

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 619 713**

51 Int. Cl.:

**C12Q 1/68** (2006.01)

**C12N 15/12** (2006.01)

**C07K 16/18** (2006.01)

**A61K 39/395** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.10.2012 PCT/US2012/061977**

87 Fecha y número de publicación internacional: **10.05.2013 WO2013066726**

96 Fecha de presentación y número de la solicitud europea: **25.10.2012 E 12844825 (5)**

97 Fecha y número de publicación de la concesión europea: **11.01.2017 EP 2773773**

54 Título: **Clonotipos de receptores de células T compartidos entre pacientes con espondilitis anquilosante**

30 Prioridad:

**04.11.2011 US 201161556125 P**  
**17.11.2011 US 201161561234 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**26.06.2017**

73 Titular/es:

**ADAPTIVE BIOTECHNOLOGIES CORPORATION**  
**(100.0%)**  
**400 East Jamie Court, Suite 301**  
**South San Francisco, CA 94080, US**

72 Inventor/es:

**FAHAM, MALEK;**  
**CARLTON, VICTORIA;**  
**MOORHEAD, MARTIN;**  
**ZHENG, JIANBIAO y**  
**ASBURY, THOMAS**

74 Agente/Representante:

**UNGRÍA LÓPEZ, Javier**

**ES 2 619 713 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

Clonotipos de receptores de células T compartidos entre pacientes con espondilitis anquilosante

5 **Antecedentes de la invención**

10 La espondilitis anquilosante (EA, del griego spondylos, vértebras; ankylos, doblado), anteriormente conocida como enfermedad de Bechterew, síndrome de Bechterew y enfermedad de Marie Strümpell, una forma de espondiloartritis, es una artritis inflamatoria crónica y una enfermedad autoinmune. Afecta principalmente a las articulaciones de la columna y el sacro-ilion en la pelvis, causando una fusión eventual de la columna. Es un miembro del grupo de las espondiloartropatías con una fuerte predisposición genética. Una fusión completa da como resultado una rigidez completa de la columna, una afección que se conoce como columna en caña de bambú.

15 El paciente típico es un varón joven, de 18-30 años de edad, cuando aparecen por vez primera los síntomas de la enfermedad, con dolor crónico y rigidez en la parte baja de la columna o a veces en toda la columna, con frecuencia con dolor referido a una u otra nalga o a la parte posterior del muslo desde la articulación sacroilíaca. Los hombres resultan afectados más que las mujeres, en una proporción de aproximadamente 3:1, teniendo la enfermedad normalmente un curso más doloroso en hombres que en mujeres. En un 40% de los casos, la espondilitis anquilosante se asocia a una inflamación del ojo (iridociclitis y uveítis), que causa enrojecimiento, dolor ocular, pérdida de visión, motas y fotofobia. Otro síntoma común es la fatiga generalizada y a veces náuseas. Menos comúnmente, pueden aparecer aortitis, fibrosis pulmonar apical y ectasia de las vainas de las raíces de los nervios sacros. Al igual que con todas las espondiloartropatías seronegativas, puede producirse levantamiento de las uñas (onicolisis).

25 No existe ninguna prueba directa para diagnosticar la EA. Un examen clínico y estudios de rayos X de la columna, que muestran los cambios característicos en la columna y sacroilitis, son las herramientas diagnósticas principales. Un inconveniente del diagnóstico por rayos X es que los signos y síntomas de la EA han estado normalmente establecidos hasta 8-10 años antes de que aparezcan cambios evidentes por rayos X en rayos X de película simple, lo que significa un retraso de hasta 10 años antes de que se puedan introducir terapias adecuadas. Las opciones para un diagnóstico más precoz son la tomografía y la imagen de resonancia magnética de las articulaciones sacroilíacas, pero la fiabilidad de estas pruebas no está aún clara. La prueba de Schober es una medida clínica útil de flexión de la columna lumbar realizada durante el examen.

35 Durante los períodos inflamatorios agudos, los pacientes de EA mostrarán a veces un aumento en la concentración sanguínea de proteína C reactiva (CRP) y un aumento en la velocidad de sedimentación de los eritrocitos (VSG), pero hay muchos pacientes con EA cuyos índices de CRP y VSE no aumentan, por lo que resultados normales de CRP y VSE no siempre se corresponden con la cantidad de inflamación que tiene realmente una persona. A veces, las personas con EA tienen resultados con niveles normales, a pesar de lo cual experimentan una cantidad significativa de inflamación en sus cuerpos.

40 Existen tres tipos principales de medicaciones empleadas para tratar la espondilitis anquilosante: 1) Fármacos antiinflamatorios, que incluyen fármacos AINE, tales como ibuprofeno, fenilbutazona, indometacina, naproxeno e inhibidores de COX-2, que reducen la inflamación y el dolor. También se ha visto gracias a la evidencia clínica que los analgésicos opioides son muy efectivos en el alivio del tipo de dolor crónico comúnmente experimentado por quienes sufren de EA, especialmente en formulaciones de liberación a lo largo del tiempo; 2) FARME, tales como ciclosporina, metotrexato, sulfasalazina y corticosteroides, usados para reducir la respuesta del sistema inmunitario por inmunosupresión; 3) los bloqueantes del TNF $\alpha$  (antagonistas), tales como etanercept, infliximab y adalimumab (también conocidos como biológicos), están indicados para el tratamiento de, y son inmunosupresores eficaces en, esta afección, al igual que en otras enfermedades autoinmunes.

50 Los bloqueantes del TNF $\alpha$  han mostrado ser el tratamiento más prometedor, ralentizando el progreso de la EA en la mayoría de los casos clínicos, ayudando a muchos pacientes a que reciban una reducción significativa, aunque no una eliminación, de la inflamación y el dolor que padecen. También han mostrado ser altamente eficaces en el tratamiento no sólo de la artritis de las articulaciones, sino también de la artritis de la columna asociada a la EA. Un inconveniente, además del frecuentemente elevado coste, es el hecho de que estos fármacos aumentan el riesgo de infecciones. Por esta razón, el protocolo para cualquiera de los bloqueantes del TNF- $\alpha$  incluye una prueba para tuberculosis (como Mantoux o Heaf) antes de iniciar el tratamiento. En caso de infecciones recurrentes, incluso de irritación de garganta recurrente, se puede suspender la terapia por la inmunosupresión que ello implica. Se aconseja a los pacientes que toman las medicaciones TNF que limiten su exposición a otras personas que sean, o puedan ser, portadoras de un virus (tal como el resfriado o la gripe) o que puedan tener una infección bacteriana o fúngica.

65 La EA produce síntomas que son muy comunes en las poblaciones sanas. Por ejemplo, un paciente que se presente quejándose de un dolor severo de espalda puede no estar experimentando un brote de EA, sino que más bien podría tener sólo un dolor de espalda rutinario. El médico se ve forzado a tomar una decisión sobre si tratar estos síntomas con fármacos caros que tienen efectos colaterales potencialmente severos sin una visión muy precisa del

estado de la enfermedad. La CRP y la VSE no proporcionan una visión muy precisa del estado de la enfermedad. Al mismo tiempo, el curso de la enfermedad no tratada puede dar como resultado un daño espinal debilitante a largo plazo. Esta situación conduce a un reto clínico difícil y se utiliza un sobretratamiento significativo. La disponibilidad de una medida objetiva que refleje la actividad de la enfermedad puede resultar de gran ayuda en el tratamiento de pacientes con EA.

Los perfiles de ácidos nucleicos codificantes de moléculas inmunes, tales como los receptores de las células T o de las células B, o sus componentes, contienen una gran cantidad de información sobre el estado de salud o de enfermedad de un organismo, por lo que el uso de dichos perfiles como indicadores diagnósticos o pronósticos ha sido propuesto para una amplia variedad de afecciones, incluyendo afecciones autoinmunes, *v.g.*, Faham y Willis, publicación de patente estadounidense 2010/0151471 y 2011/0207134; Freeman *et al.*, *Genome Research*, 19: 1817-1824 (2009); Boyd *et al.*, *Sci. Transl. Med.*, 1(12): 12ra23 (2009); He *et al.*, *Oncotarget* (8 de marzo de 2011). Dichos perfiles basados en secuencia son capaces de una sensibilidad mucho mayor que los enfoques basados en las distribuciones de tamaño de regiones codificantes de CDR amplificadas, el muestreo de secuencias por micromatrices, las curvas de cinética de hibridación a partir de amplicones de PCR u otros enfoques, *v.g.*, Morley *et al.*, patente estadounidense 5.418.134; van Dongen *et al.*, *Leukemia*, 17: 2257-2317 (2003); Ogle *et al.*, *Nucleic Acids Research*, 31: e139 (2003); Wang *et al.*, *BMC Genomics*, 8: 329 (2007); Baum *et al.*, *Nature Methods*, 3(11): 895-901 (2006).

Considerando el impacto personal y social de la EA, sería muy deseable poder disponer de medidas de actividad de la enfermedad basadas en perfiles de secuencias inmunes que pudieran correlacionarse fácilmente con los estados de salud o de enfermedad y/o la probabilidad del éxito en el tratamiento.

### Resumen de la invención

La presente invención se dirige a métodos para determinar el estado de la enfermedad de pacientes con espondilitis anquilosante mediante el análisis de los perfiles de clonotipos basados en secuencia de las cadenas  $\beta$  de los receptores de las células T de los pacientes. La invención queda ejemplificada en una serie de implementaciones y aplicaciones, algunas de las cuales son resumidas a continuación y en toda la memoria descriptiva.

En un aspecto, la invención incluye un método para determinar un estado de la enfermedad de un paciente que sufre, o que se sospecha que sufre, espondilitis anquilosante, consistente en las etapas de (i) determinar en un perfil de clonotipos de una muestra de tejido del paciente la presencia, ausencia y/o cantidad de clonotipos codificantes de segmentos de un receptor de las células T al menos un setenta por ciento homólogos a un segmento en el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) (péptido 1) y VYFCASSDSSGSDTDQYFGPGTRLTV (SEC. N° ID. 2) (péptido 2); y (ii) correlacionar la presencia, ausencia y/o cantidad de dichos clonotipos con un estado de espondilitis anquilosante en el paciente. En algunas realizaciones, dichos métodos consisten en las etapas de (i) determinar en un perfil de clonotipos de una muestra de tejido del paciente la presencia y/o cantidad de clonotipos codificantes de segmentos de un receptor de las células T al menos un noventa por ciento homólogos a un segmento en el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) (péptido 1) y VYFCASSDSSGSDTDQYFGPGTRLTV (SEC. N° ID. 2) (péptido 2); (ii) correlacionar la presencia y/o cantidad de dichos clonotipos con un estado de espondilitis anquilosante en el paciente; y (iii) tratar al paciente con una medicación para mejorar los efectos de la espondilitis anquilosante.

En otro aspecto, la invención incluye un método de tratamiento de un paciente con espondilitis anquilosante mediante la administración de una cantidad efectiva de un anticuerpo específico para un segmento de aminoácidos de un receptor de las células T, siendo seleccionado el segmento de aminoácidos entre el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y cualquier segmento de 6 a 20 aminoácidos del mismo y VYFCASSDSSGSDTDQYFGPGTRLTV (SEC. N° ID. 2) y cualquier segmento de 6 a 20 aminoácidos del mismo.

Estos aspectos anteriormente caracterizados, así como otros aspectos, de la presente invención quedan ejemplificados en una serie de implementaciones y aplicaciones ilustradas, algunas de las cuales se muestran en las figuras y se caracterizan en la sección de reivindicaciones que se dará más adelante. Sin embargo, el resumen anterior no pretende describir cada realización ilustrada o cada implementación de la invención.

### Breve descripción de los dibujos

Las nuevas características de la invención son expuestas con detalle en las reivindicaciones adjuntas. Se obtiene un mejor entendimiento de las características y ventajas de la presente invención haciendo referencia a la siguiente descripción detallada que expone realizaciones ilustrativas, en donde se utilizan los principios de la invención, y a los dibujos adjuntos, de los cuales:

Las Figs. 1A-1C muestran un esquema de PCR en dos etapas para la amplificación de genes TCR $\beta$ .

**Descripción detallada de la invención**

La práctica de la presente invención puede emplear, a menos que se indique lo contrario, técnicas convencionales y descripciones de biología molecular (incluyendo técnicas recombinantes), bioinformática, biología celular y bioquímica, que están dentro de los conocimientos de la técnica. Dichas técnicas convencionales incluyen, aunque sin limitación, la toma de muestras y el análisis de las células sanguíneas, la secuenciación y el análisis de ácidos nucleicos, la construcción y la aplicación de inmunoensayos y similares. Se puede disponer de ilustraciones específicas de técnicas adecuadas haciendo referencia al ejemplo que se describe aquí más adelante. Sin embargo, también se pueden utilizar, por supuesto, otros procedimientos convencionales equivalentes.

La invención se dirige a métodos para determinar el estado de la enfermedad de pacientes que sufren o pueden estar sufriendo espondilitis anquilosante. En un aspecto, se realiza dicha determinación detectando la presencia o ausencia o cantidad de clonotipos de los receptores beta de las células T (TCR $\beta$ ) que codifican segmentos TCR $\beta$  al menos un setenta por ciento homólogos a cualquiera de los segmentos del grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2). En otra realización, dicha detección es para clonotipos codificantes de segmentos TCR $\beta$  al menos un ochenta por ciento homólogos a cualquiera de los segmentos del grupo anterior. En otra realización, dicha detección es para clonotipos codificantes de segmentos TCR $\beta$  al menos un noventa por ciento homólogos a cualquiera de los segmentos del grupo anterior. En otra realización, dicha detección es para clonotipos codificantes de segmentos TCR $\beta$  idénticos a cualquiera de los segmentos del grupo anterior. Tal como se usa aquí, el término "péptidos relacionados con la EA" significa los péptidos LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2). En una realización de la invención, dichos clonotipos son estudiados generando un perfil de clonotipos basados en secuencia a partir de una muestra de tejido de un paciente, por ejemplo, usando el procedimiento desvelado por Faham y Willis, publicación de patente estadounidense 2011/020713 4. Resumiendo, en un aspecto, se obtiene un perfil de clonotipos basados en secuencia de un individuo y se implementa el método de la invención usando las siguientes etapas: (a) obtención de una muestra de ácido nucleico de las células T del individuo; (b) aislar espacialmente las moléculas individuales derivadas de dicha muestra de ácido nucleico; (c) secuenciar dichas moléculas individuales espacialmente aisladas; (d) determinar las abundancias de diferentes secuencias de las moléculas de ácido nucleico de la muestra de ácido nucleico para generar el perfil de clonotipos; y (e) determinar la presencia, ausencia y/o cantidad de clonotipos codificantes de segmentos de un receptor de las células T al menos un setenta por ciento homólogos a un segmento del grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2). En algunas realizaciones, la etapa de determinación incluye la determinación de la presencia, ausencia y/o cantidad de clonotipos codificantes de segmentos de un receptor de las células T al menos un ochenta por ciento o un noventa por ciento homólogos a los segmentos anteriores, o idénticos a los segmentos anteriores. En aún otras realizaciones, el método puede ser implementado mediante las siguientes etapas: (a) obtención de una muestra de un paciente que tiene células T; (b) amplificación de las moléculas de ácido nucleico de las células T de la muestra, teniendo las moléculas de ácido nucleico secuencias de ADN recombinadas de genes de los receptores de las células T; (c) secuenciación de las moléculas amplificadas de ácido nucleico para formar un perfil de clonotipos; y (d) determinar la presencia, ausencia y/o cantidad de clonotipos codificantes de segmentos de un receptor de las células T al menos un setenta por ciento homólogos a un segmento del grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2). Tal como antes, otras realizaciones pueden requerir la determinación de segmentos con diferentes homologías a las secuencias anteriores. En algunas realizaciones, los perfiles de clonotipos incluyen cada clonotipo presente a una frecuencia del 0,01 por ciento o mayor con una probabilidad del noventa y nueve por ciento. En otras realizaciones, los perfiles de clonotipos incluyen al menos  $10^4$  clonotipos, o al menos  $10^5$  clonotipos.

En otra realización, la etapa de secuenciación comprende la secuenciación bidireccional de cada una de las moléculas individuales espacialmente aisladas para producir al menos una lectura de secuencia hacia delante y al menos una lectura de secuencia inversa. Además, con respecto a esta última realización, al menos una de las lecturas de secuencia hacia delante y al menos una de las lecturas de secuencia inversas tienen una región de solapamiento, de tal forma que las bases de dicha región de solapamiento son determinadas por una relación de complementariedad inversa entre dichas lecturas de secuencia. En aún otra realización, cada una de las regiones somáticamente reorganizadas comprende una región V y una región J y la etapa de secuenciación incluye además la determinación de una secuencia de cada una de las moléculas de ácido nucleico individuales a partir de una o más de sus lecturas de secuencia hacia delante y al menos una lectura de secuencia inversa partiendo de una posición en una región J y extendiéndose en la dirección de su región V asociada.

Una muestra de un paciente puede proceder de una variedad de tejidos, pero normalmente la muestra es una muestra de sangre. Se extrae ARN de la muestra usando técnicas convencionales como fuente de ácidos nucleicos amplificados y procesados según Faham y Willis (antes citado).

En otro aspecto de la invención, se puede detectar o medir la presencia, ausencia y/o cantidad de los segmentos

TCR $\beta$  por medio de un inmunoensayo usando uno o más anticuerpos específicos para péptidos de 6 a 25 aminoácidos de longitud derivados de un segmento contiguo de LCASSLEASGSSYNEQFFGPGTRLTV (SEC. Nº ID. 1) o VYFCASSDSSGSDTQYFGPGTRLTV (SEC. Nº ID. 2). Se encuentran directrices para construir inmunoensayos en muchos tratados, incluyendo Wild, Editor, *The Immunoassay Handbook*, Tercera Edición (Elsevier Science, 2005). Se encuentran directrices para preparar anticuerpos específicos de péptidos en la patente estadounidense 5.231.012. También se pueden usar anticuerpos específicos para los segmentos anteriores para detectar y cuantificar por citometría de flujo las células T que tienen TCR con los segmentos, *v.g.*, Thiel *et al.*, *Clinical Immunology*, 111(2): 155-161 (2004); Gratama *et al.*, *Cytometry parte A*, 58A: 79-86 (2004); Sims *et al.*, *Expert Reviews of Vaccines*, 9(7): 765-774 (2010); y similares.

En otro aspecto de la invención, se pueden usar anticuerpos específicos para los TCR $\beta$  con los anteriores segmentos para inhibir la función de las células T que llevan dichos receptores, incluyendo, aunque sin limitación, efectos de relación autoinmune de dichas células T, tales como efectos relacionados con la EA. En este aspecto de la invención, se administra una cantidad efectiva de un anticuerpo terapéutico específico para el péptido LCASSLEASGSSYNEQFFGPGTRLTV (SEC. Nº ID. 1) o un segmento de 6-20 aminoácidos del mismo o VYFCASSDSSGSDTQYFGPGTRLTV (SEC. Nº ID. 2) o un segmento de 6-20 aminoácidos del mismo a un paciente que sufre EA.

#### Muestras

Se obtienen perfiles de clonotipos para uso con los métodos de la invención a partir de muestras de células T, las cuales están presentes en una amplia variedad de tejidos. Las células T incluyen las células T "helper" (células T efectoras o células Th), las células T citotóxicas (CTL), las células T de memoria y las células T reguladoras, que pueden distinguirse por medio de marcadores de la superficie celular. En un aspecto, una muestra de células T incluye al menos 1.000 células T, pero más típicamente una muestra incluye al menos 10.000 células T y más típicamente al menos 100.000 células T. En otro aspecto, una muestra incluye un número de células T en el rango de 1.000 a 1.000.000 de células.

Las muestras (a las que a veces se hace referencia como "muestras tisulares") usadas en los métodos de la invención pueden proceder de una variedad de tejidos, incluyendo, por ejemplo, la sangre y el plasma sanguíneo, el líquido linfático, el líquido cefalorraquídeo que rodea al encéfalo y a la médula espinal, el líquido sinovial que rodea a las articulaciones óseas y similares. En una realización, la muestra es una muestra de sangre. La muestra de sangre puede ser de aproximadamente 0,1, 0,2, 0,3, 0,4, 0,5, 0,6, 0,7, 0,8, 0,9, 1,0, 1,5, 2,0, 2,5, 3,0, 3,5, 4,0, 4,5 ó 5,0 ml.

Una muestra o muestra tisular incluye ácido nucleico, por ejemplo ADN (*v.g.*, ADN genómico) o ARN (*v.g.*, ARN mensajero). El ácido nucleico puede ser ADN o ARN libre de células, *v.g.*, extraído del sistema circulatorio, Vlassov *et al.*, *Curr. Mol. Med.*, 10: 142-165 (2010); Swarup *et al.*, *FEBS Lett.*, 581: 795-799 (2007). En los métodos de la invención, la cantidad de ARN o de ADN de un sujeto que se puede analizar varía ampliamente. Para generar un perfil de clonotipos, debe haber suficiente ácido nucleico en una muestra como para obtener una representación útil del repertorio de TCR de un individuo. Más concretamente, para generar un perfil de clonotipos a partir de ADN genómico, se extrae al menos 1 ng de ADN total de células T (es decir, aproximadamente 300 equivalentes de genoma diploide) de una muestra; en otra realización, se extraen al menos 2 ng de ADN total (es decir, aproximadamente 600 equivalentes de genoma diploide) de una muestra; y en otra realización se extraen al menos 3 ng de ADN total (es decir, aproximadamente 900 equivalentes de genoma diploide) de una muestra. Alguien con conocimientos ordinarios reconocería que, a medida que la fracción de linfocitos en una muestra disminuye, las cantidades mínimas anteriores de ADN deben aumentar para generar un perfil de clonotipos que contenga más de aproximadamente 1.000 clonotipos independientes. Para generar un perfil de clonotipos a partir de ARN, en una realización, se extrae una cantidad suficiente de ARN, de tal manera que se obtienen al menos 1.000 transcritos que codifican distintos TCR, o fragmentos de los mismos. La cantidad de ARN que corresponde a este límite varía ampliamente de una muestra a otra dependiendo de la fracción de linfocitos en una muestra, de la fase de desarrollo de los linfocitos y similares. En una realización, se extraen al menos 100 ng de ARN de una muestra de tejido que contiene células T para generar un perfil de clonotipos; en otra realización, se extraen al menos 500 ng de ARN de una muestra de tejido que contiene células T para la generación de un perfil de clonotipos. El ARN usado en los métodos de la invención puede ser o bien ARN total extraído de una muestra de tejido, o bien ARN poliA extraído directamente de una muestra de tejido o de ARN total extraído de una muestra de tejido. Las anteriores extracciones de ácido nucleico pueden ser llevadas a cabo usando kits comerciales, *v.g.*, de Invitrogen (Carlsbad, CA), Qiagen (San Diego, CA) o proveedores similares. Se encuentran directrices para la extracción de ARN en Liedtke *et al.*, *PCR Methods and Applications*, 4: 185-187 (1994), y referencias similares.

En algunas realizaciones, una muestra que contiene linfocitos es lo suficientemente grande como para que substancialmente cada célula T con un clonotipo distinto esté representada en ella, formando así un repertorio (tal como se usa aquí el término). En una realización, se toma una muestra que contiene con una probabilidad del noventa y nueve por ciento cada clonotipo de una población presente a una frecuencia del 0,001 por ciento o mayor. En otra realización, se toma una muestra que contiene con una probabilidad del noventa y nueve por ciento cada clonotipo de una población presente a una frecuencia del 0,0001 por ciento o mayor. En una realización, una muestra de células T incluye al menos medio millón de células, y en otra realización dicha muestra incluye al menos

un millón de células.

Siempre que una fuente de material de la que se toma una muestra sea escasa, tal como muestras para estudios clínicos o similares, se puede amplificar el ADN del material por una técnica sin sesgo, tal como amplificación del genoma completo (AGC), amplificación de múltiple desplazamiento (AMD) o una técnica similar, *v.g.*, Hawkins *et al.*, *Curr. Opin. Biotech.*, 13: 65-67 (2002); Dean *et al.*, *Genome Research*, 11: 1095-1099 (2001); Wang *et al.*, *Nucleic Acids Research*, 32: e76 (2004); Hosono *et al.*, *Genome Research*, 13: 954-964 (2003), y similares.

Las muestras de sangre tienen un interés particular y pueden ser obtenidas usando técnicas convencionales, *v.g.*, Innis *et al.*, editores, *PCR Protocols* (Academic Press, 1990), o similares. Por ejemplo, se pueden separar las células blancas de la sangre de muestras de sangre usando técnicas convencionales, *v.g.*, RosetteSep kit (Stem Cell Technologies, Vancouver, Canadá). De igual modo, se pueden aislar otras fracciones de sangre entera, tales como las células mononucleares de sangre periférica (PBMC) para uso con los métodos de la invención usando kits comerciales, (*v.g.* Miltenyi Biotec, Auburn, CA) o similares. Las muestras de sangre pueden variar en cuanto a volumen de 100  $\mu$ l a 10 ml; en un aspecto, los volúmenes de las muestras de sangre son de 200  $\mu$ l a 2 ml. Se pueden extraer entonces el ADN y/o el ARN de dicha muestra de sangre usando técnicas convencionales para uso en los métodos de la invención, *v.g.*, DNeasy Blood & Tissue Kit (Qiagen, Valencia, CA). Eventualmente, se pueden aislar además subgrupos de células blancas de la sangre, *v.g.*, linfocitos, usando técnicas convencionales, *v.g.*, clasificación de células activada por fluorescencia (FACS) (Becton Dickinson, San José, CA), clasificación de células activada magnéticamente (MACS) (Miltenyi Biotec, Auburn, CA) o similares.

#### Anticuerpos para tratamiento y detección

Se pueden usar los péptidos relacionados con la EA o segmentos de los mismos para producir anticuerpos para aplicaciones terapéuticas o de inmunoensayo usando técnicas convencionales de anticuerpos para péptidos, *v.g.*, patente estadounidense 5.231.012; patente estadounidense 4.474.754; Walter *et al.*, *Genetic Engineering*, 5: 61-91 (1983), o similares. Resumiendo, se conjuga un péptido relacionado con la EA o un segmento del mismo con una molécula transportadora, línea celular para formar hibridomas, que son cribados en cuanto a anticuerpos específicos de péptidos que tienen la afinidad y especificidad deseadas. Dichos anticuerpos pueden ser además procesados, *v.g.*, para mejorar la afinidad o la especificidad, para reducir la inmunogenicidad y similares, mediante el uso de técnicas conocidas de ingeniería de anticuerpos, tales como las desveladas en las referencias que se citan más adelante. Dicho otro procesamiento puede incluir la humanización, *v.g.*, como se desvela en las patentes estadounidenses 7.892.550 y 8.030.023.

Una vez se dispone de células B de un animal inmunizado, *v.g.*, un conejo, se producen hibridomas por técnicas bien conocidas. Normalmente, el procedimiento conlleva la fusión de una línea celular inmortalizada con un linfocito B que produce el anticuerpo deseado. De manera alternativa, son posibles técnicas que no emplean la fusión para generar líneas celulares productoras de un anticuerpo inmortal, y entran dentro del alcance de la presente invención, *v.g.*, transformación inducida por virus: Casali *et al.*, "Human Monoclonals from Antigen-Specific Selection of B Lymphocytes and Transformation by EBV", *Science*, Vol. 234, págs. 476-479 (1986). Las líneas celulares inmortalizadas son normalmente células de mamífero transformadas, en particular células de mieloma de origen en roedores, bóvidos y humanos. Más frecuentemente, se emplean líneas celulares de mieloma de rata o de ratón por razones de conveniencia y disponibilidad. Las técnicas para obtener los linfocitos apropiados de mamíferos a los que se inyecta el antígeno diana son bien conocidas. En general, se usan linfocitos de sangre periférica (PBL) si se desean células de origen humano, o se usan células de bazo o células de ganglios linfáticos si se desean fuentes de mamíferos no humanos. Se inyectan a un mamífero hospedador dosis repetidas del antígeno purificado y se permite que el mamífero genere las células productoras de anticuerpos deseadas antes de recoger éstas para fusión con la línea celular inmortalizada. Las técnicas para la fusión son también bien conocidas en este campo, y, en general, conllevan la mezcla de las células con un agente de fusión, tal como polietilenglicol. Los hibridomas son seleccionados por procedimientos estándar, tales como selección HAT. De entre estos hibridomas, se seleccionan los que segregan el anticuerpo deseado, es decir, los que son específicos para el péptido deseado, estudiando su medio de cultivo por inmunoensayos estándar, tales como Western blotting, ELISA, RIA, capacidad neutralizadora del CSIF o similares. Los anticuerpos son recuperados del medio usando técnicas de purificación de proteínas estándar, *v.g.*, Tijssen, *Practice and Theory of Enzyme Immunoassays* (Elsevier, Amsterdam, 1985). Se dispone de muchas referencias para obtener directrices en la aplicación de cualquiera de las técnicas anteriores, *v.g.*, Kohler *et al.*, *Hybridoma Techniques* (Cold Spring Harbor Laboratory, New York, 1980); Tijssen, *Practice and Theory of Enzyme Immunoassays* (Elsevier, Amsterdam, 1985); Campbell, *Monoclonal Antibody Technology* (Elsevier, Amsterdam, 1984); Hurrell, *Monoclonal Hybridoma Antibodies: Techniques and Applications* (CRC Press, Boca Ratón, Fla. 1982), y similares. También se pueden producir anticuerpos y fragmentos de anticuerpos característicos de los hibridomas de la invención por medios recombinantes extrayendo ARN mensajero, construyendo una genoteca de ADNc y seleccionando clones que codifican segmentos de la molécula de anticuerpo, *v.g.*, Huse *et al.*, *Science*, Vol. 246, págs. 1275-1281 (1989). Una vez se dispone de una secuencia nucleotídica que codifica la región variable de un anticuerpo adecuado, se pueden mejorar las propiedades de dicho anticuerpo usando técnicas convencionales, por ejemplo como se desvela en las siguientes referencias: Barbas *et al.*, *Proc. Natl. Acad. Sci.*, 88: 7978-7982 (1991), y pHEN1 y sus miembros de familia relacionados, *v.g.*, desvelado en Hoogenboom *et al.*, *Nucleic Acids Research*, 19: 4133-4137 (1991); y patentes estadounidenses 5.969.108, 6.806.079, 7.662.557 y patentes

relacionadas; y Sidhu, editor, PhageDisplay in Biotechnology and Drug Discovery (CRC Press, 2005); Lutz y Bornscheuer, Editores, Protein Engineering Handbook (Wiley-VCH, 2009); y similares.

Una vez se ha obtenido un anticuerpo terapéutico, éste puede ser sometido de nuevo a ingeniería y/o fabricado y formulado para tratar a humanos usando métodos conocidos en la técnica, v.g., como se desvela en las patentes estadounidenses 7.892.550 y 8.030.023. Normalmente, un anticuerpo terapéutico es un anticuerpo aislado de la invención que se incluye en una formulación terapéutica. En un aspecto, la invención proporciona un método de tratamiento de la espondilitis anquilosante en un sujeto, consistiendo dicho método en administrar al sujeto una cantidad efectiva de un anticuerpo de la invención, mediante lo cual se trata dicha afección. En un aspecto, la invención proporciona el uso de un anticuerpo de la invención en la preparación de un medicamento para el tratamiento terapéutico y/o profiláctico de la espondilitis anquilosante.

Se preparan formulaciones terapéuticas que contienen un anticuerpo de la invención para almacenamiento mezclando el anticuerpo que tiene el grado deseado de pureza con soportes fisiológicamente aceptables, excipientes o estabilizadores eventuales (*Remington's Pharmaceutical Sciences* 16ª edición, Osol, A. Ed. (1980)), en forma de soluciones acuosas, formulaciones liofilizadas u otras formulaciones desecadas. Los soportes, excipientes o estabilizadores aceptables son no tóxicos para los receptores a las dosificaciones y concentraciones empleadas, e incluyen tampones, tales como fosfato, citrato, histidina y otros ácidos orgánicos; antioxidantes, incluyendo ácido ascórbico y metionina; conservantes (tales como cloruro de octadecildimetilbencilamonio; cloruro de hexametonio; cloruro de benzalconio, cloruro de bencetonio; fenol, alcohol butílico o bencilico; alquilparabenos, tales como metil- o propilparabén; catecol; resorcinol; ciclohexanol; 3-pentanol; y m-cresol); polipéptidos de bajo peso molecular (menos de aproximadamente 10 residuos); proteínas, tales como seroalbúmina, gelatina o inmunoglobulinas; polímeros hidrofílicos, tales como polivinilpirrolidona; aminoácidos, tales como glicina, glutamina, asparraguina, histidina, arginina o lisina; monosacáridos, disacáridos y otros carbohidratos, incluyendo glucosa, manosa o dextrinas; agentes quelantes, tales como EDTA; azúcares, tales como sacarosa, manitol, trehalosa o sorbitol; contraiones formadores de sales, tales como sodio; complejos metálicos (v.g., complejos Zn-proteína); y/o surfactantes no iónicos, tales como TWEEN™, PLURONICS™ o polietilenglicol (PEG).

La formulación puede contener también aquí más de un compuesto activo, según sea necesario para la indicación particular que se esté tratando, preferiblemente los que tienen actividades complementarias y que no tienen efectos adversos entre sí. Dichas moléculas están adecuadamente presentes en combinación en cantidades que son efectivas para el fin pretendido.

Los principios activos pueden también estar atrapados en microcápsulas preparadas, por ejemplo, mediante técnicas de coacervación o mediante polimerización interfacial, por ejemplo, microcápsulas de hidroximetilcelulosa o gelatina y microcápsulas de poli(metilmacrilato), respectivamente, en sistemas de administración de fármacos coloidales (por ejemplo, liposomas, microesferas de albúmina, microemulsiones, nanopartículas y nanocápsulas) o en macroemulsiones. Dichas técnicas están desveladas en *Remington's Pharmaceutical Sciences* 16ª edición, Osol, A. Ed. (1980).

Las formulaciones para uso en administración *in vivo* deben ser estériles. Se consigue esto fácilmente por filtración a través de membranas de filtración estériles.

Se pueden preparar preparaciones de liberación mantenida. Como ejemplos adecuados de preparaciones de liberación mantenida, se incluyen matrices semipermeables de polímeros hidrofóbicos sólidos que contienen la inmunoglobulina de la invención, estando estas matrices en forma de artículos dotados de forma, v.g., películas o microcápsulas. Como ejemplos de matrices de liberación mantenida, se incluyen poliésteres, hidrogeles (por ejemplo, poli(2-hidroxietilo-metacrilato) o poli(alcohol vinílico)), polilactidas (Patente Estadounidense N° 3.773.919), copolímeros de ácido L-glutámico y  $\gamma$ -etil-L-glutamato, etileno-acetato de vinilo no degradable, copolímeros de ácido láctico-ácido glicólico degradables, tales como el LUPRON DEPOT™ (microesferas inyectables compuestas de copolímero de ácido láctico-ácido glicólico y acetato de leuprolida) y ácido poli-D-(-)-3-hidroxibutírico. Aunque polímeros tales como etileno-acetato de vinilo y ácido láctico-ácido glicólico permiten la liberación de moléculas durante más de 100 días, ciertos hidrogeles liberan proteínas durante periodos de tiempo más cortos. Cuando las inmunoglobulinas encapsuladas permanecen en el cuerpo durante un tiempo prolongado, pueden desnaturalizarse o agregarse como resultado de la exposición a la humedad a 37°C, para dar lugar a una pérdida de actividad biológica y a posibles cambios en la inmunogenicidad. Se pueden idear estrategias racionales para la estabilización dependiendo del mecanismo implicado. Por ejemplo, si se descubre que el mecanismo de agregación es la formación de enlaces S-S intermoleculares a través de intercambio tio-disulfuro, se puede conseguir la estabilización modificando los residuos de sulfhidrilo, liofilizando a partir de soluciones ácidas, controlando el contenido en humedad, usando aditivos apropiados y desarrollando composiciones de matrices poliméricas específicas.

En otro aspecto de la invención, se proporciona un artículo de fabricación que contiene materiales útiles para el tratamiento, la prevención y/o el diagnóstico de los trastornos antes descritos. El artículo de fabricación comprende un recipiente y una etiqueta o un prospecto de envase sobre o asociado al recipiente. Como recipientes adecuados, se incluyen, por ejemplo, botellas, viales, jeringas, etc. Los recipientes pueden estar formados con una variedad de materiales, tales como vidrio o plástico. El recipiente contiene una composición que es por sí misma, o cuando se

combina con otra composición, efectiva para tratar, prevenir y/o diagnosticar la afección, y puede tener una vía de acceso estéril (por ejemplo, el recipiente puede ser una bolsa de solución intravenosa o un vial que tenga un tapón perforable con una aguja de inyección hipodérmica). Al menos un principio activo en la composición es un anticuerpo de la invención. De manera alternativa o adicional, el artículo de fabricación puede además incluir un segundo recipiente que contenga un tampón farmacéuticamente aceptable, tal como agua bacteriostática para inyección (BWF1), solución salina tamponada con fosfatos, solución de Ringer y solución de dextrosa. Puede también incluir otros materiales deseables desde un punto de vista comercial y de usuario, incluyendo otros tampones, diluyentes, filtros, agujas y jeringas.

10 Amplificación de poblaciones de ácidos nucleicos para perfiles de clonotipos

Se pueden generar amplicones de poblaciones diana de ácidos nucleicos mediante una variedad de técnicas de amplificación. En un aspecto de la invención, se usa PCR múltiple para amplificar miembros de una mezcla de ácidos nucleicos, en particular mezclas consistentes en moléculas inmunes recombinadas, tales como receptores de las células T o porciones de los mismos. Se encuentran directrices para llevar a cabo PCR múltiple de dichas moléculas inmunes en las siguientes referencias: Morley, patente estadounidense 5.296.351; Gorski, patente estadounidense 5.837.447; Dau, patente estadounidense 6.087.096; Von Dongen *et al.*, publicación de patente estadounidense 2006/0234234; publicación de patente europea EP 1544308B1; y similares.

20 Tras amplificación del ADN a partir del genoma (o amplificación de ácido nucleico en forma de ADNc por transcripción inversa de ARN), se pueden aislar y eventualmente reamplificar las moléculas de ácido nucleico individuales, y secuenciar luego individualmente. Se pueden encontrar ejemplos de protocolos de amplificación en van Dongen *et al.*, *Leukemia*, 17: 2257-2317 (2003), o van Dongen *et al.*, publicación de patente estadounidense 2006/0234234. Resumiendo, un protocolo ejemplar es el siguiente:

25 Tampón de reacción: ABI Buffer II o ABI Gold Buffer (Life Technologies, San Diego, CA); volumen de reacción final 50 µl; muestra de 100 ng de ADN; 10 pmol de cada cebador (sujeto a ajustes para equilibrar la amplificación como se describe más adelante); dNTP a 200 µM de concentración final; MgCl<sub>2</sub> a 1,5 mM de concentración final (sujeto a optimización dependiendo de las secuencias diana y de la polimerasa); polimerasa Taq (1-2 U/tubo); 30 condiciones de ciclo: preactivación 7 min. a 95°C; hibridación a 60°C; tiempos de ciclo: 30 s desnaturalización; 30 s hibridación; 30 s extensión. Las polimerasas que se pueden emplear para la amplificación en los métodos de la invención están comercializadas e incluyen, por ejemplo, la polimerasa Taq, la polimerasa AccuPrime o Pfu. La elección de la polimerasa que se ha de utilizar puede basarse en si se prefiere fidelidad o eficacia.

35 Se pueden usar PCR en tiempo real, tinción PicoGreen, electroforesis nanofluida (v.g. LabChip) o mediciones de absorción UV en una etapa inicial para juzgar la cantidad funcional de material amplificable.

En un aspecto, se realizan amplificaciones múltiplex de tal forma que las cantidades relativas de secuencias en una población de partida sean substancialmente las mismas que las de la población amplificada, o amplicón. Es decir, se llevan a cabo múltiples amplificaciones con un sesgo mínimo de amplificación entre las secuencias miembro de una población de muestras. En una realización, dichas cantidades relativas son substancialmente las mismas si cada cantidad relativa en un amplicón está dentro de cinco veces su valor en la muestra de partida. En otra realización, dichas cantidades relativas son substancialmente las mismas si cada cantidad relativa en un amplicón está dentro de dos veces su valor en la muestra de partida. Tal como se discute más completamente a continuación, el sesgo de amplificación en la PCR puede ser detectado y corregido usando técnicas convencionales, de tal manera que se puede seleccionar un conjunto de cebadores de PCR para un repertorio predeterminado que proporcionen una amplificación sin sesgo de cualquier muestra.

50 Con respecto a muchos repertorios basados en secuencias de TCR o BCR, una amplificación múltiple usa eventualmente todos los segmentos V. Se optimiza la reacción para intentar conseguir una amplificación que mantenga la abundancia relativa de las secuencias amplificadas por diferentes cebadores de segmentos V. Algunos de los cebadores están relacionados, por lo que muchos de los cebadores pueden "interaccionar", amplificando plantillas que no están perfectamente emparejadas con ella. Se optimizan las condiciones de tal forma que cada plantilla puede ser amplificada de un modo similar independientemente de qué cebador la amplificó. En otras palabras, si hay dos plantillas, entonces, tras una amplificación de 1.000 veces, ambas plantillas pueden ser amplificadas aproximadamente 1.000 veces, y no importa que para una de las plantillas la mitad de los productos amplificados lleven un cebador diferente debido a la interacción. En un análisis ulterior de los datos de secuenciación, se elimina la secuencia cebadora del análisis, por lo que no importa qué cebador se use en la amplificación siempre que las plantillas se amplifiquen de igual modo.

60 En una realización, se puede evitar el sesgo de amplificación realizando una amplificación en dos fases (como se describe en Faham y Willis, antes citado), donde se implementan un pequeño número de ciclos de amplificación en una etapa primera, o primaria, usando cebadores que tienen colas no complementarias con respecto a las secuencias diana. Las colas incluyen sitios de unión de cebadores que se añaden a los extremos de las secuencias del amplicón primario, de tal forma que dichos sitios son usados en una amplificación de segunda etapa usando sólo un único cebador hacia delante y un único cebador inverso, eliminando así una causa primaria de sesgo de



amplificación. Preferiblemente, la PCR primaria tendrá un número suficientemente pequeño de ciclos (v.g. 5-10) para minimizar la amplificación diferencial por los diferentes cebadores. Se realiza la amplificación secundaria con un par de cebadores y por ello el problema de la amplificación diferencial es mínimo. Se lleva un uno por ciento de la PCR primaria directamente a la PCR secundaria. Treinta y cinco ciclos (equivalentes a ~28 ciclos sin la etapa de dilución de 100 veces) usados entre las dos amplificaciones fueron suficientes para mostrar una sólida amplificación independientemente de si la descomposición de ciclos fue: un ciclo primario y 34 secundarios o 25 primarios y 10 secundarios. Incluso aunque de manera ideal la realización de sólo 1 ciclo en la PCR primaria pueda disminuir el sesgo de amplificación, existen otras consideraciones. Un aspecto de esto es la representación. Ésta juega un papel cuando la cantidad de entrada de partida no está en exceso con respecto al número de lecturas finalmente obtenidas. Por ejemplo, si se obtienen 1.000.000 de lecturas y se parte de 1.000.000 de moléculas de entrada y tomando luego sólo representación a partir de 100.000 moléculas para la amplificación secundaria, se degradaría la precisión de la estimación de la abundancia relativa de las diferentes especies en la muestra original. La dilución de 100 veces entre las dos etapas significa que la representación se reduce, a menos que la amplificación por PCR primaria generara significativamente más de 100 moléculas. Esto indica que se pueden usar un mínimo de 8 ciclos (256 veces), pero más confortablemente 10 ciclos (~1.000 veces). La alternativa a ello es llevar más de un 1% de la PCR primaria a la secundaria, pero, debido a la gran concentración de cebador usada en la PCR primaria, se puede usar un gran factor de dilución para asegurarse de que estos cebadores no interfieran en la amplificación y no empeoren el sesgo de amplificación entre secuencias. Otra alternativa es añadir una etapa de purificación o enzimática para eliminar los cebadores de la PCR primaria y permitir una dilución más pequeña de ella. En este ejemplo, la PCR primaria tuvo 10 ciclos y la secundaria 25 ciclos.

#### Generación de lecturas de secuencias para clonotipos

Se puede usar cualquier técnica de alto rendimiento para secuenciar ácidos nucleicos en el método de la invención. Preferiblemente, dicha técnica tiene una capacidad de generar de un modo efectivo en cuanto a costes un volumen de datos de secuencias del que se pueden determinar al menos 1.000 clonotipos, y preferiblemente del que se pueden determinar al menos de 10.000 a 1.000.000 de clonotipos. Las técnicas de secuenciación de ADN incluyen reacciones de secuenciación dideoxi clásicas (método de Sanger) usando finalizadores o cebadores marcados y separación en gel en capas o capilar, secuenciación por síntesis usando nucleótidos marcados reversiblemente terminados, pirosecuenciación, secuenciación 454, hibridación específica de alelos con una genoteca de sondas oligonucleotídicas marcadas, secuenciación por síntesis usando hibridación específica de alelos con una genoteca de clones marcados seguida de ligación, monitorización en tiempo real de la incorporación de nucleótidos marcados durante una etapa de polimerización, secuenciación "polony" y secuenciación SOLiD. Se ha demostrado más recientemente la secuenciación de las moléculas separadas por reacciones de extensión secuenciales o únicas usando polimerasas o ligasas, así como por hibridaciones diferenciales únicas o secuenciales con genotecas de sondas. Estas reacciones han sido realizadas sobre muchas secuencias clonales en paralelo, incluyendo demostraciones en aplicaciones comerciales actuales de más de 100 millones de secuencias en paralelo. Estos enfoques de secuenciación pueden, por lo tanto, ser utilizados para estudiar el repertorio de receptores de células T (TCR) y/o de receptores de células B (BCR). En un aspecto de la invención, se emplean métodos de alto rendimiento de secuenciación que comprenden una etapa de aislamiento espacial de moléculas individuales sobre una superficie sólida, donde se secuencian en paralelo. Dichas superficies sólidas pueden incluir superficies no porosas (tales como en la secuenciación Solexa, v.g., Bentley *et al.*, Nature, 456: 53-59 (2008), o la secuenciación Complete Genomics, v.g., Drmanac *et al.*, Science, 327: 78-81 (2010)), matrices de pocillos, que pueden incluir plantillas unidas a perlas o a partículas (tales como con 454, v.g., Margulies *et al.*, Nature, 437: 376-380 (2005), o la secuenciación Ion Torrent, publicación de patente estadounidense 2010/0137143 o 2010/0304982), membranas microfabricadas (tales como con la secuenciación SMRT, v.g., Eid *et al.*, Science, 323: 133-138 (2009)), o matrices de perlas (como con la secuenciación SOLiD o la secuenciación polony, v.g., Kim *et al.*, Science, 316: 1481-1414 (2007)). En otro aspecto, dichos métodos comprenden la amplificación de las moléculas aisladas antes o después de aislarlas espacialmente sobre una superficie sólida. La amplificación previa puede incluir una amplificación basada en emulsión, tal como PCR en emulsión, o amplificación en círculo rodante. Es de particular interés la secuenciación basada en Solexa, donde se aíslan espacialmente moléculas de plantilla individuales sobre una superficie sólida, después de lo cual se amplifican en paralelo por PCR puente para formar poblaciones clonales separadas, o grupos, y se secuencian luego, como se describe en Bentley *et al.* (antes citado) y en las instrucciones del fabricante (v.g. TruSeq™ Sample Preparation Kit y Data Sheet, Illumina, Inc., San Diego, CA, 2010), y también en las siguientes referencias: patentes estadounidenses 6.090.592, 6.300.070, 7.115.400 y EP0972081B1. En una realización, las moléculas individuales dispuestas y amplificadas sobre una superficie sólida forman grupos en una densidad de al menos  $10^5$  grupos por  $\text{cm}^2$ ; o en una densidad de al menos  $5 \times 10^5$  por  $\text{cm}^2$ ; o en una densidad de al menos  $10^6$  grupos por  $\text{cm}^2$ . En una realización, se emplean químicas de secuenciación que tienen índices de error relativamente altos. En dichas realizaciones, las puntuaciones medias de calidad producidas por dichas químicas son funciones monotónicamente declinantes de longitudes de lectura de secuencias. En una realización, dicha declinación corresponde a que un 0,5 por ciento de lecturas de secuencias tienen al menos un error en las posiciones 1-75; un 1 por ciento de lecturas de secuencias tienen al menos un error en las posiciones 76-100; y un 2 por ciento de lecturas de secuencias tienen al menos un error en las posiciones 101-125.

En un aspecto, se obtiene un perfil de clonotipos basado en secuencia de un individuo usando las siguientes etapas: (a) obtener una muestra de ácido nucleico de las células T y/o de las células B del individuo; (b) aislar espacialmente

las moléculas individuales derivadas de dicha muestra de ácido nucleico, incluyendo las moléculas individuales al menos una plantilla generada a partir de un ácido nucleico en la muestra, donde dicha plantilla tiene una región somáticamente redispuesta o una porción de la misma, siendo capaz cada molécula individual de producir al menos una lectura de secuencia; (c) secuenciar dichas moléculas individuales espacialmente aisladas, y (d) determinar las abundancias de diferentes secuencias de las moléculas de ácido nucleico de la muestra de ácido nucleico para generar el perfil de clonotipos. En una realización, cada una de las regiones somáticamente redispuestas comprende una región V y una región J. En otra realización, la etapa de secuenciación comprende la secuenciación bidireccional de cada una de las moléculas individuales espacialmente aisladas para producir al menos una lectura de secuencia hacia delante y al menos una lectura de secuencia inversa. Además con respecto a esta última realización, al menos una de las lecturas de secuencia hacia delante y al menos una de las lecturas de secuencia inversas tienen una región de solapamiento, de tal modo que las bases de dicha región de solapamiento son determinadas por una relación complementaria inversa entre dichas lecturas de secuencia. En aún otra realización, cada una de las regiones somáticamente redispuestas comprende una región V y una región J, y la etapa de secuenciación incluye además la determinación de una secuencia de cada una de las moléculas de ácido nucleico individuales a partir de una o más de sus lecturas de secuencia hacia delante y al menos una lectura de secuencia inversa partiendo de una posición en una región J y extendiéndose en la dirección de su región V asociada. En otra realización, la etapa de secuenciación incluye la generación de las lecturas de secuencia que tienen puntuaciones de calidad monótonicamente decrecientes. Además con respecto a esta última realización, las puntuaciones de calidad monótonicamente decrecientes son tales que las lecturas de secuencia tienen índices de error no mejores que los siguientes: un 0,2 por ciento de las lecturas de secuencia contienen al menos un error en las posiciones de bases 1 a 50, de un 0,2 a un 1,0 por ciento de las lecturas de secuencia contienen al menos un error en las posiciones 51-75, de un 0,5 a un 1,5 por ciento de las lecturas de secuencia contienen al menos un error en las posiciones 76-100. En otra realización, el método anterior comprende las siguientes etapas: (a) obtener una muestra de ácido nucleico de las células T del individuo; (b) aislar espacialmente las moléculas individuales derivadas de dicha muestra de ácido nucleico, teniendo las moléculas individuales conjuntos anidados de plantillas generadas cada una de ellas a partir de un ácido nucleico en la muestra y conteniendo cada una una región somáticamente redispuesta o una porción de la misma, siendo capaz cada conjunto anidado de producir una pluralidad de lecturas de secuencia que se extienden cada una en la misma dirección y que parten cada una de una posición diferente en el ácido nucleico del que se generó el conjunto anidado; (c) secuenciar dichas moléculas individuales espacialmente aisladas; y (d) determinar las abundancias de diferentes secuencias de las moléculas de ácido nucleico de la muestra de ácido nucleico para generar el perfil de clonotipos. En una realización, la etapa de secuenciación incluye la producción de una pluralidad de lecturas de secuencia para cada uno de los conjuntos anidados. En otra realización, cada una de las regiones somáticamente redispuestas comprende una región V y una región J, y cada una de la pluralidad de lecturas de secuencia parte de una posición diferente en la región V y se extiende en la dirección de su región J asociada.

#### Determinación de clonotipos a partir de los datos de secuencia

La construcción de clonotipos a partir de los datos de lectura de secuencias depende en parte del método de secuenciación usado para generar dichos datos, ya que los diferentes métodos tienen diferentes longitudes de lectura esperadas y diferente calidad de datos. En un enfoque, se emplea un secuenciador Solexa para generar datos de lectura de secuencias para análisis, como se describe en Faham y Willis (antes citado). En una realización, se obtiene una muestra que proporciona al menos  $0,5-1,0 \times 10^6$  linfocitos para producir al menos 1 millón de moléculas de plantilla, las cuales, tras una amplificación eventual, pueden producir un millón o más correspondientes de poblaciones clonales de moléculas (o grupos) de plantilla. Para la mayoría de los enfoques de secuenciación de alto rendimiento, incluyendo el enfoque Solexa, es deseable dicho sobremuestreo a nivel de grupos, de tal forma que cada secuencia de plantilla es determinada con un gran grado de redundancia para aumentar la precisión de la determinación de secuencias. Para las implementaciones basadas en Solexa, preferiblemente se determina la secuencia de cada plantilla independiente 10 veces o más. Para otros enfoques de secuenciación con diferentes longitudes de lectura esperadas y diferente calidad de datos, se pueden usar diferentes niveles de redundancia para una precisión comparable de la determinación de secuencia. Quienes tienen conocimientos ordinarios en la técnica reconocen que los anteriores parámetros, *v.g.*, tamaño de la muestra, redundancia y similares, son elecciones de diseño relacionadas con aplicaciones particulares.

En un aspecto de la invención, se pueden determinar las secuencias de clonotipos combinando información de una o más lecturas de secuencias, por ejemplo, a lo largo de las regiones V(D)J de las cadenas seleccionadas. En otro aspecto, se determinan las secuencias de clonotipos combinando información de una pluralidad de lecturas de secuencias. Dichas pluralidades de lecturas de secuencias pueden incluir una o más lecturas de secuencias a lo largo de una hebra sentido (es decir, lecturas de secuencias "hacia delante") y una o más lecturas de secuencias a lo largo de su hebra complementaria (es decir, lecturas de secuencias "inversas"). Cuando se generan múltiples lecturas de secuencias a lo largo de la misma hebra, se generan primeramente plantillas separadas amplificando las moléculas de la muestra con cebadores seleccionados para las diferentes posiciones de las lecturas de secuencias. Dichas amplificaciones pueden ser llevadas a cabo en la misma reacción o en reacciones separadas. En un aspecto, siempre que se emplee PCR, se usan reacciones de amplificación separadas para generar las plantillas separadas, las cuales a su vez se combinan y se usan para generar múltiples lecturas de secuencia a lo largo de la misma hebra. Este último enfoque es preferible para evitar la necesidad de equilibrar las concentraciones de cebadores (y/u

otros parámetros de reacción) para asegurar una amplificación igual de las múltiples plantillas (a la que a veces se hace aquí referencia como "amplificación equilibrada" o "amplificación sin sesgo").

#### Análisis del repertorio de TCR $\beta$

5 En este ejemplo, se analizan las cadenas de TCR $\beta$ . El análisis incluye la amplificación, la secuenciación y el análisis de las secuencias de TCR $\beta$ . Un cebador es complementario de una secuencia común en C $\beta$ 1 y C $\beta$ 2, y existen 34 cebadores V capaces de amplificar la totalidad de los 48 segmentos V. C $\beta$ 1 o C $\beta$ 2 difieren entre sí en la posición 10 y 14 de la unión J/C. El cebador para C $\beta$ 1 y C $\beta$ 2 acaba en la posición 16 pb y no tiene preferencia para C $\beta$ 1 o C $\beta$ 2. 10 Los 34 cebadores V están modificados con respecto a un conjunto original de cebadores desvelados en Van Dongen *et al.*, publicación de patente estadounidense 2006/0234234. Los cebadores modificados están desvelados en Faham *et al.*, publicación de patente estadounidense 2010/0151471.

15 Se usa el Illumina Genome Analyzer para secuenciar el amplicón producido por los anteriores cebadores. Se realiza una amplificación en dos etapas sobre transcritos de ARN mensajero (1200), como se ilustra en las Figs. 1A-1B, empleando la primera etapa los anteriores cebadores y una segunda etapa para añadir cebadores comunes para amplificación puente y secuenciación. Tal como se muestra en la FIG. 1A, se realiza una PCR primaria usando en un lado un cebador de 20 pb (1202) cuyo extremo 3' está a 16 bases de la unión J/C (1204) y que es perfectamente 20 complementario de C $\beta$ 1 (1203) y los dos alelos de C $\beta$ 2. En la región V (1206) de los transcritos de ARN (1200), se dispone de un conjunto de cebadores (1212) que contiene secuencias cebadoras complementarias a las diferentes secuencias de la región V (34 en una realización). Los cebadores del conjunto (1212) también contienen una cola no complementaria (1214) que produce un amplicón (1216) que tiene un sitio de unión a cebador (1218) específico para cebadores P7 (1220). Tras una PCR múltiple convencional, se forma un amplicón (1216) que contiene la porción 25 altamente diversa de la región J(D)V (1206, 1208 y 1210) de los transcritos de ARNm y sitios de unión a cebador comunes (1203 y 1218) para una amplificación secundaria para añadir un marcaje de muestra (1221) y cebadores (1220 y 1222) para la formación de grupos por PCR puente. En la PCR secundaria, en el mismo lado de la plantilla, se usa un cebador (1222 en la Fig. 1B y al que se hace aquí referencia como "C10-17-P5") que tiene en su extremo 3' la secuencia de las 10 bases más próximas a la unión J/C, seguido de 17 pb con la secuencia de las posiciones 30 15-31 de la unión J/C, seguido de la secuencia P5 (1224), que tiene un papel en la formación de grupos por PCR puente en la secuenciación Solexa. (Cuando el cebador C10-17-P5 (1222) se hibrida con la plantilla generada en la primera PCR, se crea un bucle de 4 pb (posición 11-14) en la plantilla, ya que el cebador se hibrida con la secuencia de las 10 bases más próximas a la unión J/C y las bases en las posiciones 15-31 de la unión J/C. El bucle de las 35 posiciones 11-14 elimina la amplificación diferencial de plantillas que llevan C $\beta$ 1 o C $\beta$ 2. Se realiza entonces la secuenciación con un cebador complementario a la secuencia de las 10 bases más próximas a la unión J/C y las bases en las posiciones 15-31 de la unión J/C (este cebador es denominado C'). El cebador C10-17-P5 puede ser purificado por HPLC para asegurarse de que todo el material amplificado tenga extremos intactos que puedan ser eficazmente utilizados en la formación de grupos).

40 En la FIG. 1A, la longitud del saliente en los cebadores V (1212) es preferiblemente de 14 pb. Se facilita la PCR primaria con un saliente más corto (1214). De manera alternativa, en pro de la PCR secundaria, se usa el saliente en el cebador V en la PCR primaria el mayor tiempo posible ya que la PCR secundaria se ceba a partir de esta secuencia. Se investigó un tamaño mínimo de saliente (1214) que soporte una PCR secundaria eficiente. Se produjeron dos series de cebadores V (para dos segmentos V diferentes) con tamaños de saliente de 10 a 30 con 45 pasos de 2 pb. Usando las secuencias sintéticas apropiadas, se realizó la primera PCR con cada uno de los cebadores en las series y se llevó a cabo una electroforesis en gel para mostrar que todas se amplificaron.

Tal como se ilustra en la FIG. 1A, la PCR primaria usa 34 cebadores V diferentes (1212) que se hibridan con la región V (1206) de las plantillas de ARN (1200) y contienen un saliente común de 14 pb en la cola 5'. Los 14 pb son 50 la secuencia parcial de uno de los cebadores de secuenciación Illumina (denominado el cebador Read 2). El cebador de la amplificación secundaria (1220) en el mismo lado incluye la secuencia P7, un marcador (1221) y la secuencia del cebador Read 2 (1223) (este cebador es denominado Read2\_tagX\_P7). Se usa la secuencia P7 para la formación de grupos. Se usan el cebador Read 2 y su complementario para secuenciar el segmento V y el marcador, respectivamente. Se crea un conjunto de 96 de estos cebadores con marcadores numerados del 1 al 96 (véase más adelante). Estos cebadores son purificados por HPLC para asegurarse de que todo el material 55 amplificado tiene extremos intactos que puedan ser eficientemente utilizados en la formación de grupos.

60 Como se ha mencionado anteriormente, el cebador de la segunda etapa, C-10-17-P5 (1222, FIG. 1B) tiene una homología interrumpida con la plantilla generada en la PCR de la primera etapa. Se ha validado la eficacia de la amplificación usando este cebador. Un cebador alternativo a C-10-17-P5, denominado CsegP5, tiene una perfecta homología con el cebador C de la primera etapa y una cola 5' que lleva P5. Se comparó la eficacia de la utilización de C-10-17-P5 y CsegP5 en la amplificación de las plantillas de la PCR de la primera etapa realizando una PCR en tiempo real. En varias réplicas, se vio que la PCR que usaba el cebador C-10-17-P5 tenía poca o ninguna diferencia en cuanto a eficacia en comparación con la PCR que usaba el cebador CsegP5.

65 El amplicón (1230) resultante de la amplificación en dos etapas ilustrada en las Figs. 1A-1C tiene la estructura

típicamente usada con el secuenciador Illumina, tal como se muestra en la FIG. 1C. Se usan dos cebadores que se hibridan con la parte más externa de la molécula, los cebadores Illumina P5 y P7, para la amplificación en fase sólida de la molécula (formación de grupos). Se hacen tres lecturas de secuencia por molécula. Se hace la primera lectura de 100 pb con el cebador C', que tiene una temperatura de fusión que resulta apropiada para el proceso de secuenciación Illumina. La segunda lectura tiene una longitud de sólo 6 pb y tiene únicamente el fin de identificar el marcador de la muestra. Se genera usando un cebador de marcador proporcionado por el fabricante (Illumina). La lectura final es el cebador Read 2, también proporcionado por el fabricante (Illumina). Usando este cebador, se genera una lectura de 100 pb en el segmento V partiendo de la secuencia del cebador V de la PCR de 1 etapa.

10 **Ejemplo**

En este ejemplo, se generaron perfiles de clonotipos a partir de cada muestra de ARN de muestras de sangre tomadas de pacientes con EA e individuos de control como se indica más adelante. El método de generación de perfiles de clonotipos para TCRβ era esencialmente el descrito en Faham y Willis (antes citado). Después de una transcripción inversa y de una amplificación por PCR en dos etapas como se ha descrito anteriormente, se determinaron las secuencias de los amplicones resultantes en un secuenciador de ADN Illumina GA usando los protocolos sugeridos por el fabricante. Cada perfil de clonotipos comprendía aproximadamente  $2 \times 10^5$  clonotipos construidos a partir de aproximadamente  $1,3 \times 10^6$  lecturas de secuencia generadas por el secuenciador Illumina. Se analizaron los perfiles de clonotipos para detectar clonotipos o características de clonotipos que se compartían entre números significativos de las muestras de los pacientes con EA, pero no de los controles. Se descubrió que un número significativo de pacientes con EA compartían clonotipos que codificaban los siguientes segmentos peptídicos de TCRβ: LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2).

Se analizaron los perfiles de clonotipos de muestras de pacientes de control y con EA usando técnicas de minería de datos convencionales, v.g., Witten *et al.*, Data Mining: Practical Machine Learning Tools and Techniques, Tercera Edición (Morgan Kaufman, 2011), con el objetivo de determinar si los pacientes con EA tenían clonotipos que codificaran unidades de secuencia de aminoácidos comunes. Se establecieron los conjuntos de muestras de pacientes con EA como sigue: (a) se implementó un entrenamiento sobre 56 pacientes positivos para HLA B27 (1 muestra/paciente); (b) se implementaron pruebas sobre 56 pacientes positivos para HLA B27 (1 muestra/paciente); (c) se realizó la confirmación sobre 57 muestras de 16 pacientes (12 pacientes positivos para HLA B27, 2 pacientes negativos para HLA B27 y 2 pacientes con un tipo desconocido de HLA). Se establecieron los conjuntos de muestras de control como sigue: (a) se implementó un entrenamiento sobre 521 muestras de 120 pacientes con lupus y 25 individuos normales, y (b) se realizaron pruebas sobre 56 pacientes con lupus (1 muestra/paciente, con las muestras pareadas en cuanto a los recuentos de clonotipos con las muestras de ensayo de EA). Se sacaron las muestras de ensayo y de entrenamiento de pacientes con lupus del mismo conjunto de muestras, pero no contenían pacientes solapantes. El procedimiento de entrenamiento examinó una secuencia de 26 aminoácidos que abarcaba la región CDR3 de TCRβ para determinar las secuencias de aminoácidos compartidas (es decir, clones funcionales putativos) codificadas por clonotipos de pacientes con EA, pero no de controles. Se encontraron 374 clonotipos funcionales putativos compartidos por al menos 28 muestras de entrenamiento de EA. La búsqueda de estos clonotipos en el conjunto de entrenamiento de control encontró (a) 1 secuencia altamente específica (péptido 1) (observada en un 5% de las muestras de control y en un 12% de los individuos de control), (b) 1 secuencia moderadamente específica (péptido 2) (observada en un 15% de las muestras de control y en un 27% de los individuos de control) y (c) el resto de las secuencias fueron observadas en >18% de las muestras de control y en >37% de los individuos de control. En el conjunto de ensayo, el péptido 1 estaba presente en 21/56 muestras de ensayo de EA, frente a 4/56 muestras de control (valor de  $p < 10^{-4}$ ) y el péptido 2 estaba presente en 29/56 muestras de ensayo de EA, frente a 13/56 muestras de control (valor de  $p < 10^{-3}$ ). En el conjunto de confirmación, el péptido 1 estaba presente en 14 muestras de 6 pacientes, incluyendo 1 paciente positivo a B27, y el péptido 2 estaba presente en 36 muestras de 10 pacientes, incluyendo a ambos pacientes positivos a B27.

50 **Definiciones**

A menos que se defina aquí específicamente de otro modo, los términos y símbolos de la química de ácidos nucleicos, la bioquímica, la genética y la biología molecular usados en el presente documento siguen los de los tratados y textos estándar en el campo, v.g., Kornberg y Baker, DNA Replication, Segunda Edición (W.H. Freeman, New York, 1992); Lehninger, Biochemistry, Segunda Edición (Worth Publishers, New York, 1975); Strachan y Read, Human Molecular Genetics, Segunda Edición (Wiley-Liss, New York, 1999); Abbas *et al.*, Cellular and Molecular Immunology, 6ª edición (Saunders, 2007).

"Alineación" significa un método de comparación de una secuencia de ensayo, tal como una lectura de secuencia, con una o más secuencias de referencia para determinar qué secuencia de referencia o qué porción de una secuencia de referencia está más próximamente basada en alguna medida de distancia de secuencia. Un método ejemplar de alineación de secuencias de nucleótidos es el algoritmo Smith Waterman. Las medidas de distancia pueden incluir la distancia Hamming, la distancia Levenshtein o similares. Las medidas de distancia pueden incluir un componente relacionado con los valores de calidad de los nucleótidos de las secuencias que se estén comparando.

"Amplicón" significa el producto de una reacción de amplificación de polinucleótidos, es decir, una población clonal de polinucleótidos, que pueden ser de una sola hebra o de doble hebra, que se replican a partir de una o más secuencias de partida. La una o más secuencias de partida pueden ser una o más copias de la misma secuencia, o pueden ser una mezcla de diferentes secuencias. Preferiblemente, los amplicones se forman por la amplificación de una sola secuencia de partida. Los amplicones pueden ser producidos por una variedad de reacciones de amplificación, cuyos productos incluyen réplicas del uno o más ácidos nucleicos de partida, o diana. En un aspecto, las reacciones de amplificación que producen amplicones están "dirigidas por plantilla" en el sentido de que el emparejamiento de bases de los reactivos, ya sean nucleótidos u oligonucleótidos tiene complementarios en un polinucleótido de plantilla que son necesarios para la creación de productos de reacción. En un aspecto, las reacciones dirigidas por plantilla son extensiones de cebadores con una ácido nucleico polimerasa o ligaciones de oligonucleótidos con una ácido nucleico ligasa. Dichas reacciones incluyen, aunque sin limitación, reacciones en cadena de polimerasa (PCR), reacciones de polimerasa lineal, amplificación basada en secuencias de ácidos nucleicos (NASBA), amplificaciones en círculo rodante y similares, desveladas en las siguientes referencias: Mullis *et al.*, patentes estadounidenses 4.683.195, 4.965.188, 4.683.202, 4.800.159 (PCR), Gelfand *et al.*, patente estadounidense 5.210.015 (PCR en tiempo real con sondas "taqman"), Wittwer *et al.*, patente estadounidense 6.174.670, Kacian *et al.*, patente estadounidense 5.399.491 ("NASBA"), Lizardi, patente estadounidense 5.854.033, Aono *et al.*, publicación de patente japonesa JP 4-262799 (amplificación en círculo rodante), y similares. En un aspecto, los amplicones de la invención son producidos por PCR. Una reacción de amplificación puede ser una amplificación "en tiempo real" si se dispone de una química de detección que permita medir un producto de reacción a medida que la reacción de amplificación progresa, *v.g.*, "PCR en tiempo real", descrita más adelante, o "NASBA en tiempo real", como se describe en Leone *et al.*, *Nucleic Acids Research*, 26: 2150-2155 (1998), y referencias similares. Tal como se utiliza aquí, el término "amplificar" significa realizar una reacción de amplificación. Una "mezcla de reacción" significa una solución que contiene todos los reactivos necesarios para llevar a cabo una reacción, los cuales pueden incluir, aunque sin limitación, agentes tamponantes para mantener el pH a un nivel seleccionado durante una reacción, sales, cofactores, depuradores y similares.

"Anticuerpo" o "inmunoglobulina" significa una proteína, ya sea natural o sintéticamente producida por medios recombinantes o químicos, que es capaz de unirse específicamente a un antígeno o determinante antigénico particular, que puede ser una molécula diana según se emplea aquí el término. Los anticuerpos, *v.g.*, los anticuerpos IgG, son normalmente glicoproteínas heterotetraméricas de aproximadamente 150.000 daltons, compuestas por dos cadenas ligeras (L) idénticas y dos cadenas pesadas (H) idénticas. Cada cadena ligera se une a una cadena pesada por un enlace disulfuro covalente, mientras que el número de uniones disulfuro varía entre las cadenas pesadas de diferentes isotipos de inmunoglobulinas. Cada cadena pesada y ligera tiene también puentes disulfuro intracatenarios regularmente espaciados. Cada cadena pesada tiene en un extremo un dominio variable ( $V_H$ ) seguido de una serie de dominios constantes. Cada cadena ligera tiene un dominio variable en un extremo ( $V_L$ ) y un dominio constante en su otro extremo; el dominio constante de la cadena ligera está alineado con el primer dominio constante de la cadena pesada, y el dominio variable de la cadena ligera está alineado con el dominio variable de la cadena pesada. Típicamente, las características de unión, *v.g.*, especificidad, afinidad y similares, de un anticuerpo, o un compuesto de unión derivado de un anticuerpo, son determinadas por los residuos de aminoácidos en las regiones  $V_H$  y  $V_L$ , y especialmente en las regiones CDR. Los dominios constantes no están directamente implicados en la unión de un anticuerpo a un antígeno. Dependiendo de la secuencia de aminoácidos del dominio constante de sus cadenas pesadas, las inmunoglobulinas pueden ser asignadas a diferentes clases. Existen cinco clases principales de inmunoglobulinas: IgA, IgD, IgE, IgG e IgM, y varias de éstas pueden dividirse aún en subclases (isotipos), *v.g.*, IgG, IgG<sub>2</sub>, IgG<sub>3</sub>, IgG<sub>4</sub>, IgA<sub>1</sub> e IgA<sub>2</sub>. "Fragmento de anticuerpo", y todas sus variantes gramaticales, tal como se usan aquí, se definen como una porción de un anticuerpo intacto que comprende el sitio de unión a antígeno o la región variable del anticuerpo intacto, donde la porción está libre de los dominios constantes de las cadenas pesadas (es decir, CH<sub>2</sub>, CH<sub>3</sub> y CH<sub>4</sub>, dependiendo del isotipo de anticuerpo) de la región Fc del anticuerpo intacto. Como ejemplos de fragmentos de anticuerpo, se incluyen fragmentos Fab, Fab', Fab'-SH, F(ab')<sub>2</sub> y Fv; diacuerpos; cualquier fragmento de anticuerpo que sea un polipéptido que tenga una estructura primaria consistente en una secuencia ininterrumpida de residuos de aminoácidos contiguos (al que se hace aquí referencia como un "fragmento de anticuerpo de una sola cadena" o "polipéptido de una sola cadena"), incluyendo, sin limitación, (1) moléculas Fv de una sola cadena (scFv), (2) polipéptidos de una sola cadena que contienen sólo un dominio variable de cadena ligera, o un fragmento de los mismos que contiene las tres CDR del dominio variable de cadena ligera, sin un resto de cadena pesada asociado, y (3) polipéptidos de una sola cadena que contienen sólo una región variable de cadena pesada, o un fragmento de los mismos que contiene las tres CDR de la región variable de cadena pesada, sin un resto de cadena ligera asociado, y estructuras multispecíficas o multivalentes formadas a partir de fragmentos de anticuerpos. El término "anticuerpo monoclonal" (mAb), tal como se usa aquí, se refiere a un anticuerpo obtenido de una población de anticuerpos substancialmente homogéneos, es decir, que los anticuerpos individuales que constituyen la población son idénticos excepto por posibles mutaciones naturales que puedan estar presentes en cantidades menores. Los anticuerpos monoclonales son altamente específicos y se dirigen contra un solo sitio antigénico. Además, contrariamente a las preparaciones de anticuerpos convencionales (policlonales), que típicamente incluyen diferentes anticuerpos dirigidos contra diferentes determinantes (epítomos), cada mAb se dirige contra un solo determinante sobre el antígeno. Además de su especificidad, los anticuerpos monoclonales son ventajosos en el sentido de que pueden ser sintetizados por cultivo de hibridomas o por sistemas de expresión bacterianos, de levaduras o de mamíferos, sin contaminar por otras inmunoglobulinas. Un anticuerpo "aislado" es uno que ha sido identificado y separado y/o recuperado de un componente de su ambiente natural. Los

componentes contaminantes de su ambiente natural son materiales que interferirían con los usos diagnósticos o terapéuticos para el anticuerpo, y pueden incluir enzimas, hormonas y otros solutos proteicos o no proteicos. En realizaciones preferidas, el anticuerpo será purificado (1) hasta más de un 95% en peso de anticuerpo, según se determina por el método Lowry, y más preferiblemente hasta más de un 99% en peso, (2) en un grado suficiente para obtener al menos 15 residuos de secuencia de aminoácidos N-terminales o internos mediante el uso de un secuenciador de copa giratoria, o (3) hasta la homogeneidad por SDS-PAGE en condiciones reductoras o no reductoras usando azul de Coomassie o, preferiblemente, tinción de plata. El anticuerpo aislado incluye el anticuerpo *in situ* dentro de células recombinantes, ya que al menos un componente del ambiente natural del anticuerpo no estará presente. De manera ordinaria, sin embargo, el anticuerpo aislado será preparado por al menos una etapa de purificación.

"Clonalidad", tal como se usa aquí, significa una medida del grado al cual se desvía la distribución de las abundancias de clonotipos entre los clonotipos de un repertorio hacia un solo o unos cuantos clonotipos. A grosso modo, la clonalidad es una medida inversa de la diversidad de clonotipos. Se dispone de muchas medidas estadísticas en ecología que describen las relaciones especie-abundancia que pueden ser empleadas para las medidas de clonalidad según la invención, *v.g.*, Capítulos 17 y 18, en Pielou, *An Introduction to Mathematical Ecology*, (Wiley-Interscience, 1969). En un aspecto, una medida de clonalidad usada en la invención es una función de un perfil de clonotipos (es decir, el número de distintos clonotipos detectados y sus abundancias), de tal forma que, después de medir un perfil de clonotipos, se puede computar la clonalidad a partir de él para obtener un solo número. Una medida de clonalidad es la medida de Simpson, que es simplemente la probabilidad de que dos clonotipos sacados al azar sean el mismo. Otras medidas de clonalidad incluyen medidas basadas en información y el índice de diversidad de McIntosh, desvelado en Pielou (antes citado).

"Clonotipo" significa una secuencia nucleotídica recombinada de una célula T que codifica un receptor de las células T (TCR), o una porción del mismo. En un aspecto, una colección de todos los distintos clonotipos de una población de linfocitos de un individuo es un repertorio de dicha población, *v.g.*, Arstila *et al.*, *Science*, 286: 958-961 (1999); Yassai *et al.*, *Immunogenetics*, 61: 493-502 (2009); Kedzierska *et al.*, *Mol. Immunol.*, 45(3): 607-618 (2008); y similares. Tal como se usa aquí, "perfil de clonotipos" o "perfil de repertorio" es una tabulación de los clonotipos de una muestra de células T (tal como una muestra de sangre periférica que contiene dichas células) que incluye substancialmente todos los clonotipos del repertorio y sus abundancias relativas. En un aspecto de la invención, un clonotipo comprende un ácido nucleico que codifica una porción de una cadena TCR $\beta$ .

"Coalescer" significa tratar dos clonotipos candidatos con diferencias de secuencia como si fueran el mismo determinando que dichas diferencias son debidas a un error experimental o de medición y no a diferencias biológicas genuinas. En un aspecto, se compara una secuencia de un clonotipo candidato de mayor frecuencia con la de un clonotipo candidato de menor frecuencia y, si se satisfacen criterios predeterminados, entonces se añade el número de clonotipos candidatos de menor frecuencia al del clonotipo candidato de mayor frecuencia y se desecha a continuación el clonotipo candidato de menor frecuencia. Es decir, se añaden los recuentos de lectura asociados al clonotipo candidato de menor frecuencia a los del clonotipo candidato de mayor frecuencia.

"Regiones determinantes de complementariedad" (CDR) significa regiones de una inmunoglobulina (es decir, anticuerpo) o receptor de células T donde la molécula complementa la conformación de un antígeno, determinando así la especificidad de la molécula y el contacto con un antígeno específico. Los receptores de las células T y las inmunoglobulinas tienen cada uno tres CDR: CDR1 y CDR2 se encuentran en el dominio variable (V) y CDR3 incluye algunos dominios de V, todos los diversos (D) (sólo cadenas pesadas) y de la unión (J) y algunos de los constantes (C).

"Cantidad efectiva" significa una cantidad suficiente para mejorar un síntoma de una afección autoinmune. La cantidad efectiva para un paciente particular puede variar dependiendo de factores tales como el estado de la afección autoinmune que se esté tratando, la salud general del paciente, el método de administración, la severidad de los efectos colaterales y similares. En general, se administra un anticuerpo terapéutico específico para un péptido relacionado con la EA como una composición farmacéutica que tiene una cantidad efectiva de dicho anticuerpo y un soporte farmacéutico. Un soporte farmacéutico puede ser cualquier sustancia no tóxica compatible adecuada para administrar las composiciones de la invención a un paciente. En general, las composiciones útiles para administración parenteral de dichos fármacos son bien conocidas, *v.g.*, Remington's *Pharmaceutical Science*, 15<sup>a</sup> Ed. (Mack Publishing Company, Easton, Pa. 1980). De manera alternativa, se pueden introducir las composiciones de la invención en el organismo de un paciente mediante un sistema de administración de fármacos implantable o inyectable, *v.g.*, Urquhart *et al.*, *Ann. Rev. Pharmacol. Toxicol.*, Vol. 24, págs. 199-236 (1984); Lewis, ed. *Controlled Release of Pesticides and Pharmaceuticals* (Plenum Press, New York, 1981); Patente Estadounidense N° 3.773.919; Patente Estadounidense N° 3.270.960; y similares.

"Por ciento homólogo", "por ciento idéntico" o términos similares usados en relación a la comparación de una secuencia de referencia y otra secuencia ("secuencia de comparación") significan que en una alineación óptima entre las dos secuencias, la secuencia de comparación es idéntica a la secuencia de referencia en un número de posiciones subunitarias equivalentes al porcentaje indicado, siendo las subunidades nucleotídicas para las comparaciones de polinucleótidos o aminoácidos para las comparaciones de polipéptidos. Tal como se usa aquí,

una "alineación óptima" de secuencias que están siendo comparadas es una que maximice los emparejamientos entre subunidades y minimice el número de huecos empleados en la construcción de una alineación. Se puede determinar el porcentaje de identidades con implementaciones comerciales de algoritmos, como el descrito por Needleman y Wunsch, *J. Mol. Biol.*, 48: 443-453 (1970) ("GAP" program of Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, WI), o similar. Otros paquetes de programas en la técnica para construir alineaciones y calcular el porcentaje de identidad u otras medidas de similitud incluyen el programa "BestFit", basado en el algoritmo de Smith y Waterman, *Advances in Applied Mathematics*, 2: 482-489 (1981) (Wisconsin Sequence Analysis Package, Genetics Computer Group, Madison, WI). En otras palabras, por ejemplo, para obtener un polinucleótido que tenga una secuencia de nucleótidos al menos un 95 por ciento idéntica a una secuencia de nucleótidos de referencia, se pueden eliminar hasta un cinco por ciento de los nucleótidos en la secuencia de referencia o substituirlos con otro nucleótido, o se pueden insertar un número de nucleótidos hasta un cinco por ciento del número total de nucleótidos en la secuencia de referencia en la secuencia de referencia.

La "presentación en fagos" es una técnica mediante la cual se presentan polipéptidos variantes como proteínas de fusión a al menos una porción de una proteína de cubierta sobre la superficie de fagos, *v.g.*, fagos filamentosos, partículas. Una utilidad de la presentación en fagos reside en el hecho de que se pueden seleccionar rápida y eficientemente grandes bibliotecas de variantes proteicas aleatorizadas en cuanto a aquellas secuencias que se unen a una molécula diana con gran afinidad. Se ha usado la presentación de bibliotecas de péptidos y proteínas en fagos para cribar millones de polipéptidos para encontrar los que tienen propiedades de unión específica. Se han empleado métodos de presentación en fagos polivalentes para presentar pequeños péptidos aleatorios y pequeñas proteínas a través de fusiones al gen III o al gen VIII de fagos filamentosos. Wells y Lowman, *Curr. Opin. Struct. Biol.*, 3: 355-362 (1992), y las referencias allí citadas. En la presentación en fagos monovalente, se fusiona una biblioteca de proteínas o de péptidos con un gen III o una porción del mismo y se expresa a bajos niveles en presencia de la proteína del gen III de tipo salvaje, de tal forma que las partículas de fago presentan una copia o ninguna de las proteínas de fusión. Se reducen los efectos de avidéz en relación a los fagos polivalentes, de tal modo que la selección se hace en base a la afinidad intrínseca por el ligando, y se usan factores fagémidos, que simplifican las manipulaciones del ADN. Lowman y Wells, *Methods: A companion to Methods in Enzymology*, 3: 205-0216 (1991).

"Reacción en cadena de polimerasa" o "PCR" significa una reacción para la amplificación *in vitro* de secuencias específicas de ADN mediante la extensión simultánea con cebadores de hebras complementarias de ADN. En otras palabras, la PCR es una reacción para hacer múltiples copias o réplicas de un ácido nucleico diana flanqueado por sitios de unión al cebador, consistiendo dicha reacción en una o más repeticiones de las siguientes etapas: (i) desnaturalización del ácido nucleico diana, (ii) hibridación de cebadores con los sitios de unión a cebadores y (iii) extensión de los cebadores mediante una ácido nucleico polimerasa en presencia de nucleósido trifosfatos. Normalmente, la reacción es ciclada a través de diferentes temperaturas optimizadas para cada etapa en un instrumento ciclador térmico. Las temperaturas particulares, las duraciones en cada etapa y las velocidades de cambio entre etapas dependen de muchos factores bien conocidos para quienes tienen conocimientos ordinarios en la técnica, *v.g.*, ejemplificados por las referencias: McPherson *et al.*, editores, *PCR: A Practical Approach and PCR2: A Practical Approach* (IRL Press, Oxford, 1991 y 1995, respectivamente). Por ejemplo, en una PCR convencional que utiliza ADN polimerasa Taq, se puede desnaturalizar un ácido nucleico diana de doble hélice a una temperatura >90°C, hibridar los cebadores a una temperatura de 50-75°C y extender los cebadores a una temperatura de 72-78°C. El término "PCR" abarca formas derivadas de la reacción, incluyendo, aunque sin limitación, RT-PCR, PCR en tiempo real, PCR anidada, PCR cuantitativa, PCR multiplexada y similares. Los volúmenes de reacción varían de unos cuantos centenares de nanolitros, *v.g.*, 200 nl, a unos cuantos centenares de  $\mu$ l, *v.g.*, 200  $\mu$ l. "PCR de transcripción inversa" o "RT-PCR" significa una PCR que va precedida de una reacción de transcripción inversa que convierte un ARN diana en un ADN complementario de una sola hebra, el cual es entonces amplificado, *v.g.*, Tecott *et al.*, patente estadounidense 5.168.038. "PCR en tiempo real" significa una PCR para la que se monitoriza la cantidad de producto de reacción, es decir, de amplicón, a medida que la reacción procede. Existen muchas formas de PCR en tiempo real que difieren principalmente en las químicas de detección usadas para monitorizar el producto de reacción, *v.g.*, Gelfand *et al.*, patente estadounidense 5.210.015 ("taqman"); Wittwer *et al.*, patentes estadounidenses 6.174.670 y 6.569.627 (tintes intercalantes); Tyagi *et al.*, patente estadounidense 5.925.517 (balizas moleculares). Las químicas de detección para la PCR en tiempo real están revisadas en Mackay *et al.*, *Nucleic Acids Research*, 30: 1292-1305 (2002). "PCR anidada" significa una PCR en dos etapas donde el amplicón de una primera PCR se convierte en la muestra para una segunda PCR usando un nuevo conjunto de cebadores, al menos uno de los cuales se une a una localización interior del primer amplicón. Tal como se utiliza aquí, "cebadores iniciales" en relación a una reacción de amplificación anidada significa los cebadores usados para generar un primer amplicón, y "cebadores secundarios" significa el uno o más cebadores usados para generar un segundo amplicón, o amplicón anidado. "PCR multiplexada" significa una PCR donde se llevan a cabo múltiples secuencias diana (o una sola secuencia diana y una o más secuencias de referencia) simultáneamente en la misma mezcla de reacción, *v.g.*, Bernard *et al.*, *Anal. Biochem.*, 273: 221-228 (1999) (PCR en tiempo real de dos colores). Normalmente, se emplean distintos conjuntos de cebadores para cada secuencia que se esté amplificando. Típicamente, el número de secuencias diana en una PCR múltiple es de 2 a 50, o de 2 a 40, o de 2 a 30. "PCR cuantitativa" significa una PCR diseñada para medir la abundancia de una o más secuencias diana específicas en una muestra o espécimen. La PCR cuantitativa incluye tanto la cuantificación absoluta como la cuantificación relativa de dichas secuencias diana. Se hacen mediciones cuantitativas usando una o más secuencias de referencia o patrones internos que pueden ser

estudiados por separado o conjuntamente con una secuencia diana. La secuencia de referencia puede ser endógena o exógena a una muestra o espécimen, y en este último caso puede incluir una o más plantillas competidoras. Como secuencias de referencia endógenas típicas, se incluyen segmentos de transcritos de los siguientes genes:  $\beta$ -actina, GAPDH,  $\beta_2$ -microglobulina, ARN ribosomal y similares. Las técnicas para la PCR cuantitativa son bien conocidas para quienes tienen conocimientos ordinarios en la técnica, según se ejemplifica en las siguientes referencias: Freeman *et al.*, *Biotechniques*, 26: 112-126 (1999); Becker-Andre *et al.*, *Nucleic Acids Research*, 17: 9437-9447 (1989); Zimmerman *et al.*, *Biotechniques*, 21: 268-279 (1996); Diviacco *et al.*, *Gene*, 122: 3013-3020 (1992); Becker-Andre *et al.*, *Nucleic Acids Research*, 17: 9437-9446 (1989); y similares.

"Cebador" significa un oligonucleótido, ya sea natural o sintético, que es capaz, al formar un dúplex con una plantilla de polinucleótido, de actuar como un punto de iniciación de la síntesis de ácidos nucleicos y de extenderse desde su extremo 3' a lo largo de la plantilla, de tal modo que se forma un dúplex extendido. La extensión de un cebador es normalmente llevada a cabo con una ácido nucleico polimerasa, tal como una ADN o ARN polimerasa. Se determina la secuencia de nucleótidos añadidos en el proceso de extensión por la secuencia del polinucleótido de plantilla. Normalmente, los cebadores son extendidos por una ADN polimerasa. Los cebadores normalmente tienen una longitud de 14 a 40 nucleótidos, o de 18 a 36 nucleótidos. Los cebadores son empleados en una variedad de reacciones de amplificación nucleica, por ejemplo, reacciones de amplificación lineal usando un solo cebador, o reacciones en cadena de polimerasa, empleando dos o más cebadores. Las directrices para seleccionar las longitudes y las secuencias de cebadores para aplicaciones particulares son bien conocidas para quienes tienen conocimientos ordinarios en la técnica, como evidencian las siguientes referencias: Dieffenbach, editor, *PCR Primer: A Laboratory Manual*, 2ª Edición (Cold Spring Harbor Press, New York, 2003).

"Puntuación de calidad" significa una medida de la probabilidad de que una asignación de bases en una localización de secuencia particular sea correcta. Una variedad de métodos son bien conocidos para quienes tienen conocimientos ordinarios para el cálculo de las puntuaciones de calidad para circunstancias particulares, tales como para bases identificadas como resultado de diferentes químicas de secuenciación, sistemas de detección, algoritmos de identificación de bases, etc. En general, los valores de las puntuaciones de calidad se relacionan monotónicamente con las probabilidades de una correcta identificación de bases. Por ejemplo, una puntuación de calidad, o Q, de 10 puede significar que hay un 90 por ciento de probabilidad de que una base sea identificada correctamente, una Q de 20 puede significar que hay un 99 por ciento de probabilidad de que una base sea identificada correctamente, etc. Para algunas plataformas de secuenciación, particularmente para las que usan químicas de secuenciación por síntesis, las puntuaciones medias de calidad disminuyen en función de la longitud de lectura de secuencia, de tal modo que las puntuaciones de calidad al comienzo de una lectura de secuencia son mayores que las del final de una lectura de secuencia, siendo debidas dichas disminuciones a fenómenos tales como extensiones incompletas, extensiones hacia delante, pérdida de plantilla, pérdida de polimerasa, fallos de terminación, fallos de desprotección y similares.

"Repertorio" o "repertorio inmune" o "repertorio de receptores inmunes" significa un conjunto de distintas secuencias nucleotídicas recombinadas, o clonotipos, que codifican receptores de las células T (TCR) o fragmentos de los mismos en una población de linfocitos de un individuo. Las poblaciones de linfocitos de los que se determina un repertorio pueden ser tomadas de diferentes muestras de tejidos, para producir diferentes repertorios inmunes. En algunos aspectos de la invención, la población de linfocitos correspondientes a un repertorio puede ser de células T circulantes, o puede tratarse de subpoblaciones de las anteriores poblaciones, incluyendo, aunque sin limitación, células T CD4+, o células T CD8+, u otras subpoblaciones definidas por marcadores de la superficie celular, o similares. Dichas subpoblaciones pueden ser adquiridas tomando muestras de tejidos particulares, v.g., médula ósea o ganglios linfáticos o similares, o clasificando o enriqueciendo las células de una muestra (tal como sangre periférica) en base a uno o más marcadores de la superficie celular, al tamaño, a la morfología o similares. En aún otros aspectos, la población de linfocitos correspondientes a un repertorio puede derivar de tejidos enfermos, tales como un tejido tumoral, un tejido infectado o similares. En una realización, un repertorio que comprende cadenas TCR $\beta$  humanas o fragmentos de las mismas comprende un número de distintas secuencias nucleotídicas en el rango de  $0,1 \times 10^6$  a  $1,8 \times 10^6$ , o en el rango de  $0,5 \times 10^6$  a  $1,5 \times 10^6$ , o en el rango de  $0,8 \times 10^6$  a  $1,2 \times 10^6$ . En una realización particular, un repertorio de la invención comprende un conjunto de secuencias nucleotídicas codificantes substancialmente de todos los segmentos de la región V(D)J de la cadena TCR $\beta$ . En un aspecto, "substancialmente todos", tal como se usa aquí, significa cada segmento que tiene una abundancia relativa del 0,001 por ciento o superior; o en otro aspecto, "substancialmente todos", tal como se usa aquí, significa cada segmento que tiene una abundancia relativa del 0,0001 por ciento o superior. En otra realización, un repertorio de la invención comprende un conjunto de secuencias nucleotídicas que tienen longitudes en el rango de 25-200 nucleótidos y que incluyen segmentos de las regiones V, D y J de una cadena TCR $\beta$ . En otra realización, un repertorio de la invención comprende un número de distintas secuencias nucleotídicas que es substancialmente equivalente al número de linfocitos que expresan una cadena TCR $\beta$  distinta. En aún otra realización, "substancialmente equivalentes" significa que, con una probabilidad del noventa y nueve por ciento, un repertorio de secuencias nucleotídicas incluirá una secuencia nucleotídica codificante de una TCR $\beta$  o porción de la misma llevada o expresada por cada linfocito de una población de un individuo a una frecuencia del 0,001 por ciento o superior. En aún otra realización, "substancialmente equivalentes" significa que, con una probabilidad del noventa y nueve por ciento, un repertorio de secuencias nucleotídicas incluirá una secuencia nucleotídica codificante de una TCR $\beta$  o porción de la misma llevada



o expresada por cada linfocito presente a una frecuencia del 0,0001 por ciento o superior. Se hace aquí a veces referencia a los conjuntos de clonotipos descritos en las dos oraciones anteriores como representantes de "todo el repertorio" de secuencias TCR $\beta$ . Tal como se ha mencionado anteriormente, cuando se mide o genera un perfil de clonotipos (o perfil de repertorio), se obtiene una muestra de linfocitos lo suficientemente grande como para que dicho perfil proporcione una representación razonablemente precisa de un repertorio para una aplicación particular. En un aspecto, se emplean muestras que contienen de 10<sup>5</sup> a 10<sup>7</sup> linfocitos, especialmente cuando se obtienen de muestras de sangre periférica de 1-10 ml.

"Lectura de secuencias" significa una secuencia de nucleótidos determinada a partir de una secuencia o de un torrente de datos generados mediante una técnica de secuenciación, donde dicha determinación es realizada, por ejemplo, por medio de un programa de determinación de bases asociado a la técnica, v.g., programa de determinación de bases de un proveedor comercial de una plataforma de secuenciación de ADN. Una lectura de secuencia normalmente incluye puntuaciones de calidad para cada nucleótido de la secuencia. Típicamente, se hacen las lecturas de secuencia extendiendo un cebador a lo largo de un ácido nucleico de plantilla, v.g., con una ADN polimerasa o una ADN ligasa. Se generan los datos registrando las señales, tales como señales ópticas, químicas (v.g. cambio de pH) o eléctricas, asociadas a dicha extensión. Se convierten dichos datos iniciales en una lectura de secuencia. Típicamente, se genera un clonotipo por coalescencia de múltiples lecturas de secuencia.

"Árbol de secuencias" significa una estructura de datos arborea para representar secuencias nucleotídicas. En un aspecto, una estructura de datos arborea de la invención es un árbol dirigido con raíces que comprende nodulos y bordes que no incluyen ciclos, o rutas cíclicas. Los bordes de los nodulos de las estructuras de datos arboreas de la invención están normalmente ordenados. Los nodulos y/o los bordes son estructuras que pueden contener, o asociarse a, un valor. Cada nódulo en un árbol tiene cero o más nodulos hijos, los cuales son mostrados por convención por debajo de él en el árbol. Un nódulo que tiene un hijo es denominado el nódulo parental del hijo. Un nódulo que no tiene un hijo es denominado el nódulo hoja. El nódulo que se encuentra más arriba en un árbol es denominado el nódulo raíz. Al ser el nódulo que se encuentra más arriba, el nódulo raíz no tendrá parentales. Es el nódulo en el que comúnmente comienzan las operaciones en el árbol (aunque algunos algoritmos comienzan con los nodulos hoja y trabajan acabando en el raíz). El resto de los nodulos pueden ser alcanzados desde él siguiendo los bordes o enlaces.

**Lista de secuencias**

<110> Sequenta, Inc.  
 Faham, Malek  
 Carlton, Victoria  
 Moorhead, Martin  
 Zheng, Jianbiao  
 Asbury, Thomas

<120> Clonotipos de receptores de las células T compartidos entre pacientes con espondilitis anquilosante

<130> 820WO00

<150> 61/556125  
 <151> 04-11-2011

<150> 61/561234  
 <151> 17-11-2011

<160> 2

<170> PatentIn version 3.5

<210> 1  
 <211> 26  
 <212> PRT  
 <213> Homo sapiens

<400> 1

Leu Cys Ala Ser Ser Leu Glu Ala Ser Gly Ser Ser Tyr Asn Glu Gln  
 1 5 10 15

Phe Phe Gly Pro Gly Thr Arg Leu Thr Val  
 20 25

ES 2 619 713 T3

<210> 2  
<211> 26  
<212> PRT  
<213> Homo sapiens

5

<400> 2

Val Tyr Phe Cys Ala Ser Ser Asp Ser Ser Gly Ser Thr Asp Thr Gln  
1 5 10 15

Tyr Phe Gly Pro Gly Thr Arg Leu Thr Val  
20 25

## REIVINDICACIONES

1. Un método para determinar un estado de la enfermedad de un paciente que sufre, o es sospechoso de sufrir, espondilitis anquilosante, donde el método comprende las etapas de:
- 5 (a) amplificar moléculas de ácido nucleico de las células T de una muestra obtenida del paciente, teniendo las moléculas de ácido nucleico secuencias de ADN recombinadas de genes de los receptores de las células T;
- (b) secuenciar las moléculas amplificadas de ácido nucleico para formar un perfil de clonotipos;
- 10 (c) determinar a partir del perfil de clonotipos una presencia, una ausencia y/o un nivel de clonotipos codificantes de segmentos de un receptor de las células T al menos un setenta por ciento idéntico a un segmento en el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2); y
- (d) correlacionar la presencia, ausencia y/o nivel de dichos clonotipos con un estado de espondilitis anquilosante en el paciente.
- 15 2. El método de la reivindicación 1, donde dicha etapa de determinación consiste en determinar dicho nivel de dichos clonotipos codificantes de dichos segmentos de un receptor de las células T al menos un noventa por ciento idéntico a un segmento en el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2), y donde dicha etapa de correlación consiste en correlacionar una presencia o un nivel elevado de dichos clonotipos con espondilitis anquilosante en dicho paciente.
- 20 3. El método de la reivindicación 1 o de la reivindicación 2, donde dicha muestra es una muestra de sangre periférica.
- 25 4. Una medicación para mejorar los efectos de la espondilitis anquilosante para uso en el tratamiento de un paciente con un elevado nivel de clonotipos codificantes de segmentos de un receptor de las células T al menos un noventa por ciento idéntico a un segmento en el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2), donde el nivel de los clonotipos es determinado por el método de la reivindicación 2.
- 30 5. La medicación de la reivindicación 4 para uso como se define en la reivindicación 4, donde dicha medicación es seleccionada entre el grupo consistente en un fármaco antiinflamatorio, un fármaco antirreumático modificador de la enfermedad (FARME) y un bloqueante del TNF $\alpha$ .
- 35 6. Una medicación para uso en el tratamiento de un paciente que sufre espondilitis anquilosante, donde se determina el estado de la enfermedad del paciente mediante un método que comprende las etapas de: determinar en un perfil de clonotipos de una muestra de tejido del paciente la presencia, ausencia y/o cantidad de clonotipos codificantes de segmentos de un receptor de las células T al menos un noventa por ciento idéntico a un segmento en el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2); y correlacionar una presencia o una elevación en un nivel de dichos clonotipos con una presencia de espondilitis anquilosante en el paciente; y donde la medicación es seleccionada entre el grupo consistente en un fármaco antiinflamatorio, un fármaco antirreumático modificador de la enfermedad (FARME) y un bloqueante del TNF $\alpha$ .
- 40 7. La medicación de la reivindicación 6 para uso como se define en la reivindicación 6, donde dicho elevado nivel de dicho clonotipo es al menos un 0,00001 por ciento de los clonotipos en dicho perfil de clonotipos.
- 45 8. La medicación de la reivindicación 7 para uso como se define en la reivindicación 7, donde dicho elevado nivel de dicho clonotipo es al menos un 0,0001 por ciento de los clonotipos en dicho perfil de clonotipos.
- 50 9. La medicación de la reivindicación 8 para uso como se define en la reivindicación 8, donde dicho elevado nivel de dicho clonotipo es al menos un 0,001 por ciento de los clonotipos en dicho perfil de clonotipos.
- 55 10. Un anticuerpo aislado específico para un segmento de aminoácidos de un receptor de las células T, donde el segmento de aminoácidos es seleccionado entre el grupo consistente en LCASSLEASGSSYNEQFFGPGTRLTV (SEC. N° ID. 1) y cualquier segmento de 6 a 20 aminoácidos del mismo y VYFCASSDSSGSDTQYFGPGTRLTV (SEC. N° ID. 2) y cualquier segmento de 6 a 20 aminoácidos del mismo.
- 60 11. El anticuerpo de la reivindicación 10 para uso en el tratamiento de la espondilitis anquilosante.

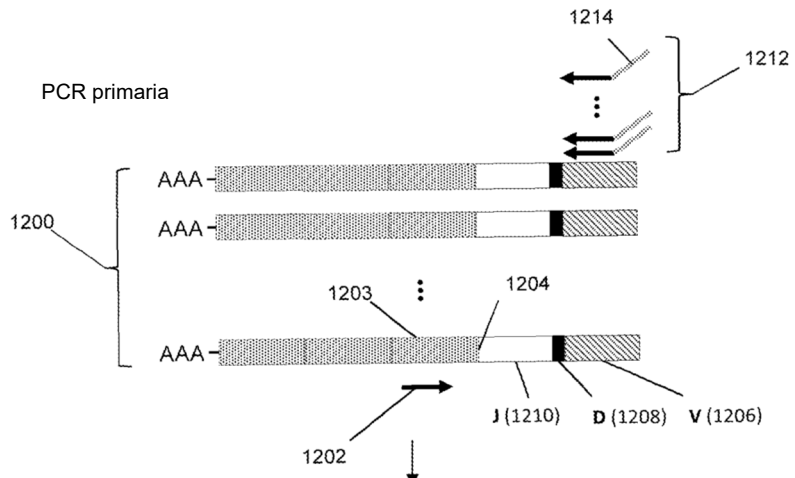


Fig. 1A

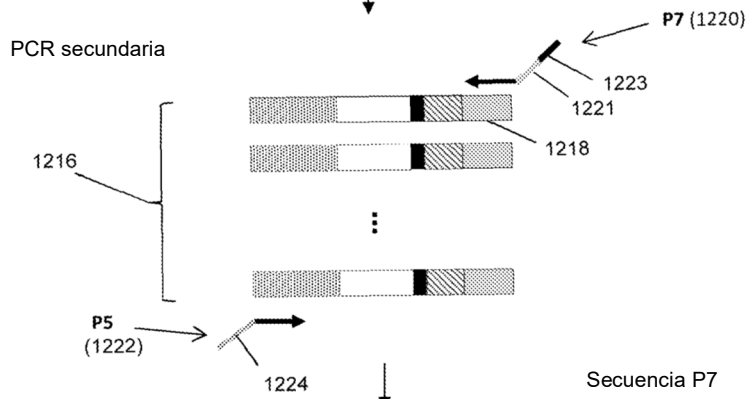


Fig. 1B

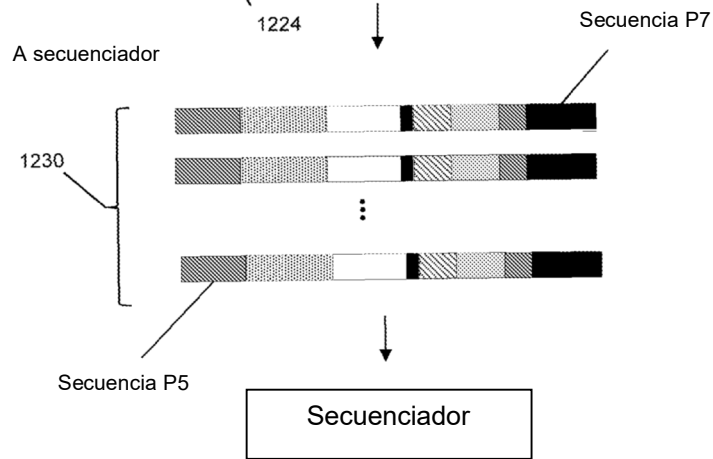


Fig. 1C