAAG TGATCA
	Secuencia de inserción de región constante IGKC-Km3 (sitios de restricción XmaI, EcoRI, BstEII, DraIII para empalmar en la región variable introducida mediante mutaciones silenciosas)	AAGCTTAATTAAGGCGCGCCGAACAGTGGCTGCTCCTTCCGTGT TCATATTCCCCCATCCGACGAGCAGCTTAAATCTGGGACTGCT AGCGTCTGTGTGCTTGAATAATTTTATCCCCGGGAGGCTAA GGTACAGTGGAAGGTGGACAACGCCCTCCAATCAGGGAATTCCC AGGAGTCGGTCACCGAACAGGACAGCAAGGACTCAACCTACTCT CTGTCACTCCACTCTCACACTCAGCAAAGCCGACTATGAAAAACA CAAAGTGTATGCTTGGCAGGTGACTCATCAAGGGCTCTCCAGTC CTGTGACTAAATCCTTCAACCGAGGCGAATGCTGATCA
	Secuencia de inserción de la región constante IGLC-McG-Ke-Oz (sitios de restricción DraIII, XmaI, BstEII, PspXI para el empalme en la región variable introducida mediante mutaciones silenciosas)	AAGCTTGGCGCGCCTTAATTAAGGCCAGCCTAAAGCCGCACCCA GTGTGACCCTGTTTCCCTCCCTCCTCTGAAGAGCTCCAGGCAAAC AAAGCTACTCTGGTGTGTCTTATTAGCGATTTCTATCCCGGGGC GGTGACCGTGGCTTGAAGGCCGACTCGAGCCCAGTGAAGGCCG GAGTGAAACTACAACCCCTTCCAAACAGTCAAACAATAAATAC GCCGCTAGCAGCTATCTCTCTCTCACCCCAGAACAGTGGAATC CCACAGGTCCTATTCTTGCCAGGTCACACACGAGGGGTCAACCG TTGAGAAGACTGTTGCCCAACAGAGTGCAGCTGATCA

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados							
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo	
	LC1	Cadena liviana	K8B8	IGKV3-11*01 F	IGKJ3*01 F		
	LC1	Cadena pesada	G8B8	IGHV1-46*01 F, o IGHV1-46*03 F	IGHJ3*01 F, o IGHJ3*02 F	IGHD5-24*01 ORF	
	LC2	Cadena liviana	K8C11	IGKV3-11*01 F	IGKJ3*01 F, o IGHJ4*01 F		
	LC2	Pesado cadena	G8C11	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ4*02 F	IGHD6-25*01 F	
	LC3	Cadena liviana	K8D6	IGKV3-20*01 F	IGKJ1*01 F		
	LC3	Cadena pesada	G8D6	IGHV1-46*01 F, o IGHV1-46*03 F	IGHJ3*01 F, o IGHJ3*02 F	IGHD5-24*01 ORF	
	LC5	Cadena pesada	K10G5	IGKV3-11*01 F	IGKJ4*01 F		
	LC5	Pesado	G10G5	IGHV3-33*05 F	IGHJ4*02 F	IGHD6-13*01 F	
		Cadena					
	LC6	Cadena liviana	K8D6	IGKV3-20*01 F	IGKJ1*01 F		
	LC6	Cadena pesada	G10H2	IGHV3-11*01 F	IGHJ4*02 F	IGHD3-9*01 F	
	LC7	Cadena liviana	L8D9	IGLV2-8*01 F	IGLJ2*01 F, o IGLJ3*01 For IGLJ3*02 F		
	LC7	Cadena pesada	G8D9	IGHV3-53*02 F	IGHJ4*02 F	IGHD4-17*01 F	
	LC9	Cadena liviana	L10A1	IGLV10-54*01 F	IGLJ3*02 F		

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados						
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo
	LC9	Cadena pesada	G10A1	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ3*02 F	IGHD2-2*01 F
	LC10	Cadena liviana	K9C11	IGKV3-11*01 F	IGKJ4*01 F	
	LC10	Cadena pesada	G9C11	IGHV3-33*01 F, o IGHV3-33*06 F	IGHJ6*02 F	IGHD5-18*01 F
	LC11	Cadena liviana	L10A6	IGLV2-14*01 F	IGLJ3*02 F	
	LC11	Cadena pesada	G10A6	IGHV3-21*01 F, o IGHV3-21*04 F	IGHJ4*02 F	IGHD1-20*01 F
	LC12	Cadena liviana	L9C9	IGLV2-14*01 F	IGLJ3*02 F	
	LC12	Cadena pesada	G9C9	IGHV3-15*01 F	IGHJ4*02 F	IGHD1-26*01 F
	LC13	Cadena liviana	L9B1	IGLV2-8*01 F	IGLJ1*01 F	
	LC13	Cadena pesada	G9B1	IGHV3-66*01 F, o IGHV3-66*04 F	IGHJ3*02 F	IGHD2-8*01 F
	LC14	Cadena liviana	L9A1	IGLV2-14*01 F	IGLJ3*02 F	
	LC14	Cadena pesada	G9A1	IGHV3-15*01 F	IGHJ4*02 F	IGHD1-26*01 F
	LC15	Cadena liviana	K10A9	IGKV3-11*01 F	IGKJ3*01 F	
	LC15	Cadena pesada	G10A9	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ4*02 F	IGHD6-25*01 F
	LC16	Cadena liviana	K10D2	IGKV3-20*01 F	IGKJ2*02 F	
	LC16	Cadena pesada	G10D2	IGHV3-15*01 F	IGHJ4*02 F	IGHD4-23*01 ORF

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados							
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo	
	LC17	Cadena liviana	K8D5	IGKV3-11*01 F	IGKJ3*01 F		
	LC17	Cadena pesada	G8D5	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ4*02 F	IGHD4-23*01 ORF	
	LC18	Cadena liviana	L10D5	IGLV2-8*01 F	IGLJ3*02 F		
	LC18	Cadena pesada	G10D5	IGHV3-53*02 F	IGHJ4*02 F	IGHD4-17*01 F	
	Gripe 14	Cadena liviana	L51A6	IGLV3-25*03 F	IGLJI*01 F		
	Gripe 14	Cadena pesada	G51A6	IGHV3-30-3*01 F	IGHJ3*01 F	IGHD1-14*01 ORF	
	Gripe 15	Cadena liviana	L51C4	IGLV3-1*01 F	IGLJ2*01 F, o IGLJ3*01 For IGLJ3*02 F		
	Gripe 15	Cadena pesada	G51C4	IGHV3-30*04 F	IGHJ6*02 F	IGHD3-10*01 F	
	Gripe 16	Cadena liviana	K51G11	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		
	Gripe 16	Cadena pesada	G51G11	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ6*02 F	IGHD4-17*01 F	
	Gripe 17	Cadena liviana	K49F7	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		
	Gripe 17	Cadena pesada	G49F7	IGHV1-69*02 F, o IGHV1-69*04 F	IGHJ3*02 F	IGHD3-22*01 F	
	Gripe 18	Cadena liviana	K51D7	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		
	Gripe 18	Cadena pesada	G51D7	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ6*02 F	IGHD4-11*01 ORF	

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados							
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo	
	Gripe 19	Cadena liviana	K51D8	IGKV1-39*01 F, o IGKV1D-39*01 F	IGKJ4*01 F		
	Gripe 19	Cadena pesada	G51D8	IGHV3-49*04 F	IGHJ6*02 F	IGHD3-22*01 F	
	Gripe 20	Cadena liviana	K51G10	IGKV3-15*01 F	IGKJ2*01 F		
		Cadena pesada	G51G10	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ6*01 F	IGHD3-16*02 F	
	Gripe 21	Cadena liviana	L49A9	IGLV3-21*02 F	IGLJ3*02 F		
	Gripe 21	Cadena pesada	G49A9	IGHV3-74*01 F	IGHJ2*01 F	IGHD2-21*01 F	
	Gripe 22	Cadena liviana	L52A6	IGLV2-14*01 F	IGLJ3*02 F		
	Gripe 22	Cadena pesada	G52A6	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ6*02 F	IGHD4-23*01 ORF	
	Gripe 23	Cadena liviana	K49F11	IGKV3-11*01 F	IGKJ3*01 F		
	Gripe 23	Cadena pesada	G49F11	IGHV3-30*04 F, o IGHV3-30*08 For IGHV3-30-3*01 F	IGHJ4*02 F	IGHD3-16*01 F	
	Gripe 24	Cadena liviana	K51C8	IGKV3-15*01 F	IGKJ2*01 F		
	Gripe 24	Cadena liviana	G51C8	IGHV3-30*14 F	IGHJ5*02 F	IGHD3-22*01 F	



ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados							
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo	
	Gripe 25	Cadena liviana	K51H1	IGKV1-39*01 F, o IGKV1D-39*01 F	IGKJ4*01 F		
	Gripe 25	Cadena liviana					
	Gripe 25	Cadena liviana	K52A2	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		
	Gripe 26	Cadena pesada	G52A2	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ6*02 F	IGHD4-23*01 ORF	
	Gripe 27	Cadena liviana	L49A11	IGLV2-14*01 F	IGLJ3*02 F		
	Gripe 27	Cadena pesada	G49A11	IGHV3-7*03 F	IGHJ5*02 F	IGHD3-3*01 F	
	Gripe 28	Cadena liviana	L49C4	IGLV3-10*01 F	IGLJ2*01 F, o IGLJ3*01 F		
	Gripe 28	Cadena pesada	G49C4	IGHV3-43*01 F	IGHJ4*02 F	IGHD4-17*01 F	
	Gripe 29	Cadena liviana	L51H5	IGLV3-27*01 F	IGLJ2*01 F, o IGLJ3*01 F		
	Gripe 29	Cadena pesada	G51H5	IGHV1-2*04 F	IGHJ6*02 F	IGHD2-21*02 F	
	Gripe 30	Cadena liviana	L52B8	IGLV2-14*01 F	IGLJ3*02 F		
	Gripe 30	Cadena pesada	G52B8	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ6*02 F	IGHD4-17*01 F	
	Gripe 31	Cadena liviana	K49A5	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		
	Gripe 31	Cadena pesada	G49A5	IGHV3-30-3*01 F	IGHJ4*02 F	IGHD5-12*01 F	
	Gripe 32	Cadena liviana	K49C11	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados						
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo
	Gripe 32	Cadena pesada	G49C11	IGHV3-66*01 F, o IGHV3-66*04 F	IGHJ4*02 F	IGHD6-19*01 F
	Gripe 33	Cadena liviana	L51E9	IGLV2-14*01 F	IGLJ3*02 F	
	Gripe 33	Cadena pesada	G51E9	IGHV3-23*01 F	IGHJ4*02 F	IGHD3-16*01 F
	Gripe 34	Cadena pesada	L52G10	IGLV2-14*01 F	IGLJ3*02 F	
	Gripe 34	Cadena pesada	G52G10	IGHV3-30-3*01 F	IGHJ3*01 F, o IGHJ3*02 F	IGHD6-19*01 F
	Gripe 35	Cadena liviana	L53F10	IGLV3-21*02 F	IGLJ3*02 F	
	Gripe 35	Cadena pesada	G53F10	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ6*02 F	IGHD3-3*01 F
	Gripe 36	Cadena liviana	L52G7	IGLV2-8*01 F	IGLJ2*01 F, o IGLJ3*01 For IGLJ3*02 F	
	Gripe 36	Cadena pesada	G52G7	IGHV3-66*02 F	IGHJ4*02 F	IGHD2-15*01 F
	Gripe 37	Cadena liviana	K51E8	IGKV3-15*01 F	IGKJ1*01 F	
	Gripe 37	Cadena pesada	G51E8	IGHV1-2*04 F	IGHJ6*02 F	IGHD2-21*02 F
	Gripe 39	Cadena liviana	K53G7	IGKV1-5*03 F	IGKJ1*01 F	
	Gripe 39	Cadena liviana	G53G7	IGHV1-46*01 F, o IGHV1-46*03 F	IGHJ4*02 F	IGHD3-9*01 F
	Gripe 40	Cadena liviana	L51A5	IGLV3-10*01 F	IGLJ2*01 F, o IGLJ3*01 For IGLJ3*02 F	
	Gripe 40	Cadena liviana	G51A5	IGHV4-34*01 F	IGHJ6*02 F	IGHD5-24*01 ORF

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados							
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo	
	Gripe 41	Cadena liviana	L51B1	IGLV3-25*03 F	IGLJ1*01 F		
	Gripe 41	Cadena pesada	G51B1	IGHV1-2*02 F	IGHJ4*02 F	IGHD3-10*01 F	
	Gripe 43	Cadena liviana	L51D3	IGLV3-22*01 F	IGLJ3*02 F		
	Gripe 43	Cadena pesada	G51D3	IGHV4-39*01 F	IGHJ3*02 F	IGHD1-26*01 F	
	Gripe 44	Cadena liviana	L51D4	IGLV2-14*01 F	IGLJ3*02 F		
	Gripe 44	Cadena pesada	G51D4	IGHV3-11*01 F	IGHJ6*02 F	IGHD4-17*01 F	
	Gripe 45	Cadena liviana	L52D4	IGLV2-14*01 F	IGLJ3*02 F		
	Gripe 45	Cadena pesada	G52D4	IGHV3-30-3*01 F	IGHJ3*01 F, o IGHJ3*02 F	IGHD6-19*01 F	
	Gripe 46	Cadena liviana	L52H4	IGLV1-51*01 F	IGLJ3*02 F		
	Gripe 46	Cadena pesada	G52H4	IGHV3-23*01 F, o IGHV3-23*04 F	IGHJ4*02 F	IGHD6-13*01 F	
	S1	Cadena liviana	K3G4	IGKV3-11*01 F	IGKJ3*01 F		
	S1	Cadena pesada	G3G4	IGHV3-53*02 F	IGHJ4*02 F	IGHD2-21*01 F	
	S2	Cadena liviana	K4C4	IGKV2-28*01 F, o IGKV2D-28*01 F	IGKJ4*01 F		
	S2	Cadena pesada	G4C4	IGHV3-23*04 F	IGHJ5*02 F	IGHD6-19*01 F	
	S3	Cadena liviana	K15C6	IGKV1-5*01 F	IGKJ4*01 F		
	S3	Cadena pesada	G15C6	IGHV3-7*01 F	IGHJ1*01 F	IGHD2-2*01 F	

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados						
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo
			K15G1	IGKV1-6*01 F	IGKJ1*01 F	
	S4	Cadena pesada	G15G1	IGHV3-7*03 F	IGHJ3*02 F	IGHD6-13*01 F
	S5	Cadena liviana	K17C3	IGKV1-39*01 F, o IGKV1D-39*01 F	IGKJ2*02 F	
	S5	Cadena pesada	G17C3	IGHV1-8*02 F	IGHJ4*02 F	IGHD3-16*01 F
	S6	Cadena liviana	K3E11	IGKV3-15*01 F	IGKJ1*01 F	
	S6	Cadena pesada	G3E11	IGHV3-30*04 F, o IGHV3-30-3*01 F	IGHJ4*02 F	IGHD4-23*01 ORF
	S7	Cadena liviana	L4B8	IGLV2-8*01 F	IGLJ2*01 F, o IGLJ3*01 F	
	S7	Cadena pesada	G4B8	IGHV3-30*04 F, o IGHV3-30*10 F	IGHJ4*01 F, o IGHJ4*02 F	IGHD5-18*01 F
	S8	Cadena liviana	L4D2	IGLV2-23*01 F, o IGLV2-23*03 F	IGLJ3*02 F	
	S8	Cadena pesada	G4D2	IGHV3-33*03 F	IGHJ6*02 F	IGHD3-10*01 F
	S9	Cadena liviana	L4D6	IGLV2-8*01 F	IGLJ2*01 F, o IGLJ3*01 F	
	S9	Cadena pesada	G4D6	IGHV3-20*01 F	IGHJ4*02 F	IGHD2-2*01 F
	S10	Cadena liviana	L4F4	IGLV3-1*01 F	IGLJ3*02 F	
	S10	Cadena pesada	G4F4	IGHV4-59*01 F, o IGHV4-59*08 F	IGHJ4*02 F	IGHD3-3*01 F
	S11	Cadena liviana	L15D1	IGLV8-61*01 F	IGLJ3*02 F	
	S11	Cadena pesada	G15D1	IGHV3-7*01 F	IGHJ4*02 F	IGHD3-10*01 F

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados							
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo	
	S12	Cadena liviana	L17C6	IGLV1-47*01 F, o IGLV1-47*02 F	IGLJ3*02 F		
	S12	Cadena pesada	G17C6	IGHV3-7*03 F	IGHJ4*02 F	IGHD5-18*01 F	
	S13	Cadena liviana	L17C9	IGLV7-46*01 F	IGLJ3*02 F		
	S13	Cadena pesada	G17C9	IGHV5-a*03 F	IGHJ6*02 F	IGHD6-13*01 F	
	RA1	Cadena liviana	K11G5	IGKV3-11*01 F	IGKJ5*01 F		
	RA1	Cadena pesada	G11G5	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ4*02 F	IGHD4-23*01 ORF	
	RA2	Cadena liviana	K22C7	IGKV3-15*01 F	IGKJ1*01 F		
	RA2	Cadena pesada	G22C7	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ4*02 F	IGHD6-25*01 F	
	RA3	Cadena liviana	K26B1	IGKV3-15*01 F	IGKJ4*01 F		
	RA3	Cadena pesada	G26B1	IGHV4-39*07 F	IGHJ4*02 F	IGHD4-23*01 ORF	
	RA4	Cadena liviana	K26F5	IGKV3-15*01 F	IGKJ5*01 F		
	RA4	Cadena pesada	G26F5	IGHV3-23*01 F	IGHJ4*02 F	IGHD6-19*01 F	
	RA5	Cadena liviana	K26H1	IGKV3-11*01 F	IGKJ4*01 F		
	RA5	Cadena pesada	G26H1	IGHV3-9*01 F	IGHJ4*02 F	IGHD6-13*01 F	
	RA6	Cadena liviana	K40C5	IGKV1-39*01 F, o IGKV1D-39*01 F	IGKJ3*01 F		
	RA6	Cadena pesada	G40C5	IGHV3-30*04 F, o IGHV3-30*08 For IGHV3-30-3*01 F	IGHJ4*02 F	IGHD3-16*01 F	

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados							
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo	
	RA7	Cadena liviana	K40G1	IGKV3-11*01 F	IGKJ1*01 F		
	RA7	Cadena pesada	G40G1	IGHV4-39*01 F	IGHJ6*02 F	IGHD2-15*01 F	
	RA8	Cadena liviana	K40H4	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		
	RA8	Cadena pesada	G40H4	IGHV1-2*02 F	IGHJ5*02 F	IGHD2-21*02 F	
	RA9	Cadena liviana	K41A2	IGKV1-33*01 F, o IGKV1D-33*01 F	IGKJ4*01 F		
	RA9	Cadena pesada	G41A2	IGHV1-2*02 F	IGHJ6*02 F	IGHD1-1*01 F	
	RA10	Cadena liviana	K47A2	IGKV3-15*01 F	IGKJ5*01 F		
	RA10	Cadena pesada	G47A2	IGHV3-23*01 F	IGHJ4*02 F	IGHD6-19*01 F	
	RA11	Cadena liviana	K47E2	IGKV3-15*01 F	IGKJ5*01 F		
	RA11	Cadena pesada	G47E2	IGHV3-23*01 F	IGHJ4*02 F	IGHD6-19*01 F	
	RA12	Cadena liviana	K47F9	IGKV3-11*01 F	IGKJ4*01 F		
	RA12	Cadena pesada	G47F9	IGHV4-39*02 F	IGHJ3*02 F		
	RA13	Cadena liviana	L13B10	IGLV3-1*01 F	IGLJ1*01 F		
	RA13	Cadena pesada	G13B10	IGHV5-51*01 F	IGHJ6*02 F	IGHD6-25*01 F	
	RA14	Cadena liviana	L13G5	IGLV3-1*01 F	IGLJ1*01 F		
	RA14	Cadena pesada	G13G5	IGHV5-51*01 F	IGHJ6*02 F	IGHD6-25*01 F	

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados						
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo
	RA15	Cadena liviana	K40D6	IGKV3-20*01 F	IGKJ5*01 F	
	RA15	Cadena pesada	G40D6	IGHV4-39*07 F	IGHJ4*02 F	IGHD4-23*01 ORF
	RA16	Cadena liviana	K25D6	IGKV3-11*01 F	IGKJ3*01 F	
	RA16	Cadena pesada	G26D6	IGHV3-72*01 F	IGHJ6*03 F	IGHD4-17*01 F
	RA17	Cadena liviana	K25E9	IGKV3-11*01 F	IGKJ3*01 F	
	RA17	Cadena pesada	G25E9	IGHV3-30*03 F, o IGHV3-30*18 F o IGHV3-33*05 F	IGHJ5*02 F	IGHD3-22*01 F
	RA18	Cadena liviana	K25G4	IGKV1-27*01 F	IGKJ2*03 F	
	RA18	Cadena pesada	G25G4	IGHV3-30*09 F	IGHJ4*02 F	IGHD1-1*01 F
	RA19	Cadena liviana	K45D9	IGKV3-15*01 F	IGKJ1*01 F	
	RA19	Cadena pesada	G45D9	IGHV3-30*03 F, o IGHV3-30*18 F	IGHJ4*02 F	IGHD3-22*01 F
	RA21	Cadena liviana	L13E11	IGLV2-23*02 F	IGLJ1*01 F	
	RA21	Cadena pesada	G13E11	IGHV3-15*01 F	IGHJ4*02 F	IGHD6-13*01 F
	RA22	Cadena liviana	L13G5	IGLV3-1*01 F	IGLJ1*01 F	
	RA22	Cadena pesada	G13G5	IGHV3-7*01 F	IGHJ5*02 F	IGHD5-12*01 F
	RA23	Cadena liviana	L44C5	IGLV2-23*01 F, o IGLV2-23*02 For IGLV2-23*03 F	IGLJ1*01 F	
	RA23	Cadena pesada	G44C5	IGHV3-30*14 F	IGHJ5*02 F	IGHD7-27*01 F

ES 2 716 013 T3

Tabla 18. Todos los anticuerpos expresados						
SEQ ID N°	Anticuerpo	Tipo de cadena	Clon	V-GENE y alelo	J-GENE y alelo	D-GENE y alelo
	RA24	Cadena liviana	L44D6	IGLV2-23*01 F, o IGLV2-23*02 For IGLV2-23*03 F	IGLJ1*01 F	
	RA24	Cadena pesada	G44D6	IGHV3-30*04 F	IGHJ6*03 F	IGHD3-10*01 F
* Identidad VDJ según lo dado por V-QUEST.						

Tabla 19 - anticuerpos utilizados en ELISA Fluzone		
Anticuerpo	Tipo de cadena	Clon
Gripe 14	Cadena liviana	L51A6
Gripe 14	Cadena pesada	G51A6
Gripe 15	Cadena liviana	L51C4
Gripe 15	Cadena pesada	G51C4
Gripe 16	Cadena liviana	K51G11
Gripe 16	Cadena pesada	G51G11
Gripe 17	Cadena liviana	K49F7
Gripe 17	Cadena pesada	G49F7
Gripe 18	Cadena liviana	K51D7
Gripe 18	Cadena pesada	G51D7
Gripe 19	Cadena liviana	K51D8
Gripe 19	Cadena pesada	G51D8
Gripe 20	Cadena liviana	K51G10
Gripe 20	Cadena pesada	G51G10
Gripe 21	Cadena liviana	L49A9
Gripe 21	Cadena pesada	G49A9



ES 2 716 013 T3

Tabla 19 - anticuerpos utilizados en ELISA Fluzone		
Anticuerpo	Tipo de cadena	Clon
Gripe 22	Cadena liviana	L52A6
Gripe 22	Cadena pesada	G52A6
Gripe 23	Cadena liviana	K49F11
Gripe 23	Cadena pesada	G49F11
Gripe 25	Cadena liviana	K51H1
Gripe 25	Cadena pesada	G51H1
Gripe 26	Cadena liviana	K52A2
Gripe 26	Cadena pesada	G52A2
Gripe 27	Cadena liviana	L49A11
Gripe 27	Cadena pesada	G49A11
Gripe 29	Cadena liviana	L51H5
Gripe 29	Cadena pesada	G51H5
Gripe 30	Cadena liviana	L52B8
Gripe 30	Cadena pesada	G52B8
Gripe 33	Cadena pesada	G51E9
Gripe 34	Cadena liviana	L52G10
Gripe 34	Cadena pesada	G52G10
Gripe 35	Cadena liviana	L53F10
Gripe 35	Cadena pesada	G53F10
Gripe 37	Cadena liviana	K51E8
Gripe 37	Cadena pesada	G51E8
Gripe 39	Cadena liviana	K53G7

ES 2 716 013 T3

Tabla 19 - anticuerpos utilizados en ELISA Fluzone		
Anticuerpo	Tipo de cadena	Clon
Gripe 39	Cadena pesada	G53G7
Gripe 40	Cadena liviana	L51A5
Gripe 40	Cadena pesada	G51A5
Gripe 41	Cadena liviana	L51B1
Gripe 41	Cadena pesada	G51B1
Gripe 43	Cadena liviana	L51D3
Gripe 43	Cadena pesada	G51D3
Gripe 44	Cadena liviana	L51D4
Gripe 44	Cadena pesada	G51D4
Gripe 45	Cadena liviana	L52D4
Gripe 45	Cadena pesada	G52D4
Gripe 46	Cadena liviana	L52H4
Gripe 46	Cadena pesada	G52H4
S1	Cadena liviana	K3G4
S1	Cadena pesada	G3G4
S2	Cadena liviana	K4C4
S2	Cadena pesada	G4C4

Tabla 20 - Anticuerpos utilizados en la resonancia de plasmón superficial		
Anticuerpo	Tipo de Cadena	Clon
Gripe 14	Cadena liviana	L51A6
Gripe 14	Cadena pesada	G51A6

ES 2 716 013 T3

Tabla 20 - Anticuerpos utilizados en la resonancia de plasmón superficial		
Anticuerpo	Tipo de Cadena	Clon
Gripe 15	Cadena liviana	L51C4
Gripe 15	Cadena pesada	G51C4
Gripe 16	Cadena liviana	K51G11
Gripe 16	Cadena pesada	G51G11
Gripe 17	Cadena liviana	K49F7
Gripe 17	Cadena pesada	G49F7
Gripe 18	Cadena liviana	K51D7
Gripe 18	Cadena pesada	G51D7
Gripe 19	Cadena liviana	K51D8
Gripe 19	Cadena pesada	G51D8
Gripe 20	Cadena liviana	K51G10
Gripe 20	Cadena pesada	G51G10
Gripe 21	Cadena liviana	L49A9
Gripe 21	Cadena pesada	G49A9
Gripe 22	Cadena liviana	L52A6
Gripe 22	Cadena pesada	G52A6
Gripe 26	Cadena liviana	K52A2
Gripe 26	Cadena pesada	G52A2
Gripe 29	Cadena liviana	L51H5
Gripe 29	Cadena pesada	G51H5
Gripe 34	Cadena liviana	L52G10
Gripe 34	Cadena pesada	G52G10

Tabla 20 - Anticuerpos utilizados en la resonancia de plasmón superficial		
Anticuerpo	Tipo de Cadena	Clon
Gripe 35	Cadena liviana	L53F10
Gripe 35	Cadena liviana	G53F10
Gripe 46	Cadena liviana	L52H4
Gripe 46	Cadena pesada	G52H4

Tabla 21 - Anticuerpos utilizados en la matriz de antígenos RA		
Anticuerpo	Tipo de Cadena	Clon
RA1	Cadena liviana	K11G5
RA1	Cadena pesada	G11G5
RA2	Cadena liviana	K22C7
RA2	Cadena pesada	G22C7
RA4	Cadena liviana	K26F5
RA4	Cadena pesada	G26F5
RA5	Cadena liviana	K26H1
RA5	Cadena pesada	G26H1
RA8	Cadena liviana	K40H4
RA8	Cadena pesada	G40H4
RA9	Cadena liviana	K41A2
RA9	Cadena pesada	G41A2
RA10	Cadena liviana	K47A2
RA10	Cadena pesada	G47A2
RA11	Cadena liviana	K47E2
RA11	Cadena pesada	G47E2

ES 2 716 013 T3

Tabla 21 - Anticuerpos utilizados en la matriz de antígenos RA		
Anticuerpo	Tipo de Cadena	Clon
RA12	Cadena liviana	K47F9
RA12	Cadena pesada	G47F9
RA13	Cadena liviana	L13B10
RA13	Cadena pesada	G13B10
RA16	Cadena liviana	K25D6
RA16	Cadena pesada	G26D6
RA19	Cadena liviana	K45D9
RA19	Cadena pesada	G45D9
RA22	Cadena liviana	L13G5
RA22	Cadena pesada	G13G5
RA23	Cadena liviana	L44C5
RA23	Cadena pesada	G44C5
Gripe 14	Cadena liviana	L51A6
Gripe 14	Cadena pesada	G51A6
Gripe 26	Cadena liviana	K52A2
Gripe 26	Cadena pesada	G52A2

Tabla 22 - Anticuerpos utilizados en ELISA para histonas 2A y ELISA para CCP		
Anticuerpo	Tipo de Cadena	Clon
RA1	Cadena liviana	K11G5
RA1	Cadena pesada	G11G5
RA2	Cadena liviana	K22C7

ES 2 716 013 T3

Tabla 22 - Anticuerpos utilizados en ELISA para histonas 2A y ELISA para CCP		
Anticuerpo	Tipo de Cadena	Clon
RA2	Cadena pesada	G22C7
RA4	Cadena liviana	K26F5
RA4	Cadena pesada	G26F5
RA5	Cadena liviana	K26H1
RA5	Cadena pesada	G26H1
RA6	Cadena liviana	K40C5
RA6	Cadena pesada	G40C5
RA7	Cadena liviana	K40G1
RA7	Cadena pesada	G40G1
RA8	Cadena liviana	K40H4
RA8	Cadena pesada	G40H4
RA9	Cadena liviana	K41A2
RA9	Cadena pesada	G41A2
RA10	Cadena liviana	K47A2
RA10	Cadena pesada	G47A2
RA11	Cadena liviana	K47E2
RA11	Cadena pesada	G47E2
RA12	Cadena liviana	K47F9
RA12	Cadena pesada	G47F9
RA13	Cadena liviana	L13B10
RA13	Cadena pesada	G13B10
RA16	Cadena liviana	K25D6

ES 2 716 013 T3

Tabla 22 - Anticuerpos utilizados en ELISA para histonas 2A y ELISA para CCP		
Anticuerpo	Tipo de Cadena	Clon
RA16	Cadena pesada	G26D6
RA17	Cadena liviana	K25E9
RA17	Cadena pesada	G25E9
RA18	Cadena liviana	K25G4
RA18	Cadena pesada	G25G4
RA19	Cadena liviana	K45D9
RA19	Cadena pesada	G45D9
RA22	Cadena liviana	L13G5
RA22	Cadena pesada	G13G5
RA23	Cadena liviana	L44C5
RA23	Cadena pesada	G44C5
RA24	Cadena liviana	L44D6
RA24	Cadena pesada	G44D6

Tabla 23 - Anticuerpos utilizados en ELISA de RF		
Anticuerpo	Tipo de Cadena	Clon
RA1	Cadena liviana	K11G5
RA1	Cadena pesada	G11G5
RA2	Cadena liviana	K22C7
RA2	Cadena pesada	G22C7
RA4	Cadena liviana	K26F5
RA4	Cadena pesada	G26F5
RA5	Cadena liviana	K26H1

Tabla 23 - Anticuerpos utilizados en ELISA de RF		
Anticuerpo	Tipo de Cadena	Clon
RA5	Cadena pesada	G26H1
RA6	Cadena liviana	K40C5
RA6	Cadena pesada	G40C5
RA8	Cadena liviana	K40H4
RA8	Cadena pesada	G40H4
RA9	Cadena liviana	K41A2
RA9	Cadena pesada	G41A2
RA10	Cadena liviana	K47A2
RA10	Cadena pesada	G47A2
RA11	Cadena liviana	K47E2
RA11	Cadena pesada	G47E2
RA12	Cadena liviana	K47F9
RA12	Cadena pesada	G47F9
RA14	Cadena liviana	L13G5
RA14	Cadena pesada	G13G5

Tabla 24 - Anticuerpos utilizados en el cáncer de pulmón IHC tisular y citometría de flujo de líneas celulares de cáncer de pulmón		
Anticuerpo	Tipo de Cadena	Clon
LC1	Cadena liviana	K8B8
LC1	Cadena pesada	G8B8
LC5	Cadena liviana	K10G5
LC5	Cadena pesada	G10G5



ES 2 716 013 T3

Tabla 24 - Anticuerpos utilizados en el cáncer de pulmón IHC tisular y citometría de flujo de líneas celulares de cáncer de pulmón		
Anticuerpo	Tipo de Cadena	Clon
LC6	Cadena liviana	K8D6
LC6	Cadena pesada	G10H2
LC7	Cadena liviana	L8D9
LC7	Cadena pesada	G8D9
LC9	Cadena liviana	L10A1
LC9	Cadena pesada	G10A1
LC10	Cadena liviana	K9C11
LC10	Cadena pesada	G9C11
LC11	Cadena liviana	L10A6
LC11	Cadena pesada	G10A6
LC12	Cadena liviana	L9C9
LC12	Cadena pesada	G9C9
LC13	Cadena liviana	L9B1
LC13	Cadena pesada	G9B1
LC14	Cadena liviana	L9A1
LC14	Cadena pesada	G9A1
LC15	Cadena liviana	K10A9
LC15	Cadena pesada	G10A9
LC16	Cadena liviana	K10D2
LC16	Cadena pesada	G10D2
LC17	Cadena liviana	K8D5
LC17	Cadena pesada	G8D5

Tabla 24 - Anticuerpos utilizados en el cáncer de pulmón IHC tisular y citometría de flujo de líneas celulares de cáncer de pulmón

Anticuerpo	Tipo de Cadena	Clon
LC18	Cadena liviana	L10D5
LC18	Cadena pesada	G10D5
Flu14	Cadena liviana	L51A6
Flu14	Cadena pesada	G51A6

Tabla 25 - Anticuerpos utilizados en la tinción de la superficie de *S. aureus*

Anticuerpo	Tipo de Cadena	Clon
S1	Cadena liviana	K3G4
S1	Cadena pesada	G3G4
S2	Cadena liviana	K4C4
S2	Cadena pesada	G4C4
S3	Cadena liviana	K15C6
S3	Cadena pesada	G15C6
S4	Cadena liviana	K15G1
S4	Cadena pesada	G15G1
S6	Cadena liviana	K3E11
S6	Cadena pesada	G3E11
S7	Cadena liviana	L4B8
S7	Cadena pesada	G4B8
S8	Cadena liviana	L4D2
S8	Cadena pesada	G4D2
S9	Cadena liviana	L4D6
S9	Cadena pesada	G4D6

ES 2 716 013 T3

Tabla 25 - Anticuerpos utilizados en la tinción de la superficie de <i>S. aureus</i>		
Anticuerpo	Tipo de Cadena	Clon
S10	Cadena liviana	L4F4
S10	Cadena pesada	G4F4
S11	Cadena liviana	L15D1
S11	Cadena pesada	G15D1
S12	Cadena liviana	L17C6
S12	Cadena pesada	G17C6
S13	Cadena liviana	L17C9
S13	Cadena pesada	G17C9
Gripe 14	Cadena liviana	L51A6
Gripe 14	Cadena pesada	G51A6
Gripe 26	Cadena liviana	K52A2
Gripe 26	Cadena pesada	G52A2

Tabla 26 - Anticuerpos utilizados en el ensayo de microneutralización		
Anticuerpo	Tipo de Cadena	
Gripe 15	Cadena liviana	L51C4
Gripe 15	Cadena pesada	G51C4
Gripe 16	Cadena liviana	K51G11
Gripe 16	Cadena pesada	G51G11
Gripe 18	Cadena liviana	K51D7
Gripe 18	Cadena pesada	G51D7
Gripe 19	Cadena liviana	K51D8

ES 2 716 013 T3

Tabla 26 - Anticuerpos utilizados en el ensayo de microneutralización		
Anticuerpo	Tipo de Cadena	
Gripe 19	Cadena pesada	G51D8
Gripe 20	Cadena liviana	K51G10
Gripe 20	Cadena pesada	G51G10
Gripe 21	Cadena liviana	L49A9
Gripe 21	Cadena pesada	G49A9

Tabla 27 - anticuerpos utilizados en el ensayo de inhibición de estafilococos		
Anticuerpo	Tipo de Cadena	Clon
S6	Cadena liviana	K3E11
S6	Cadena pesada	G3E11
S9	Cadena liviana	L4D6
S9	Cadena pesada	G4D6
LC1	Cadena liviana	K8B8
LC1		G8B8

Tabla 28 - anticuerpos utilizados en estafilococo IP		
Anticuerpo	Tipo de Cadena	Clon
S1	Cadena liviana	K3G4
S1	Cadena pesada	G3G4
S2	Cadena liviana	K4C4
S2	Cadena pesada	G4C4
S3	Cadena liviana	K15C6
S3	Cadena pesada	G15C6

ES 2 716 013 T3

Tabla 28 - anticuerpos utilizados en estafilococo IP		
Anticuerpo	Tipo de Cadena	Clon
S4	Cadena liviana	K15G1
S4	Cadena pesada	G15G1
S5	Cadena liviana	K17C3
S5	Cadena pesada	G17C3
S6	Cadena liviana	K3E11
S6	Cadena pesada	G3E11
S7	Cadena liviana	L4B8
S7	Cadena pesada	G4B8
S8	Cadena liviana	L4D2
S8	Cadena pesada	G4D2
S9	Cadena liviana	L4D6
S9	Cadena pesada	G4D6
S10	Cadena liviana	L4F4
S10	Cadena pesada	G4F4
S11	Cadena liviana	L15D1
S11	Cadena pesada	G15D1
S12	Cadena liviana	L17C6
S12	Cadena pesada	G17C6
S13	Cadena liviana	L17C9
S13	Cadena pesada	G17C9
Gripe 14	Cadena liviana	L51A6
Gripe 14	Cadena pesada	G51A6

ES 2 716 013 T3

Tabla 29 - Anticuerpo usado en espec. de masas de estafilococos		
Anticuerpo	Tipo de Cadena	Clon
S4	Cadena liviana	K15G1
S4	Cadena pesada	G15G1

Tabla 30.	
Nombre	Secuencia
RT oligo	CACGACCGGTGCTCGATTTAGTTAATTAA[muestralID]AGCGATCGCTGGG
RT oligo'	CTAAATCGAGCACCGGTCGTGTGGG
De avance (cadena forkappa)	CGATTGGAGGGCGTTATCCAC
De avance (para cadena lambda)	TYTGTGGGACTTCCACTGCTC

Tabla 31.	
Nombre	Secuencia
RT oligo	CACGACCGGTGCTCGATTTAGTTAATTAA[muestra-ID]AGCGA TCGCTGGG
Cebador de superposición y extensión	CGTATCGCTCCTAGGAGCGATACGCACGACCGGTGCTCGATTTAG
Cebador LC (para cadena kappa)	CGATTGGAGGGCGTTATCCAC
Cebador LC (para cadena lambda)	TYTGTGGGACTTCCACTGCTC
Cebador HC	TCTTGTCACCTTGGTGTTGCTG

Tabla 32.	
Nombre	Secuencia
RT oligo	CGTATCGCTCCTAGGAGCGATACGTTAATTAA[muestra-ID]AGCGA TCGCTGGG
Cebador LC (para cadena kappa)	CGATTGGAGGGCGTTATCCAC

ES 2 716 013 T3

Tabla 32.	
Nombre	Secuencia
Cebador LC (para cadena lambda)	TYTGTGGGACTTCCACTGCTC
Cebador HC	TCTTGTCCACCTTGGTGTGCTG

Tabla 33.	
Nombre	Secuencia
Univ_seq_2	AACGCGTGACGAGAGACTGACAG
VK	ATGAGGSTCCCYGCTCAGCTGCTG G
VL	GGTCCTGGGCCAGTCTGCCCTG
IgKC_v3_barcode	AGGCCCTTACGACTGCGTCTTG AACAATAC CAGATGGCGGGAAGATGAAGAC
IgLC_v5_barcode	AGGCCCTTACGACTGCGTCTTGAACAATAC CTCCCGGGTAGAAGTCAC
Fixed_PCR3	AGGCCCTTACGACTGCGTCTTG

Tabla 34. Genes coexpresados asociados con la diferenciación de células B en células B de memoria, células plasmáticas de vida corta, células plasmáticas de vida larga y células secretoras de anticuerpos.				
Generación de células B de memoria.	Generación de células plasmáticas de corta vida.	Generación de células plasmáticas de larga vida.	Células secretoras de anticuerpos	
PAX-5	Blimp-1	Blimp-1	SLC7A7	CD36
. factor de transcripción asociado con Microftalmia (MITF)	Proteína de unión a caja X 1 (XBP-1)	Proteína de unión a caja X 1 (XBP-1)	IL6R	BCL2L1
			RPN2	IL21R
	IRF-4	IRF-4	PDIA4	IKZF1
		BCMA	IG HD	BACH 2

Tabla 35. Genes coexpresados asociados con la diferenciación de células T en células Treg, Th1, Th2, Th17.				
Generación de Th1	Generación de Th2	Generación de Th 17	Generación de Tregs	
T-bet	Gata-3	RORyt	FoxP3	GITR

Tabla 35. Genes coexpresados asociados con la diferenciación de células T en células Treg, Th1, Th2, Th17.

Generación de Th1	Generación de Th2	Generación de Th 17	Generación de Tregs

Tabla 36. Genes coexpresados asociados con la formación de plasmoblastos en tejidos específicos.

Alojamiento de plasmoblastos en el intestino delgado.	Alojamiento de plasmoblastos en tejidos mucosos.	Alojamiento de plasmoblastos para la piel.
CCR9	CCR10	Antígeno cutáneo asociado a linfocitos (CLA)
$\alpha 4\beta 7$		

Referencias generales

5 Burbelo, P. D., S. K. Browne, et al. (2010). "Anti-cytokine autoantibodies are associated with opportunistic infection in patients with thymic neoplasia." *Blood* 116(23): 4848-4858.

Hua, J., K. Kirou, et al. (2006). "Functional assay of type I interferon in systemic lupus erythematosus plasma and association with anti-RNA binding protein autoantibodies." *Arthritis Rheum* 54(6): 1906-1916.

May, L. T., R. Neta, et al. (1993). "Antibodies chaperone circulating IL-6. Paradoxical effects of anti-IL-6 "neutralizing" antibodies in vivo." *J Immunol* 151(6): 3225-3236.

10 Mostbock, S. (2009). "Cytokine/Antibody complexes: an emerging class of immunostimulants." *Curr Pharm Des* 15(7): 809-825.

Robinson, W. H., C. DiGennaro, et al. (2002). "Autoantigen microarrays for multiplex characterization of autoantibody responses." *Nat Med* 8(3): 295-301.

15 Watanabe, M., K. Uchida, et al. (2007). "Anti-cytokine autoantibodies are ubiquitous in healthy individuals." *FEBS Lett* 581(10): 2017-2021.

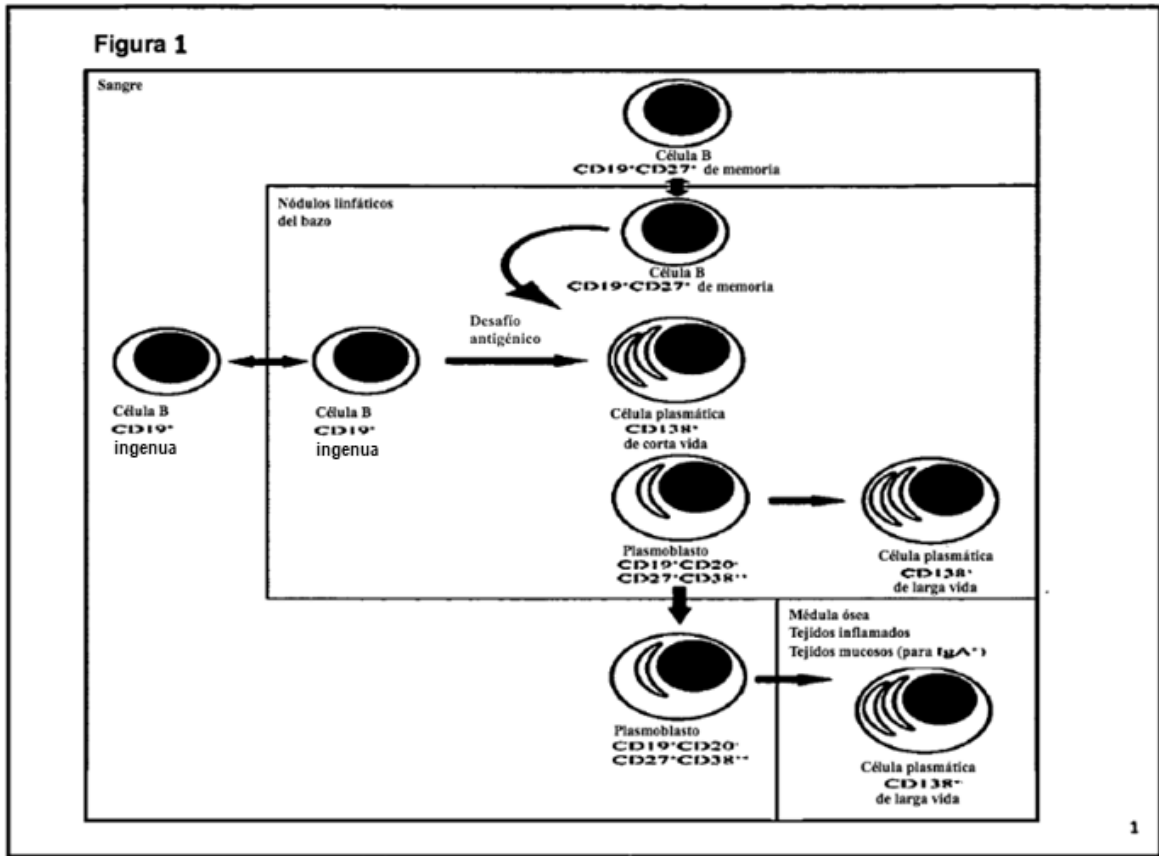
Wildbaum, G., M. A. Nahir, et al. (2003). "Beneficial autoimmunity to proinflammatory mediators restrains the consequences of self-destructive immunity." *Immunity* 19(5): 679-688.

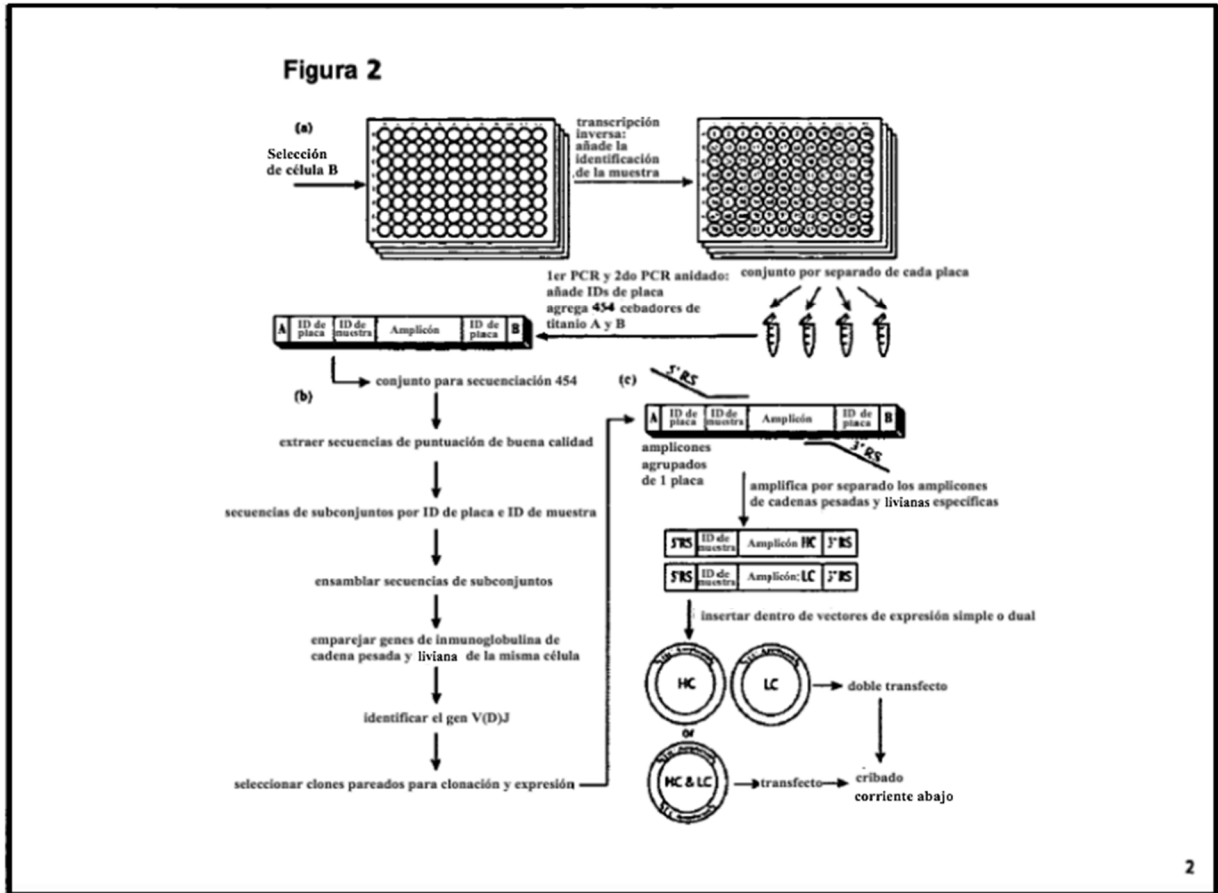
Wrarmert, J., K. Smith, et al. (2008). "Rapid cloning of high-affinity human monoclonal antibodies against influenza virus." *Nature* 453(7195): 667-671.



**REIVINDICACIONES**

1. Una biblioteca de polinucleótidos que comprende una pluralidad de composiciones, en donde la biblioteca comprende ADNc que codifican pares afines de regiones variables de cadena pesada y liviana de inmunoglobulina y en donde:
- 5 cada composición está presente en un recipiente separado, cada composición comprende
- (i) polinucleótidos que comprenden moléculas de ADNc derivadas de un solo plasmoblasto, que comprenden moléculas de ADNc que codifican un par análogo de región variable de cadena pesada de inmunoglobulina y región variable de cadena liviana de inmunoglobulina del plasmoblasto, y una región de identificación de muestra unida a las moléculas de ADNc, en donde
- 10 (ii) la secuencia de nucleótidos de la región de identificación de la muestra es distinta de la secuencia de nucleótidos de la región de identificación de la muestra de las otras composiciones presentes en otros recipientes separados en la biblioteca.
2. La biblioteca de la reivindicación 1, en donde las moléculas de ADNc están unidas a la región de identificación de la muestra por una región adaptadora.
- 15 3. La biblioteca de la reivindicación 2, en donde la región del adaptador comprende el nucleótido G en su extremo 3' y las moléculas de ADNc comprenden el nucleótido C en el extremo 3'.
4. La biblioteca de una cualquiera de las reivindicaciones 1 a 3, en donde cada composición comprende además una región de cebador universal unida a la región de identificación de la muestra.
5. La biblioteca de la reivindicación 1, en donde el plasmoblasto es una célula B murina.
- 20 6. Un método para producir uno o más polinucleótidos de interés, que comprende: obtener la biblioteca de una cualquiera de las reivindicaciones 2 a 5; y amplificar la biblioteca de polinucleótidos con un conjunto de cebadores para producir uno o más polinucleótidos de interés, en donde el polinucleótido de interés es el par análogo de regiones variables de cadena liviana y cadena pesada de inmunoglobulina.





**Figura 3**

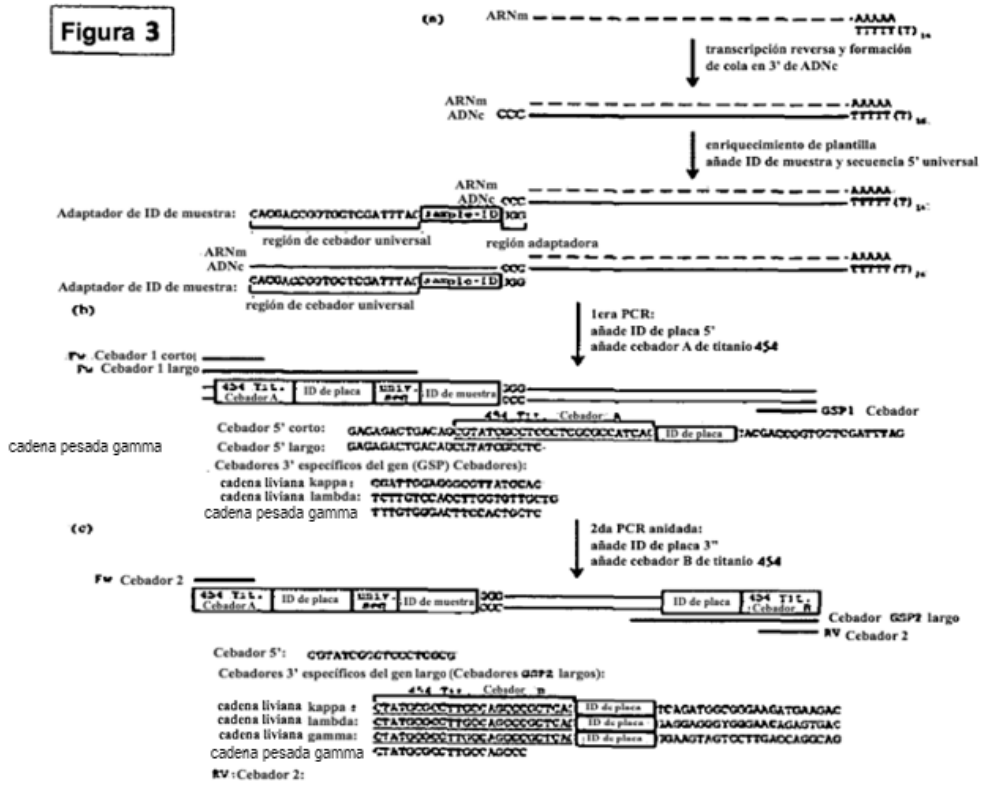


Figura 4

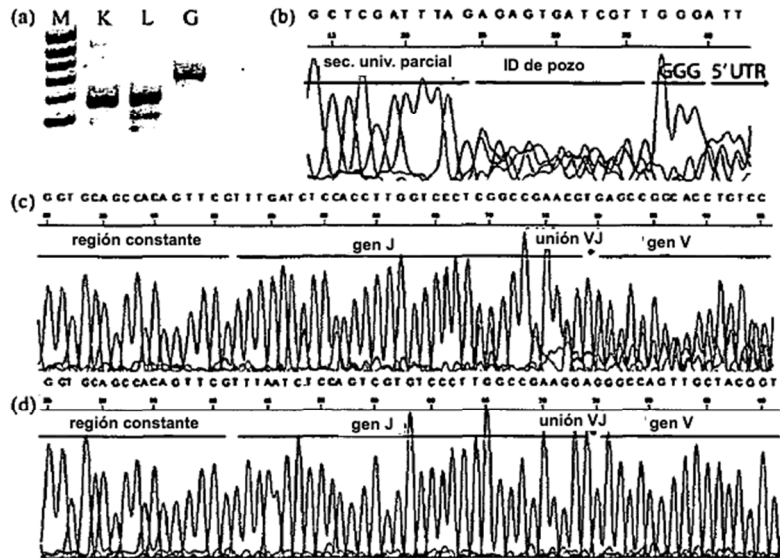


Figura 5

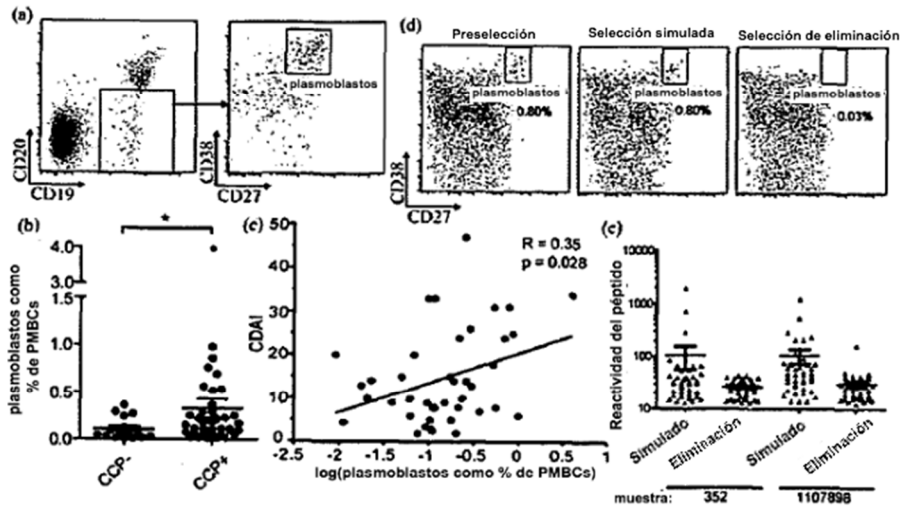


Figura 6

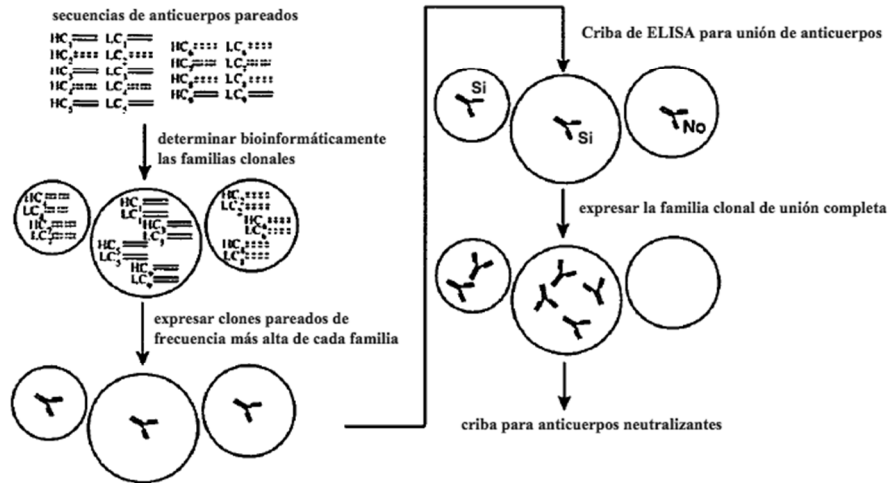
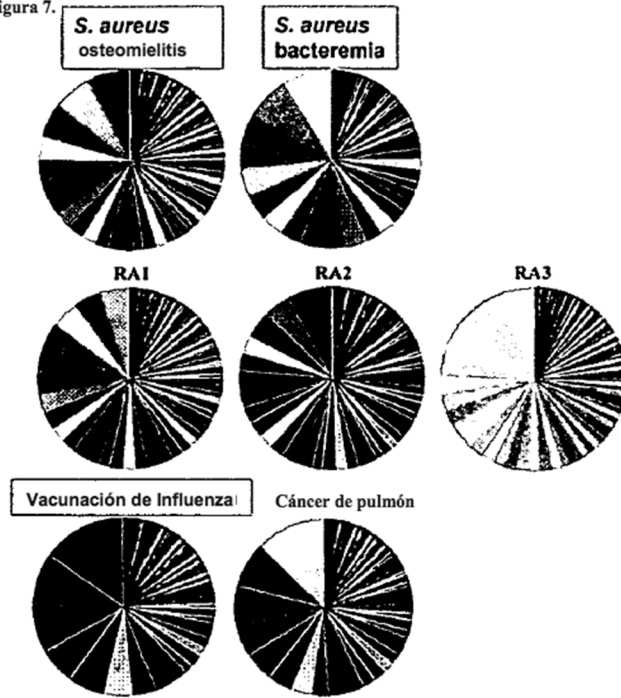


Figura 7.





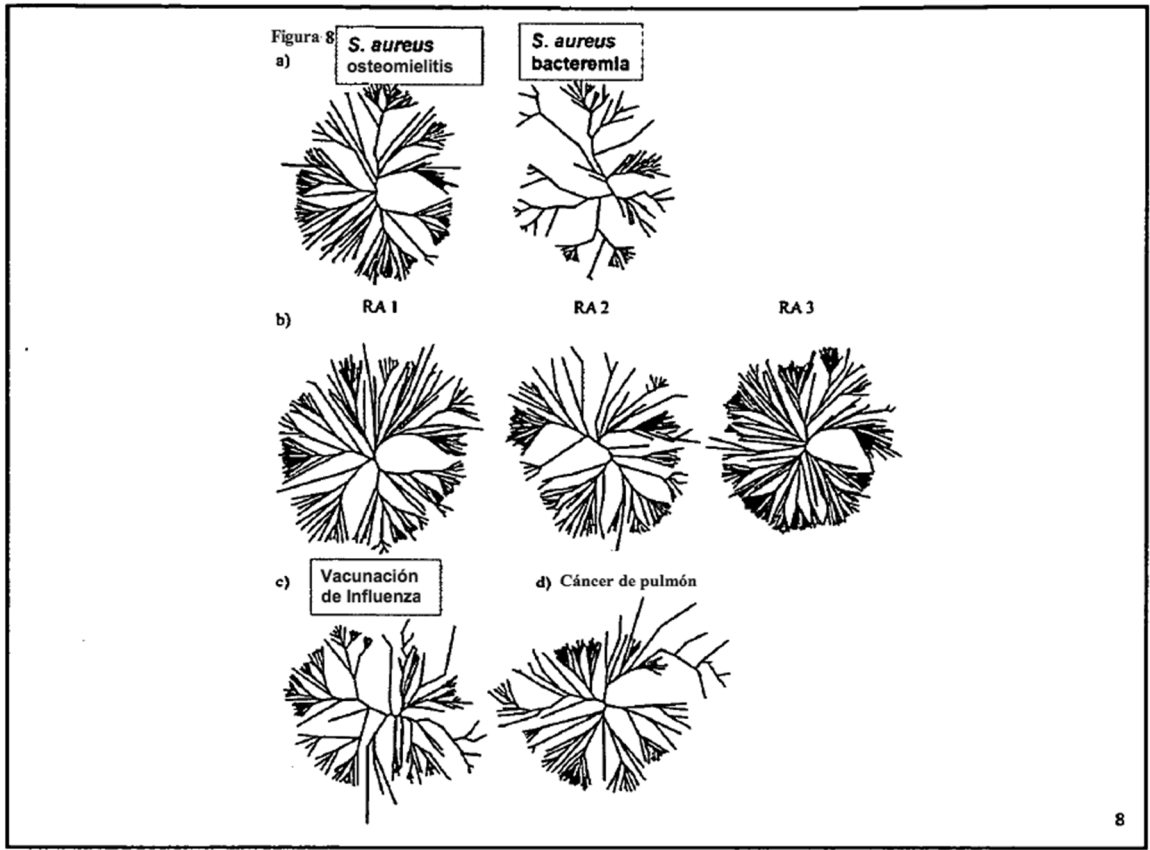


Figura 9

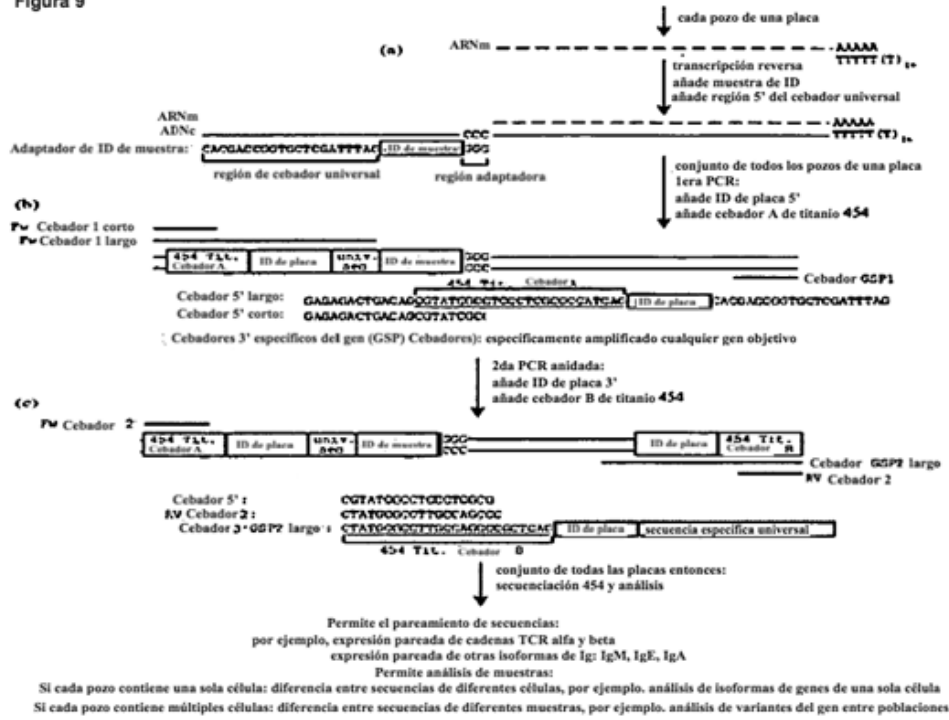


Figura 10

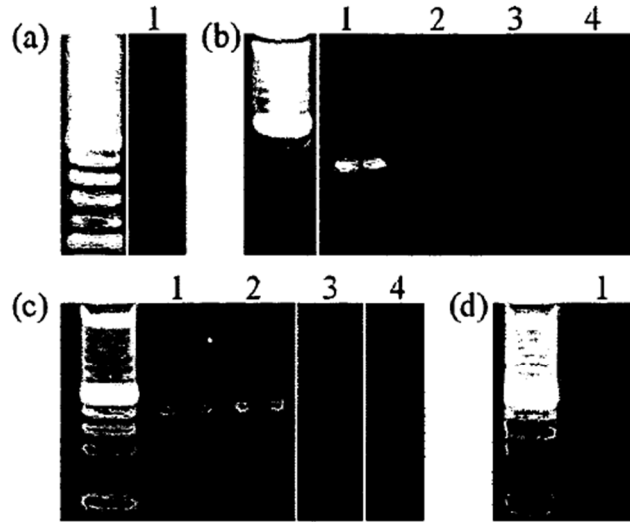
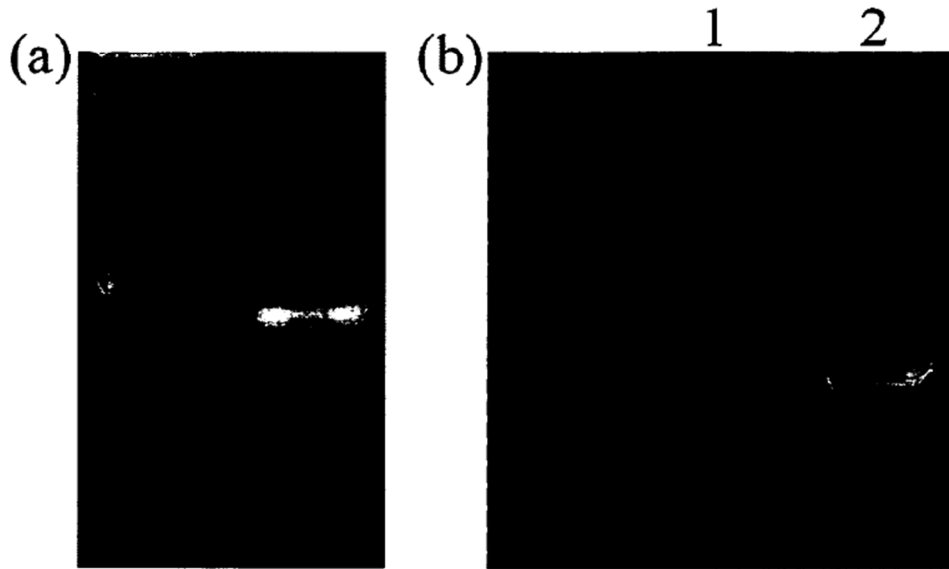


Figura 11



11

Figura 12

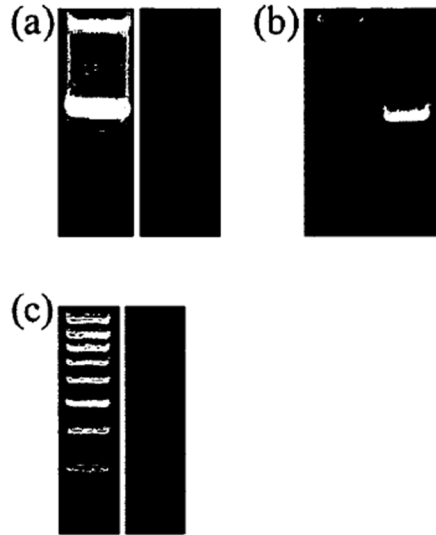


Figura 13

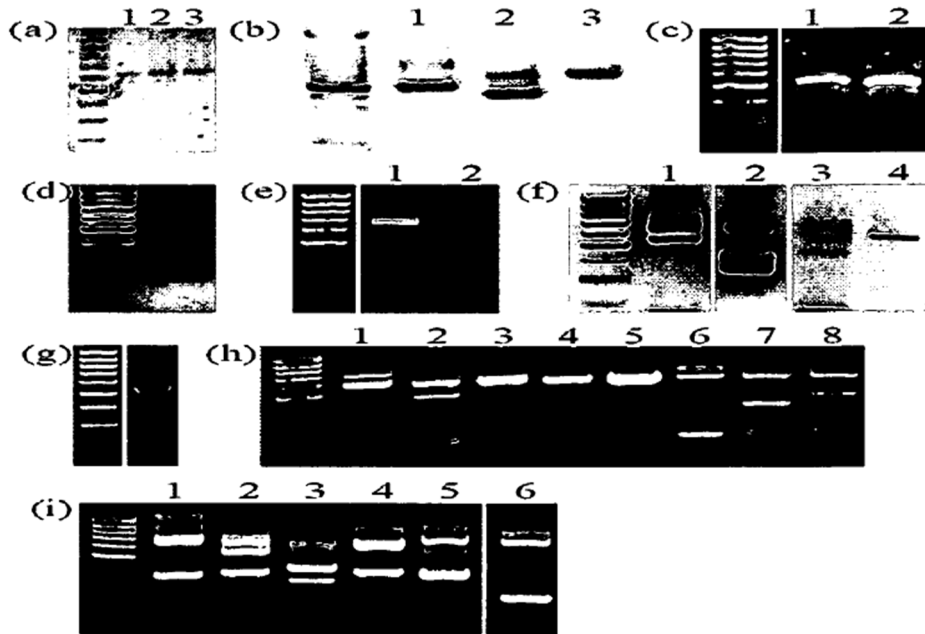


Figura 14

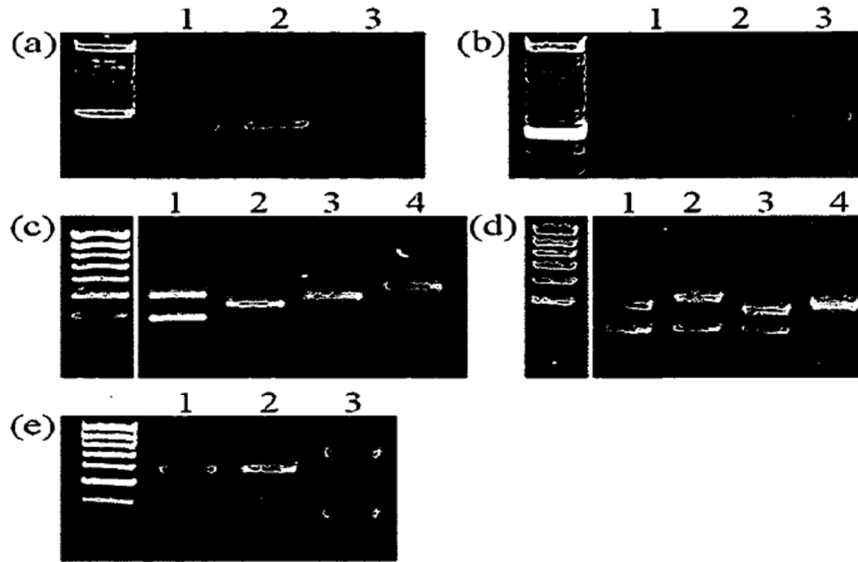
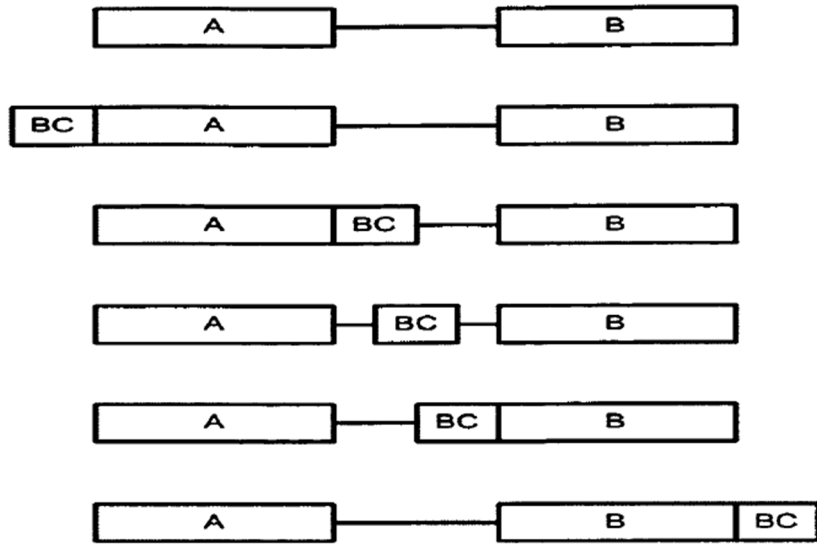


Figura 15





**Figura 16**

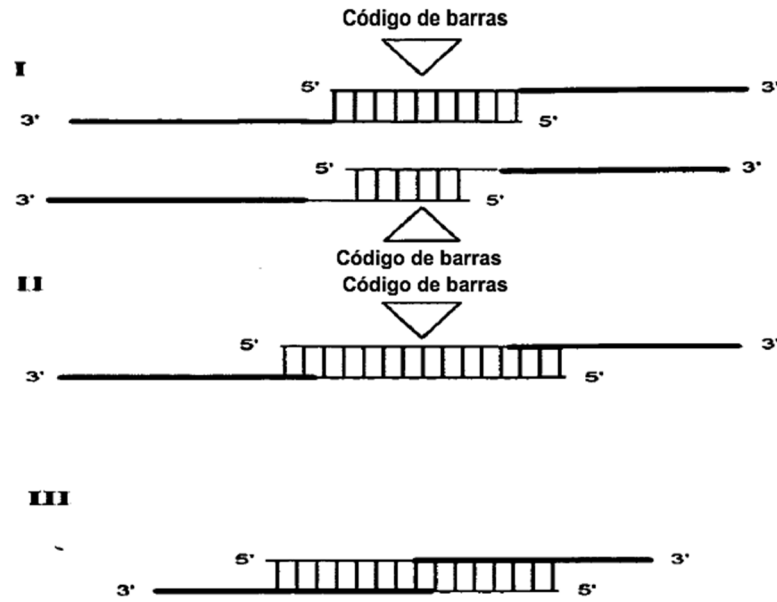


Figura 17

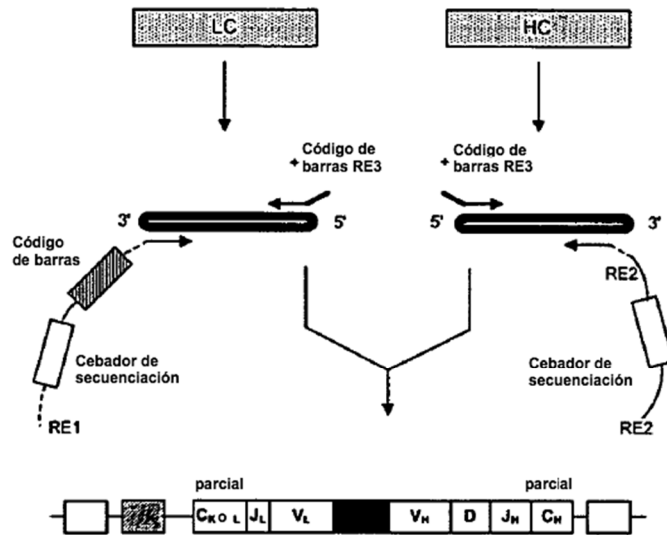
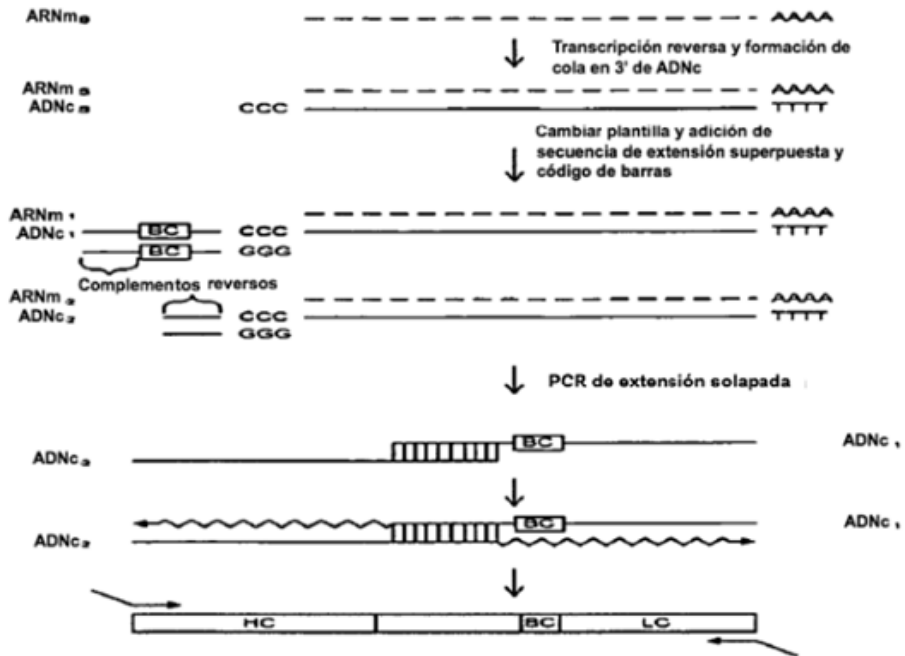
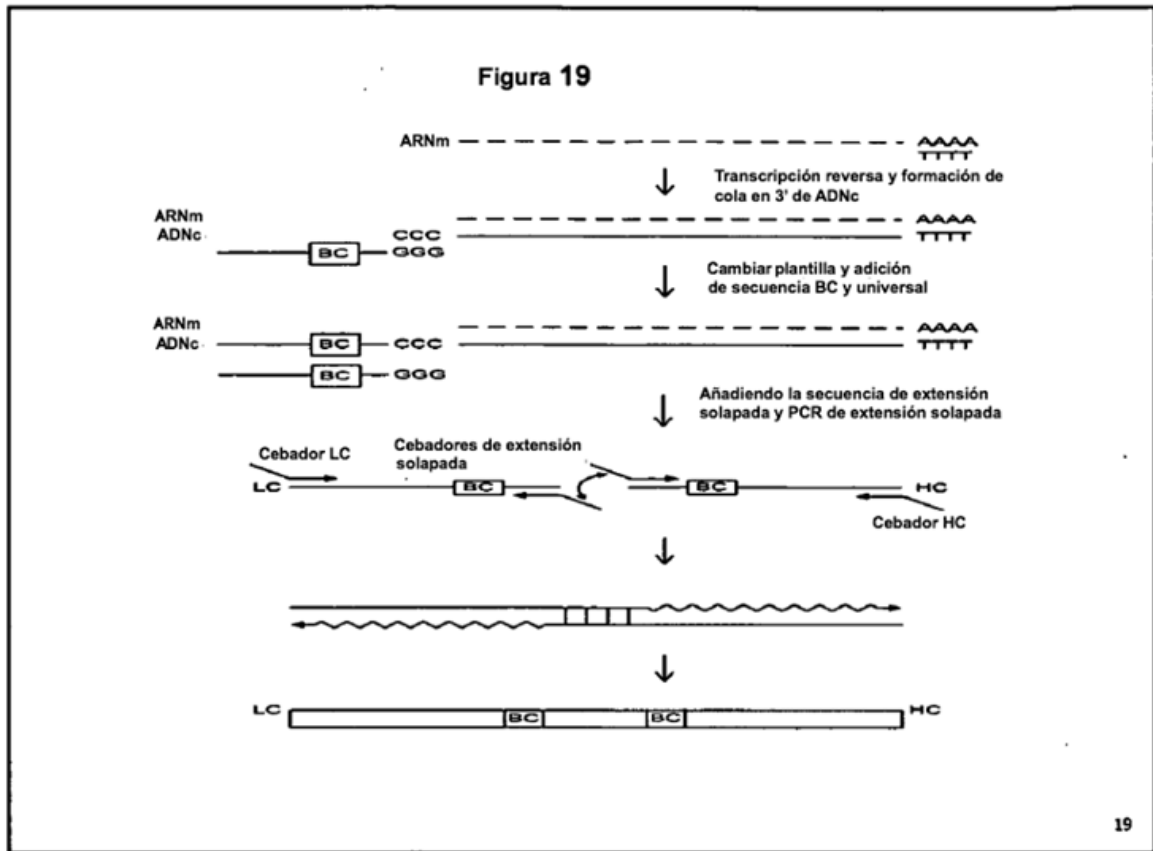
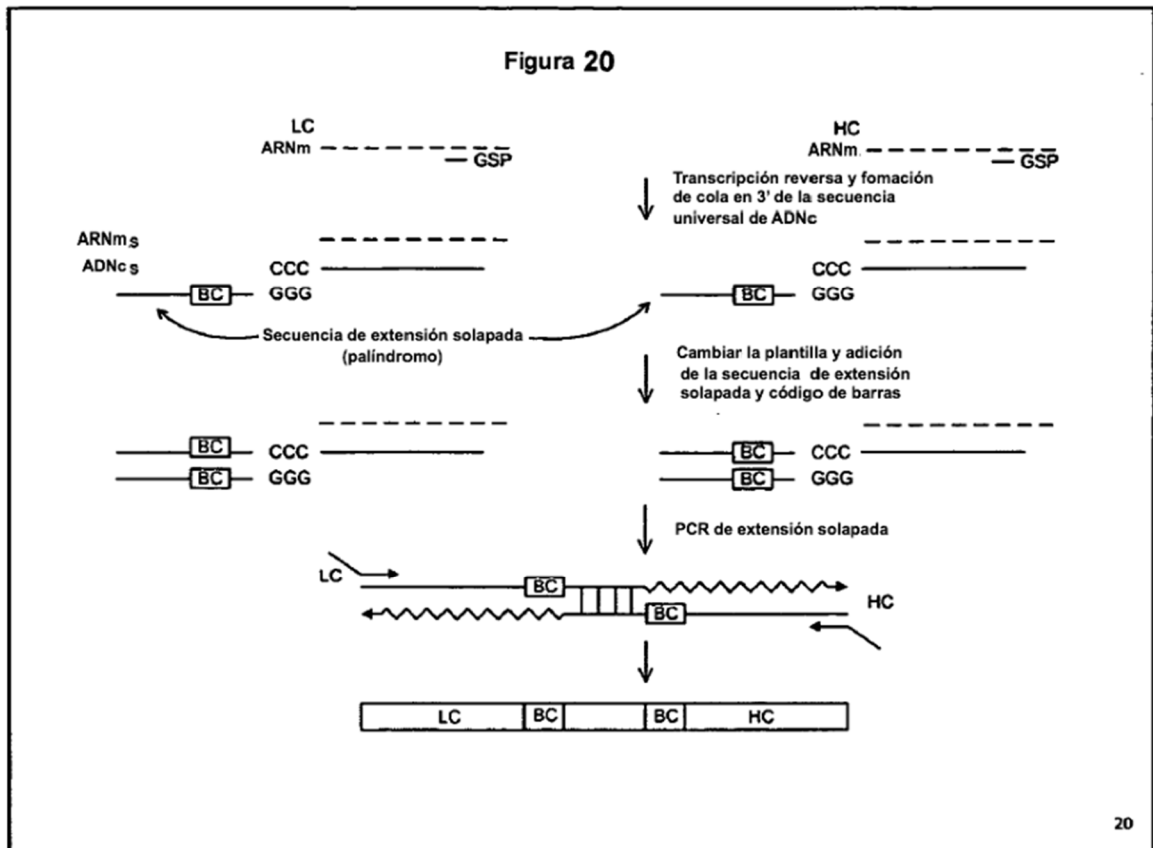


Figura 18







**Figura 21**

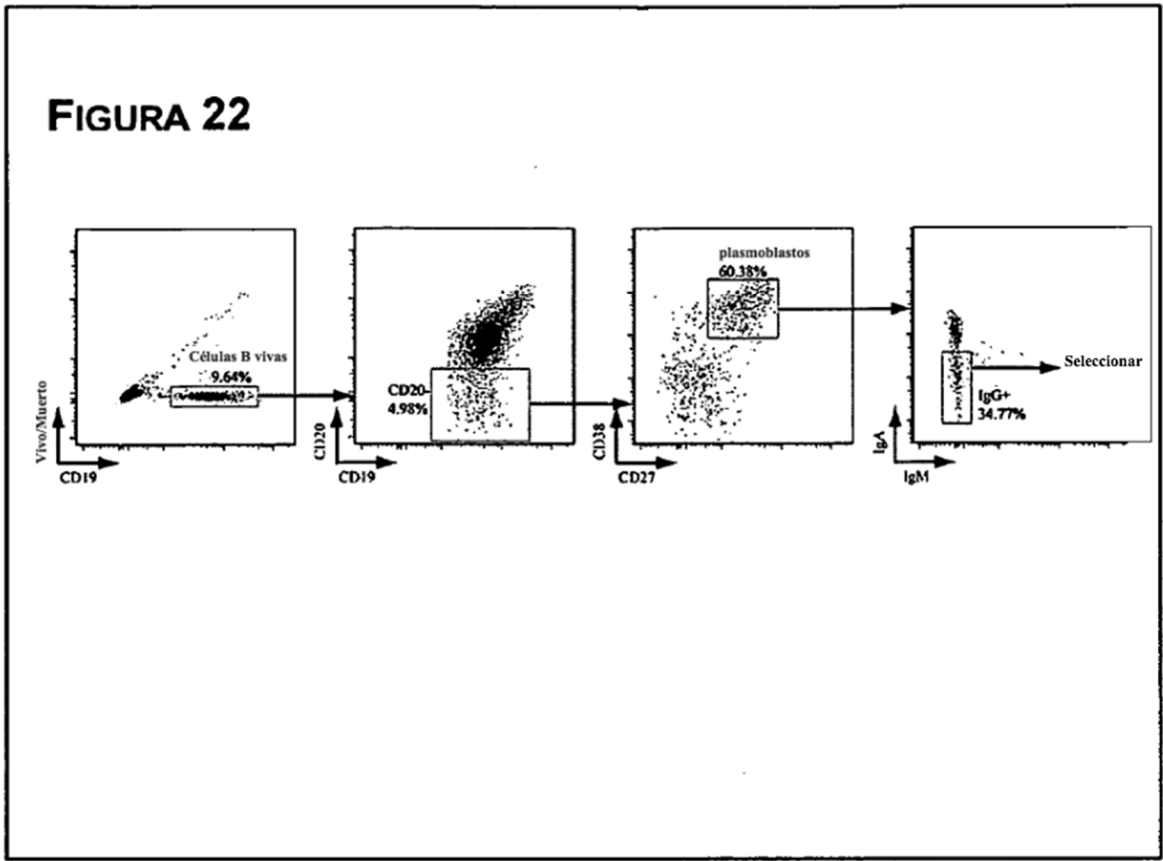
M 1 2 3 4



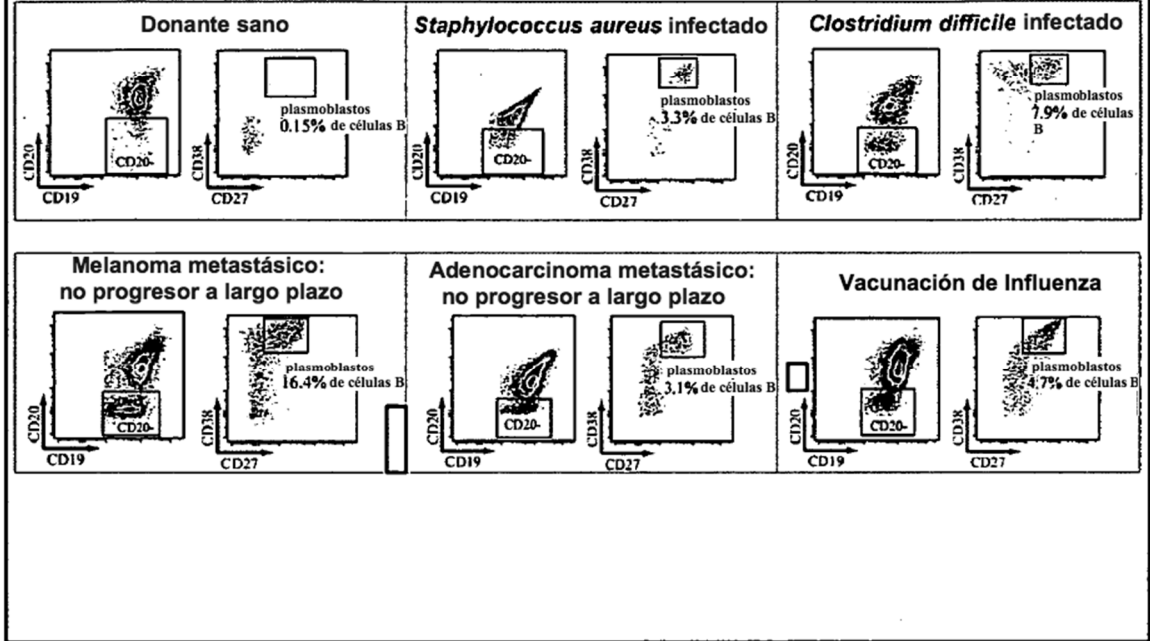
Secuencias:  
Univ\_sec\_2  
VR  
VL  
IgKC\_v3 código de barras  
IgLC\_v5 código de barras  
Fijado\_PCR3

AACGCGTGACGAGAGACTGACAG  
ATGAGGSTCCYGCTCAGCTGCTG G  
GGT CCTGGGCCAGT CTGCCCTG  
AGGCCCTTACGACTGCGTCTTGAACAATACCAGATGGCGGGAAGATGAAGAC  
AGGCCCTTACGACTGCGTCTTGAACAATACTCCOGGTAGAAATCAC  
AGGCCCTTACGACTGCGTCTTG

**FIGURA 22**



**FIGURA 23**





**FIGURA 24**

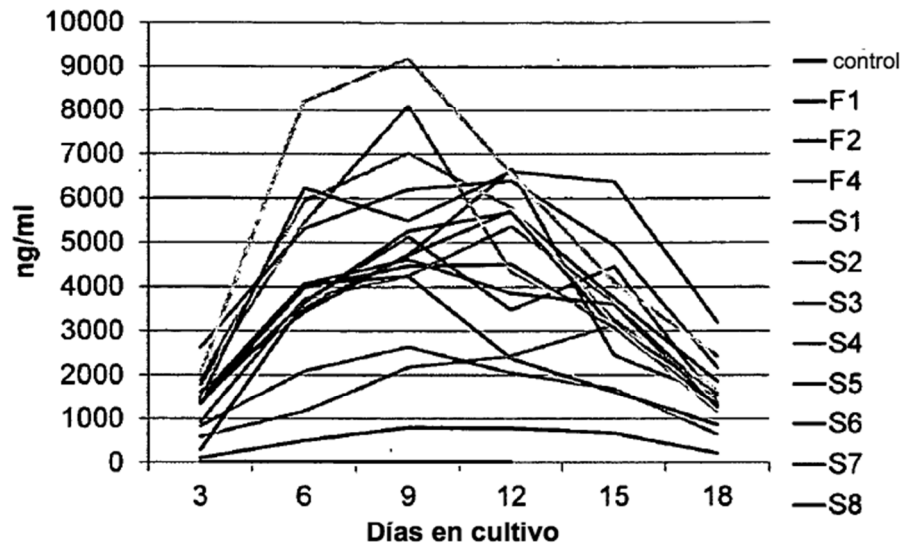
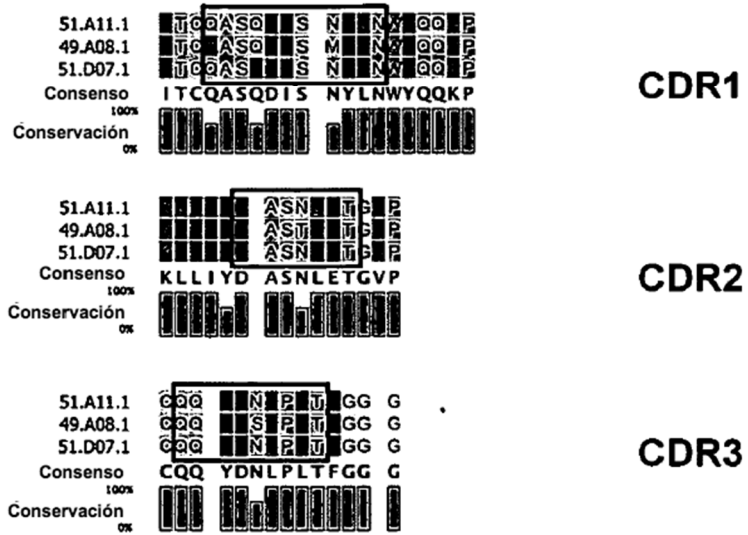


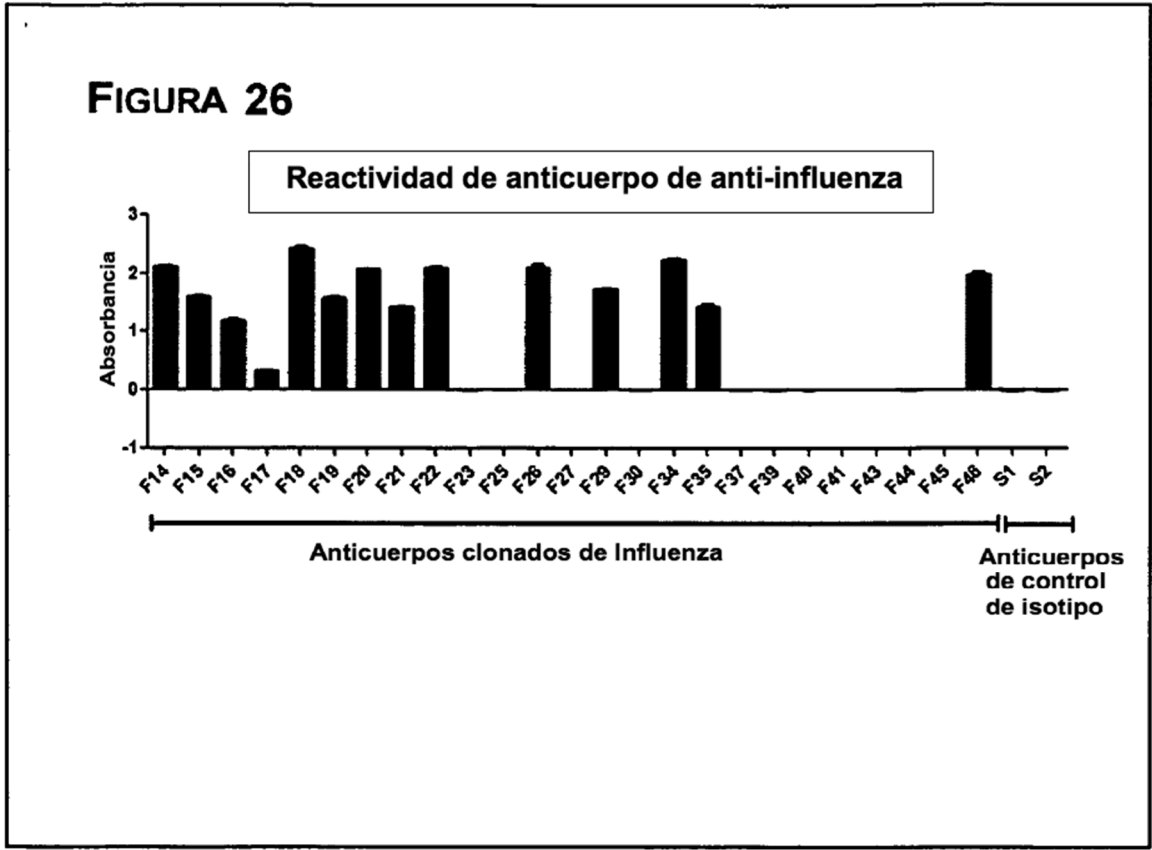
FIGURA 25A





**FIGURA 25C**





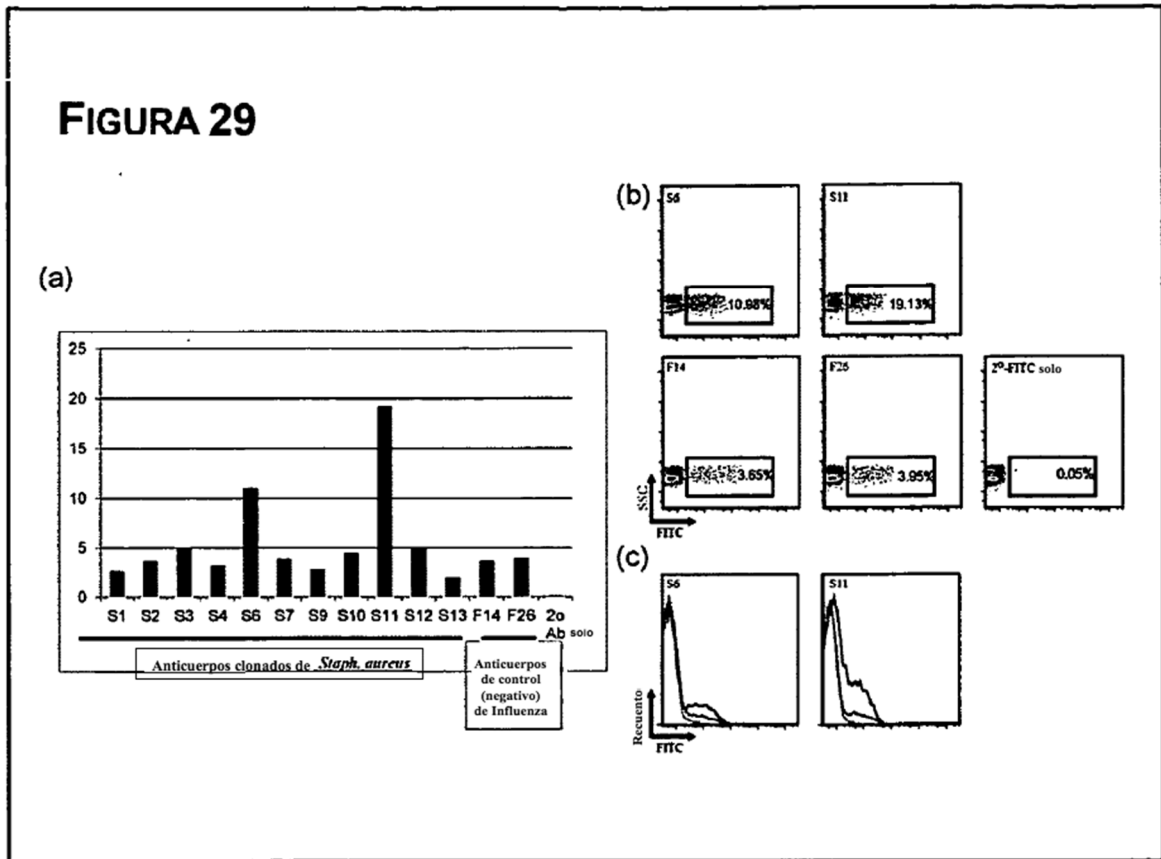
**FIGURA 27**

Anticuerpo	¿H3 Perth unido?	¿H1 California unido?	K <sub>a</sub>	K <sub>d</sub>	K <sub>n</sub>
FLU 14	Y	N	7.12E+05	1.13E-02	3.04E-08
FLU 16	Y	N	5.35E+05	5.86E-03	1.36E-08
FLU 18	Y	N	2.89E+05	1.60E-03	6.00E-09
FLU 19	Y	N	1.96E+05	5.08E-02	2.59E-07
FLU 20	Y	N	3.77E+05	6.28E-02	4.62E-07
FLU 22	Y	N	1.39E+05	6.75E-03	5.38E-08
FLU 26	Y	N	5.24E+05	1.52E-04	3.13E-10
FLU 34	Y	N	1.40E+05	1.26E-02	9.25E-08
FLU 35	Y	N	1.96E+04	2.35E-02	1.20E-06
FLU 46	Y	N	3.39E+04	4.13E-02	1.24E-06
FLU 21	N	Y	2.23E+05	5.29E-02	2.38E-07
FLU 15	N	N	--	--	--
FLU 17	N	N	--	--	--
FLU 29	N	N	--	--	--

**FIGURA 28**

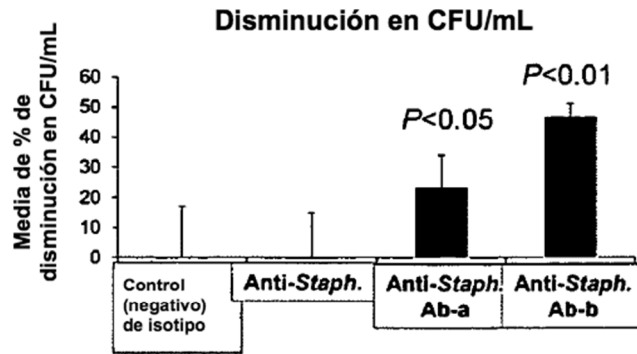
Anticuerpo probado	Neutralización de infectividad	
	H1N1- A/California/07/2009	H3N2 A/Perth/16/2009
mAb F15	No	No
mAb F16	No	Si. Neutraliza a $\leq 1.5 \mu\text{g/ml}$
mAb F18	No	Si. Neutraliza a $\leq 1.5 \mu\text{g/ml}$
mAb F19	No	Si. Neutraliza a $3.1 \mu\text{g/ml}$
mAb F20	No	Si. Neutraliza a $12.5 \mu\text{g/ml}$
mAb F21	Si. Neutraliza a $\geq 1.5 \mu\text{g/ml}$	Si. Neutraliza a $100 \mu\text{g/ml}$
Control policlonal	Si. Neutraliza a $\geq 1/128 \text{ diln}$	Si. Neutraliza a $\geq 1/128 \text{ diln}$

**FIGURA 29**





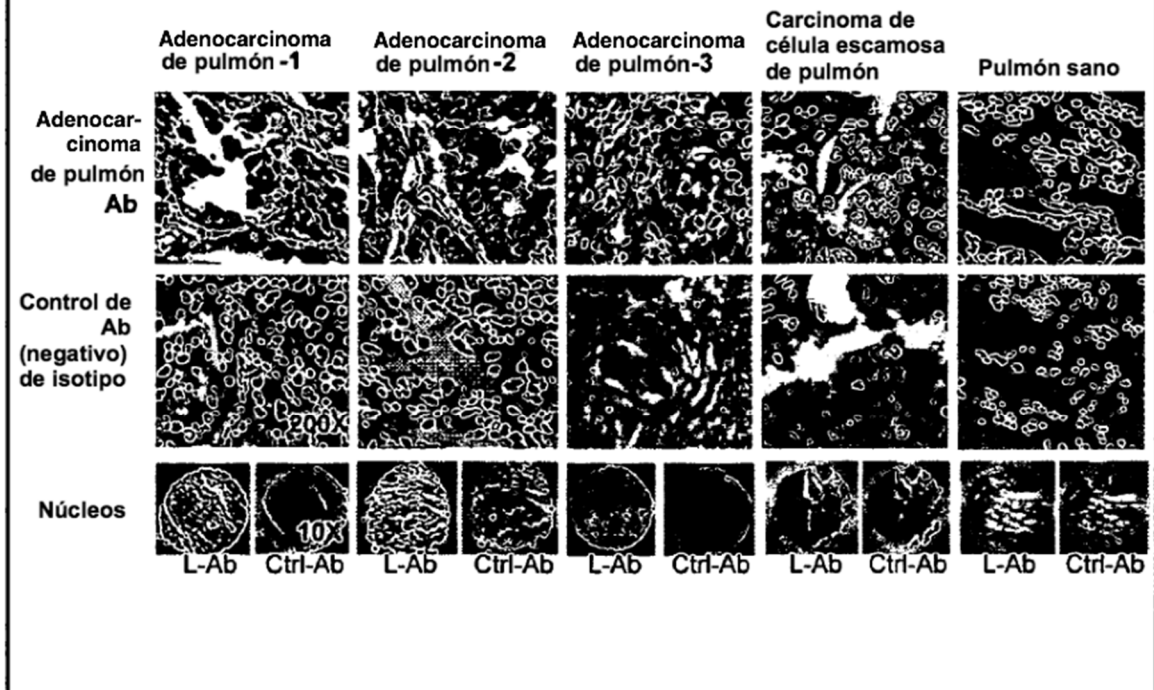
**FIGURA 30**



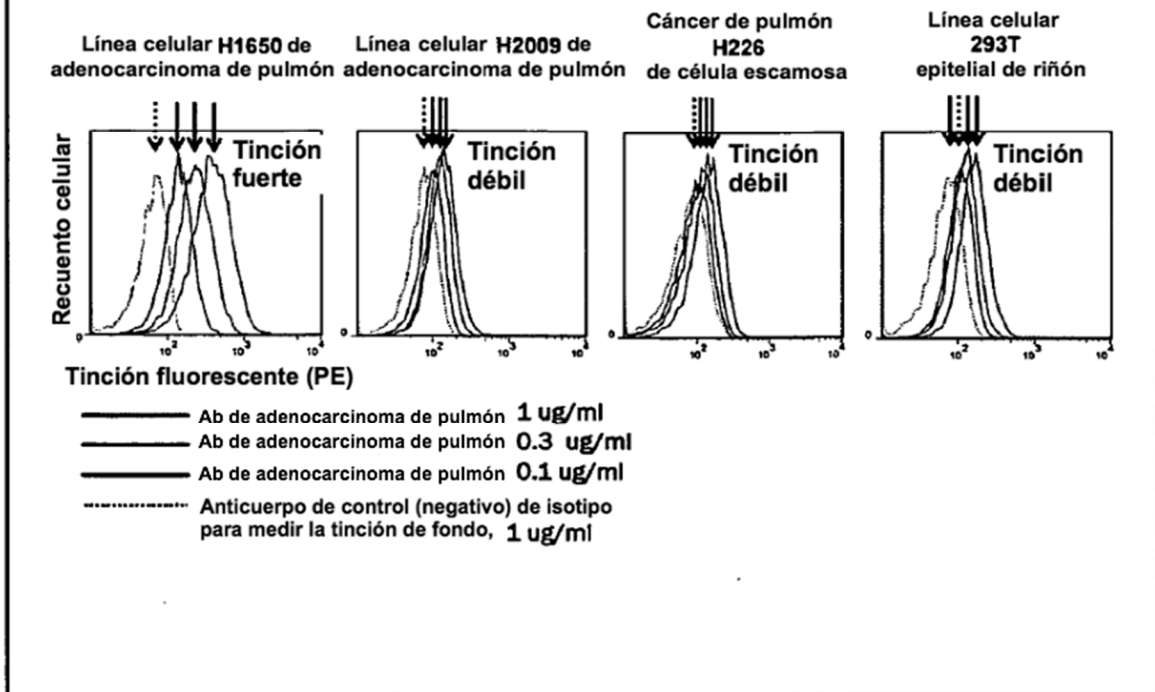
**FIGURA 31**

Anticuerpo	Adhesión de objetivo #	Nombre de proteína
S04	Q6G7S2	Péptido de modulina alfa 1 soluble en fenol
S04	P0C7Y8	Hemolisina delta

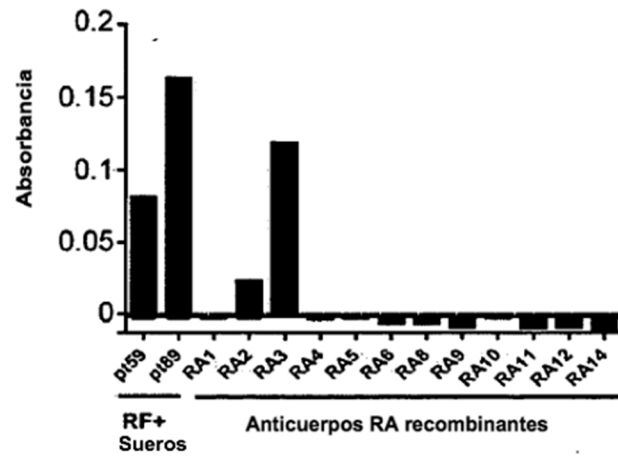
**FIGURA 32**



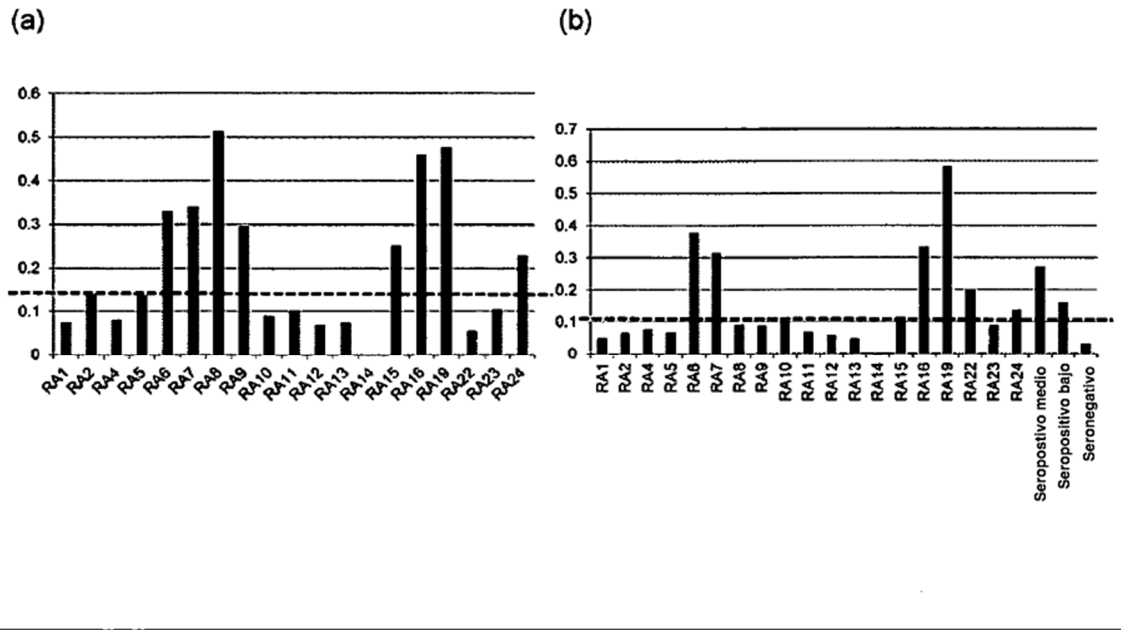
**FIGURA 33**



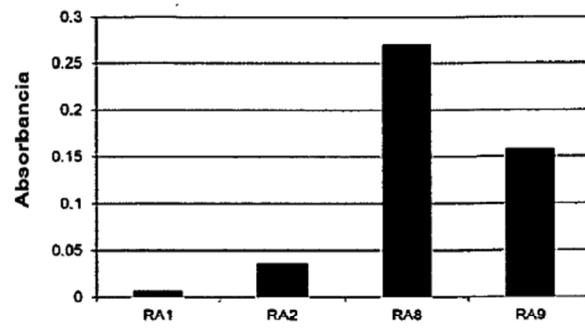
**FIGURA 34**



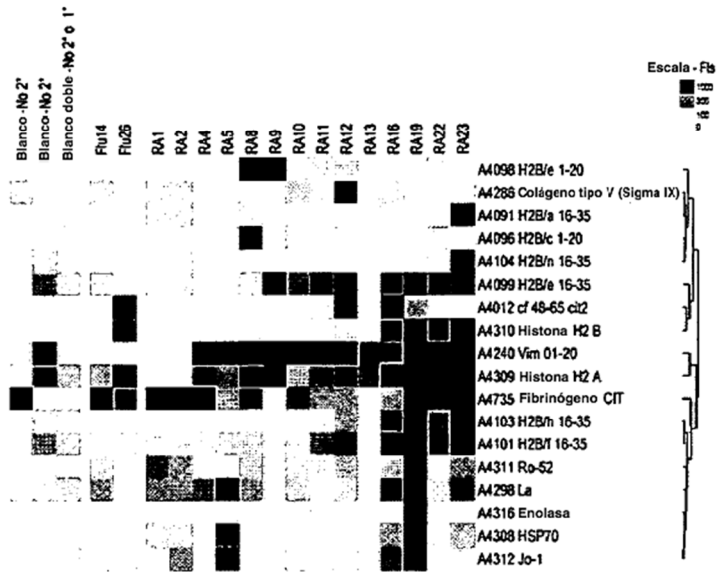
**FIGURA 35**



**FIGURA 36**



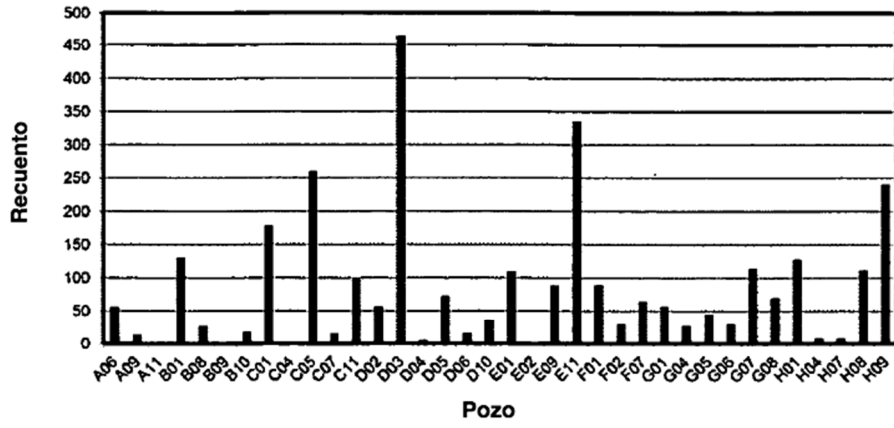
**FIGURA 37**



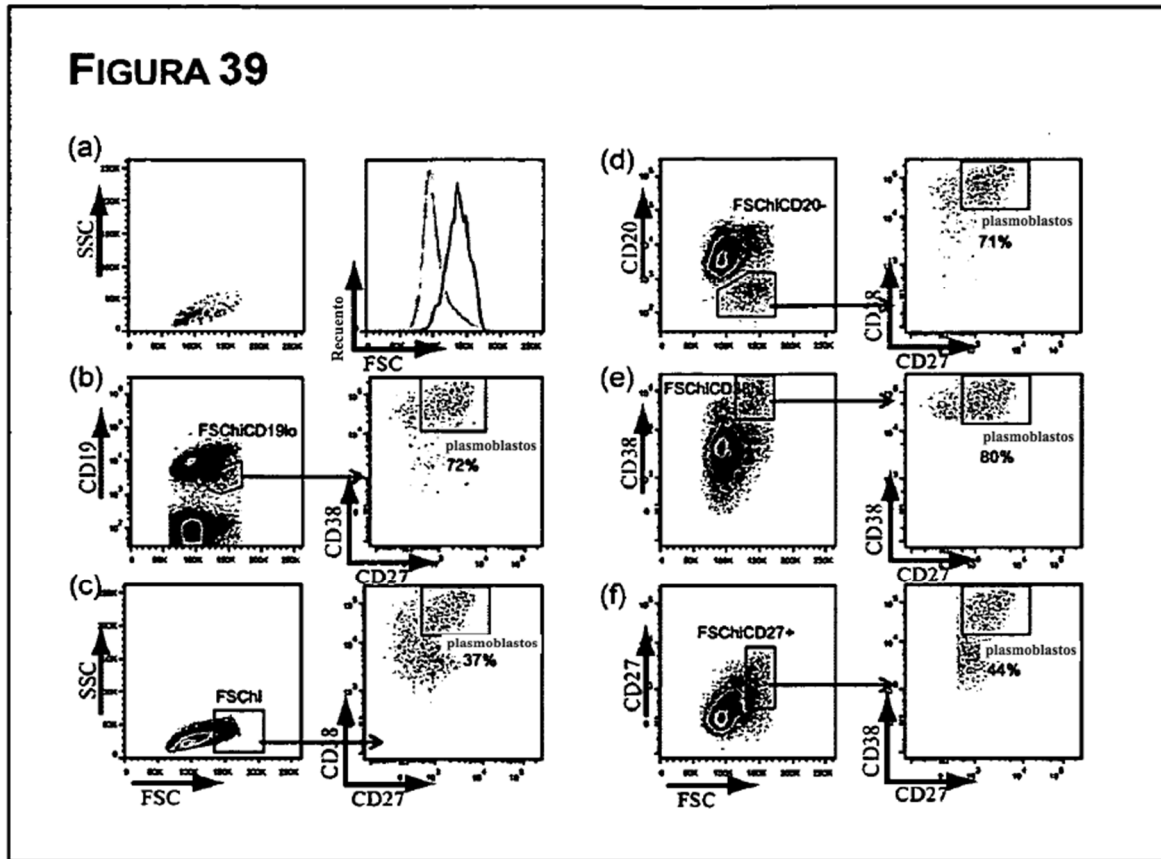


**FIGURA 38**

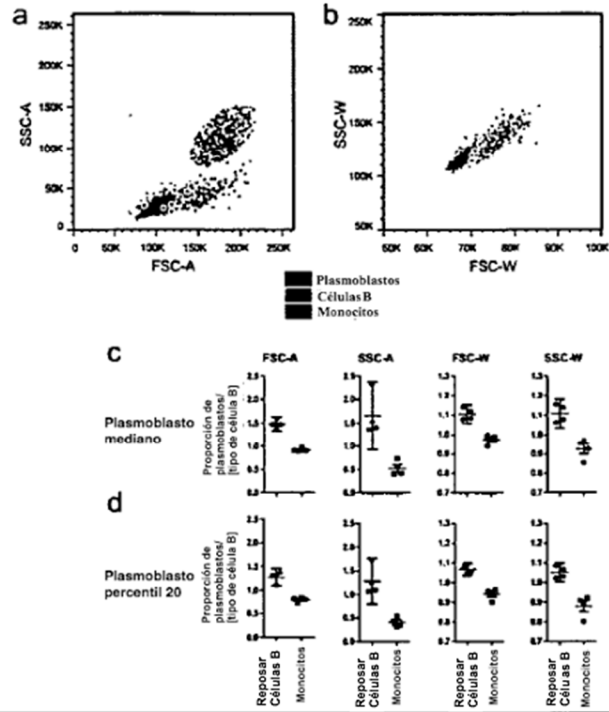
**Número de lecturas CCS**

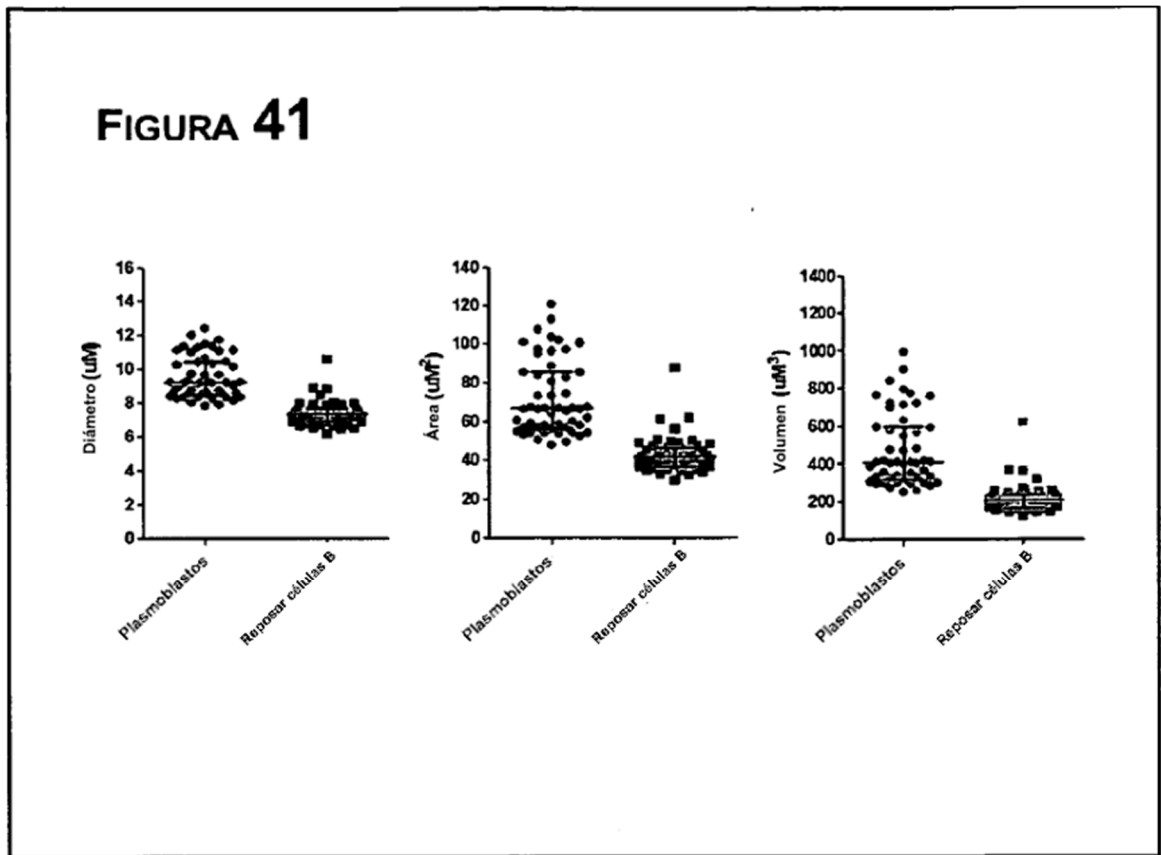


**FIGURA 39**

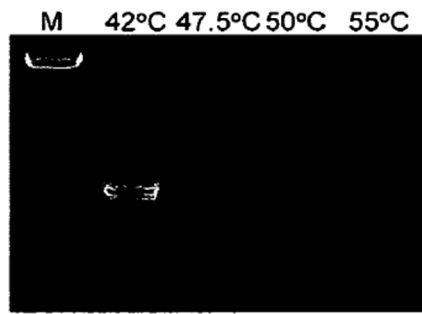


**FIGURA 40**





**FIGURA 42**

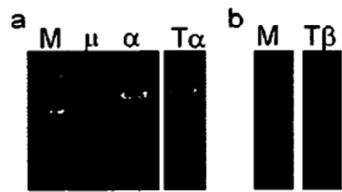


**FIGURA 43**

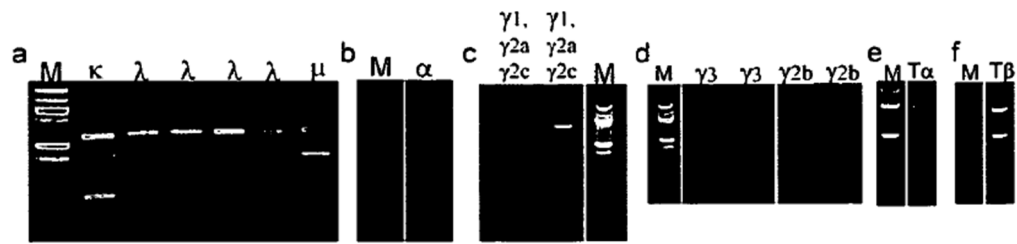
M K1 K1e K1f L1 L1x L1y



**FIGURA 44**



**FIGURA 45**





**FIGURA 46**

