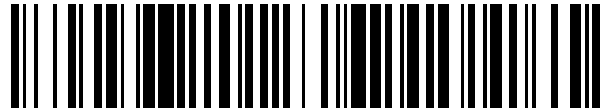


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 530 743**

51 Int. Cl.:

**C12Q 1/68**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.09.2009 E 09815707 (6)**

97 Fecha y número de publicación de la concesión europea: **19.11.2014 EP 2333110**

54 Título: **Método in vitro y kit para el pronóstico o predicción de la respuesta por parte de pacientes con artritis reumatoide al tratamiento con agentes bloqueantes del factor TNF $\alpha$**

30 Prioridad:

**24.09.2008 ES 200802713**

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

**05.03.2015**

73 Titular/es:

**FUNDACIÓ INSTITUT DE RECERCA DE  
L'HOSPITAL UNIVERSITARI VALL D'HEBRON  
(100.0%)**

**Passeig Vall d'Hebron, 119-129  
08035 Barcelona, ES**

72 Inventor/es:

**JULIÁ CANO, ANTONIO y  
MARSAL BARRIL, SARA**

74 Agente/Representante:

**ILLESCAS TABOADA, Manuel**

**Observaciones :**

**Véase nota informativa (Remarks) en el folleto original publicado por la Oficina Europea de Patentes**

**ES 2 530 743 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****MÉTODO *IN VITRO* Y KIT PARA EL PRONÓSTICO O PREDICCIÓN DE LA RESPUESTA POR PARTE DE PACIENTES CON ARTRITIS REUMATOIDE AL TRATAMIENTO CON AGENTES BLOQUEANTES DEL FACTOR TNF $\alpha$** **CAMPO DE LA INVENCIÓN**

La presente invención se refiere a un método *in vitro* (en adelante método de la invención) para el pronóstico o predicción de la respuesta al tratamiento con agentes bloqueantes del factor necrosis tumoral alfa (TNF $\alpha$ ), como por ejemplo infliximab o adalimumab, por parte de pacientes con artritis reumatoide. Así, la presente invención puede englobarse dentro del campo de la medicina personalizada, de la reumatología o de la genética humana como campo que estudia las enfermedades de base genética.

**ESTADO DE LA TÉCNICA**

La artritis reumatoide (AR) es una de las enfermedades autoinmunes más frecuentes en el mundo (prevalencia mundial ~ 1%). La AR conduce a la inflamación crónica de las articulaciones sinoviales y al desarrollo de daño articular progresivo que puede dar lugar a una marcada incapacidad funcional. Además, la AR es una enfermedad muy heterogénea y compleja en todos sus aspectos, incluyendo tanto sus manifestaciones clínicas como la variabilidad en su respuesta a las diferentes terapias.

Fruto de la intensa investigación llevada a cabo durante los últimos años, se han identificado varios tratamientos para el control de la AR. Durante la última década las terapias biológicas que inhiben el TNF $\alpha$  han tenido un mayor desarrollo, con respecto al control de la AR, en comparación con terapias basadas en DMARD (medicamentos antirreumáticos modificadores de la enfermedad).

Actualmente el infliximab es una de las terapias más usadas para el tratamiento de la AR. El infliximab es un anticuerpo monoclonal quimérico (IgG) derivado de un ADN recombinante de origen humano y murino, que se une y neutraliza al TNF $\alpha$ , logrando interrumpir la cascada secuencial de activación de las vías inflamatorias mediadas por esta citoquina.

Sin embargo, existe un porcentaje de pacientes que no responden al tratamiento con agentes bloqueantes del factor TNF $\alpha$  y deben ser redirigidos hacia terapias alternativas. Por lo tanto, los métodos para el pronóstico o predicción de los pacientes que se pueden beneficiar de este tratamiento y su diferenciación de los pacientes no respondedores a dicho tratamiento que serán focalizados hacia terapias alternativas, son una necesidad creciente dirigida hacia el reto que supone la medicina individualizada.

En el documento [*“Gene profiling in white blood cells predicts infliximab responsiveness in rheumatoid arthritis”*. Thierry Lequerré et al.] se divulga el estudio del perfil de expresión génica (mediante la cuantificación del nivel de ARNm) en células mononucleares periféricas (PBMCs) (lo que supone una fracción de menos del 60% de las células de la sangre), como base para la predicción de la respuesta a infliximab por parte de pacientes con AR. Se determinaron un grupo de genes correlacionados con la respuesta al tratamiento. Los genes estudiados se muestran en las diferentes tablas y figuras de este documento, siendo el listado más completo el comprendido en la Tabla 4.

En el documento [*“Effect of infliximab therapy on gene expression levels of tumor necrosis factor alpha, tristetraprolin, T cell intracellular antigen 1, and Hu antigen R in patients with rheumatoid arthritis”*. Sugihara, Makoto et al.] se llevó a cabo la identificación de parámetros que puedan predecir, en pacientes con AR, la eficacia de la terapia anti-TNF $\alpha$ . Para ello se analizó el nivel de expresión de los genes TNF $\alpha$ , TTP, TIA-1 y HuR concluyendo que las diferencias de expresión en genes ABP (TTP, TIA-1 y HuR) puede afectar a la expresión del gen TNF $\alpha$ . Un ratio de expresión TIA-1:HuR elevado podría estar relacionado con la respuesta del paciente al tratamiento con infliximab.

En el documento [*“Effect of infliximab on mRNA expression profiles in synovial tissue of rheumatoid arthritis patients”*. Johan Lindberg et al.] se divulga el examen realizado del perfil de expresión génica de pacientes aquejados con AR para investigar si dichos perfiles pueden ser utilizados para predecir la respuesta de los pacientes al tratamiento con infliximab. Para ello se estudió el perfil de expresión de una serie de genes como MMP-3 y otros citados en la página 5 (columna izquierda segundo párrafo). Los documentos WO2007/135568 y WO2007/038501 divulgan genes para la predicción de la respuesta a agentes bloqueantes de TNF $\alpha$  en muestras de sangre.

Ninguno de los documentos localizados en el estado de la técnica describe ninguno de los ocho genes, cuyo nivel de expresión es analizado en la presente invención: HLA-DRB3 (NCBI: 3125), EAT2 (NCBI: 117157), GNLY (NCBI: 10578), CAMP (NCBI: 820), SLC2A3 (NCBI: 6515), IL2RB (NCBI: 3560), MXD4 (NCBI: 10608) y TLR5 (NCBI: 7100). Por lo tanto, no existe en el estado de la técnica ninguna evidencia relacionada con la utilización de los ocho genes arriba citados, los cuales tal y como se explica en la descripción de la invención han sido específicamente seleccionados en la presente invención mediante un exhaustivo proceso de cribado entre miles de genes.

Es importante hacer notar que aunque en la presente invención el modelo predictor óptimo es el que considera la expresión del conjunto de los ocho genes arriba citados a la vez, cada uno de estos ocho genes tiene potencial predictivo por sí mismo. Por lo tanto en el método divulgado podría utilizarse, con suficientes visos de efectividad, cualquiera de los ocho genes arriba mencionados, cualquier sub-combinación de los ocho genes con un número de genes  $\geq 2$  o el conjunto de los ocho genes, en su totalidad.

Además, otra diferencia importante del método de la invención con respecto a los métodos divulgados en el estado de la técnica se refiere a la naturaleza de la muestra biológica usada para el análisis de la expresión génica. En la presente invención y a diferencia del estado de la técnica, se aísla y se analiza el ARNm de sangre total, sin realizar ningún fraccionamiento previo. El perfil de expresión, por tanto, incluye todas las poblaciones celulares de la sangre. Este aspecto adquiere importancia al tener en cuenta que cada tipo o población celular pueden tener asociados diferentes perfiles de expresión génica.

El tipo de muestra tomada en la presente invención (sangre total) hace que el método sea muy poco invasivo para el paciente y constituye otra diferencia respecto a algunos de los métodos conocidos donde se realizan biopsias sinoviales para la extracción de las muestras.

## DESCRIPCIÓN DE LA INVENCION

### BREVE DESCRIPCIÓN DE LA INVENCION

La presente invención se refiere a un método *in vitro* para el pronóstico o predicción de la respuesta al tratamiento con agentes bloqueantes del factor TNF $\alpha$  por parte de pacientes con artritis reumatoide, como se divulga en las reivindicaciones adjuntas 1-10. Además en el presente documento se describen dichos métodos, que comprenden la determinación en una muestra de sangre de dichos pacientes y mediante la cuantificación del nivel de expresión transcripcional (ARNm), de al menos uno de los ocho genes seleccionado del grupo: HLA-DRB3 (NCBI: 3125), EAT2 (NCBI: 117157), GNLY (NCBI: 10578), CAMP (NCBI: 820), SLC2A3 (NCBI: 6515), IL2RB (NCBI: 3560), MXD4 (NCBI: 10608) y TLR5 (NCBI: 7100); y la comparación de dicho nivel de expresión con respecto a los valores de expresión previamente obtenidos de pacientes que demostraron ser respondedores al tratamiento o no respondedores al mismo.

Además la presente divulgación se refiere a un kit para el pronóstico o predicción de la respuesta al tratamiento con agentes bloqueantes del factor TNF $\alpha$  por parte de pacientes con artritis reumatoide, que comprende al menos una sonda que hibride con al menos una de las secuencias nucleotídicas de los siguientes genes: HLA-DRB3 (NCBI: 3125), EAT2 (NCBI: 117157), GNLY (NCBI: 10578), CAMP (NCBI: 820), SLC2A3 (NCBI: 6515), IL2RB (NCBI: 3560), MXD4 (NCBI: 10608) y TLR5 (NCBI: 7100).

Las sondas comprendidas en el kit se caracterizan por las secuencias SEQ ID NO: 1-8.

Tal y como se observa en las Figuras 2-9, debe tenerse en cuenta que los genes sobreexpresados en pacientes no respondedores antes de iniciar el tratamiento son: MXD4, HLA-DRB3, IL2RB, EAT2, GNLY y los genes sobreexpresados en pacientes respondedores antes de iniciar el tratamiento son: TLR5, CAMP, SLC2A3.

Por otro lado comentar que los genes HLA-DRB3, EAT2, GNLY, CAMP, IL2RB y TLR5, están todos ellos relacionados con la respuesta inmune.

Así, el problema técnico resuelto por la presente invención se refiere a un método *in vitro* para el pronóstico o predicción de la respuesta al tratamiento con agentes bloqueantes del factor necrosis tumoral de tipo alfa TNF $\alpha$  (como por ejemplo infliximab o adalimumab) por parte de pacientes aquejados de artritis reumatoide, a partir de muestras de sangre total de dichos pacientes (i.e. comprenden el 100% de las poblaciones celulares en ella presentes y no sólo las células PBMCs), que es alternativo a los existentes en el estado de la técnica (los genes analizados son diferentes a los utilizados en los métodos divulgados), sencillo (se basa en el análisis de un pequeño número de genes) y poco invasivo para el paciente (no requiere la realización de biopsias).

La predicción de la respuesta al tratamiento por parte de pacientes aquejados de AR será útil en dos sentidos principales:

1. Predicción un paciente no respondedor al tratamiento con agentes bloqueantes del factor TNF $\alpha$ : de esta forma se podría dirigir al paciente hacia tratamientos alternativos al bloqueo del TNF $\alpha$  sin necesidad de aplicarle, como primer tratamiento, la terapia con bloqueantes del factor TNF $\alpha$ . Algunos de los posibles efectos secundarios debidos al tratamiento con bloqueantes del factor TNF $\alpha$  son los siguientes (Ledingham et al. Rheumatology 2005):
  - a. Infecciones graves excluyendo tuberculosis
  - b. Tuberculosis
  - c. Linfomas
  - d. Aparición de Lupus Eritematoso Sistémico
  - e. Insuficiencia cardíaca

- f. Desmielinización y complicaciones neurológicas
  - g. Complicaciones hematológicas
2. Predicción de un paciente respondedor al tratamiento con agentes bloqueantes del factor TNF $\alpha$ : esta herramienta es muy útil para empezar el tratamiento de forma temprana, ya que en el caso particular de la AR se ha demostrado que el tratamiento en fases precoces de la enfermedad pueden ser muy efectivas evitando el avance de las erosiones articulares.

## DESCRIPCIÓN DE LAS FIGURAS

**Figura 1.** Muestra el número de veces que cada gen es seleccionado entre el grupo de 8 genes de la invención, después de 29 rondas de LOOCV (Leave One Out Cross Validation). Se observa que el grupo de ocho genes de la invención es muy frecuentemente seleccionado, mientras que el resto de los genes parecen estar sometidos a procesos estocásticos.

**Figura 2.** Perfil de expresión del gen MXD4 en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

**Figura 3.** Perfil de expresión del gen HLA-DRB3 en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

**Figura 4.** Perfil de expresión del gen TLR5 en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

**Figura 5.** Perfil de expresión del gen IL2RB en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

**Figura 6.** Perfil de expresión del gen EAT2 en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

**Figura 7.** Perfil de expresión del gen CAMP en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

**Figura 8.** Perfil de expresión del gen SLC2A3 en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

**Figura 9.** Perfil de expresión del gen GNLY en 43 individuos analizados. Los puntos negros corresponden a pacientes no respondedores y los grises corresponden a pacientes respondedores.

Tras las observación de las Figuras 2-9, puede deducirse claramente que los genes sobreexpresados en pacientes no respondedores antes de iniciar el tratamiento son: MXD4, HLA-DRB3, IL2RB, EAT2, GNLY, y los genes sobreexpresados en pacientes respondedores antes de iniciar el tratamiento son: TLR5, CAMP, SLC2A3.

## DESCRIPCIÓN DETALLADA DE LA INVENCION

La presente invención se refiere a un método *in vitro* para el pronóstico o predicción de la respuesta al tratamiento con agentes bloqueantes del factor TNF $\alpha$  por parte de pacientes con artritis reumatoide, que comprende, en una muestra de sangre de dichos pacientes, y mediante la cuantificación del nivel de mRNA, determinar el nivel de expresión del gen GNLY y la comparación de dicho nivel de expresión con respecto a los valores de expresión previamente obtenidos de pacientes que demostraron ser respondedores al tratamiento o no respondedores al mismo.

### Selección de pacientes

Desde Enero de 2005 hasta Junio de 2007, los pacientes definidos con un índice DAS28 ("Disease Activity Score" > 3.2), y por lo tanto con AR activa, fueron seleccionados para tratamiento con infliximab.

Los criterios para la inclusión de los pacientes en la terapia fueron:

1. El cumplimiento de los criterios ACR "American College of Rheumatology" 1987 para el diagnóstico de AR.
2. Tolerancia el tratamiento con metotrexato ( $\leq$  20 mg / semana).
3. Pueden recibir prednisolona (dosis  $\leq$  10 mg / día) y tratamiento con NSAID ("Non Steroidal Anti-Inflammatory Drug").
4. Las dosis de metotrexato, glucocorticoides y NSAID deben permanecer estables durante las últimas 4 semanas previas al tratamiento.
5. No haber recibido ningún otro DMARD durante las 4 semanas anteriores a la inclusión en el estudio.
6. Deben cumplirse todos los criterios estándares definidos para la inclusión de los pacientes en el tratamiento con infliximab.

Los criterios para la exclusión de los pacientes del tratamiento son:

1. Pacientes con hepatitis B (activa o inactiva)
2. Deben cumplirse todos los criterios estándares definidos para la exclusión de los pacientes del tratamiento con infliximab.

5

### Procedimientos clínicos para la evaluación de la respuesta al infliximab

En el momento de la primera infusión (semana 0), en la segunda semana y en la semana 14 se registraron variables clínicas correspondientes a la actividad de la enfermedad. Estas variables incluyen: el número de articulaciones dolorosas, número de articulaciones tumefactas, la proteína C-reactiva (CRP), velocidad de sedimentación globular (VSG) y el HAQ ("Health Assessment Questionnaire").

10

La respuesta a la terapia con infliximab se evaluó tanto por los criterios de respuesta EULAR ("European League Against Rheumatism") como por los criterios de respuesta ACR ("American College of Rheumatology").

15

EULAR define los criterios tipo en función de dos variables:

1. La variación en la puntuación de actividad DAS28 ("Disease Activity Score").
2. El resultado inicial DAS "Disease Activity Score".

20

ACR define el nivel de respuesta sobre la base de un porcentaje de reducción en diversas variables clínicas (ACR50 es equivalente a una reducción del 50% en las variables clínicas anteriormente citadas).

En la presente invención se utilizó una respuesta binaria variable en la que los no respondedores fueron las personas que se clasificaron como no respondedores en ambos sistemas (es decir, "Mala Respuesta" en EULAR y "ACR0" en ACR). Por tanto EULAR "Moderado", EULAR "Bueno", "ACR20", "ACR50" y "ACR70" fueron incluidos en el grupo de respondedores al tratamiento.

25

### Recogida de sangre y extracción de RNA

Las muestras de sangre para la extracción de ARN fueron extraídas de los pacientes antes de iniciar el tratamiento con infliximab. Uno de los factores que influyen en el perfil de ARN extraído de células sanguíneas es la hora del día a la cual se toma la muestra. Con el fin de minimizar este fuente de variabilidad, todas las muestras de sangre se extrajeron entre 8:30-9:30 a.m.

30

En el presente estudio se utilizó el sistema "PAXGene" (PreAnalytix, Suiza) para la preservación inmediata del ARN procedente de muestras de sangre total. Este sistema utiliza un agente preservante del ARN contenido en el mismo tubo de extracción de la sangre y, por lo tanto, los perfiles de sangre obtenidos son representativos de las condiciones *in vivo*. Todos los tubos "PAXGene" se almacenaron congelados a -80 ° C hasta el aislamiento de ARN. El ARN fue extraído por medio del kit de aislamiento total "PAXGene" (Qiagen, USA). Finalmente la calidad del ARN de todas las muestras se evaluó usando el sistema de gel "BioAnalyzer" (Agilent, USA).

35

### Análisis de expresión de los genes

El análisis de la expresión de genes de todo el genoma fue llevado a cabo utilizando el sistema "Illumina Human-6 v1 Beadchip array system". Mediante el uso de la tecnología "bead-based", el microarray empleado en la presente invención evaluó la expresión de 47.000 transcritos ([www.illumina.com](http://www.illumina.com)). Los ARN de buena calidad (ratio 28S/18S > 1.8, número RNA Integrity Number >9) fueron posteriormente procesados usando dicho protocolo estándar de Illumina. Después del marcaje e hibridación de la muestra, los "arrays" fueron leídos utilizando el lector de arrays "Illumina's BeadArray Reader". Los datos fueron extraídos utilizando el programa informático "Illumina BeadStudio". El resto de pasos se realizaron utilizando el programa informático "R" (<http://cran.r-project.org/>) y su extensión para el análisis de datos genómicos en el programa informático "Bioconductor". En el momento del análisis de los datos, se puso en marcha una versión actualizada del array "IlluminaHuman-6 v2, Beadchip array". Esta nueva versión incorpora un sustancial rediseño de la secuencia de las sondas y, por lo tanto, el análisis del microarray fue restringido a la porción de las sondas que fueron consideradas "válidas" según la nueva versión del microarray (datos disponibles en [www.illumina.com](http://www.illumina.com)). Los valores de la expresión final de 17.454 transcritos fueron transformados a log2 y normalizados utilizando el método quantile normalization (Bolstad Bioinformatics '03). Antes de llevar a cabo cualquier análisis estadístico, se realizó un "clustering" jerárquico para detectar y descartar cualquier posible microarray con valores extremos ("outlier").

40

50

### Construcción del predictor, respuesta y validación

Antes de aplicar cualquier técnica de modelado, se filtró el conjunto de datos obtenidos del microarray con el fin de excluir genes no informativos. Las sondas que dieron valores de expresión de genes por debajo del quinto percentil de los valores de expresión global fueron consideradas como no expresadas y fueron descartadas (N=4,150). Las sondas con una baja variabilidad (coeficiente de variación < 0.03) también fueron eliminadas (N=14.701).

60

El proceso de construcción de un predictor de la expresión génica se puede dividir en dos grandes pasos. En primer lugar, se necesita una buena estimación de la capacidad del predictor para ser generalizado, es decir, del nivel de fiabilidad que se puede esperar del predictor en el futuro. La forma adecuada de estimación de esta medida es tener un sistema de validación independiente que nunca haya sido usado en el proceso de construcción del predictor (Ntzani Lancet 2005). En segundo lugar, se tiene que hacer una selección imparcial (no sesgada) de los mejores parámetros del modelo de clasificación. Esto generalmente se lleva a cabo mediante la realización de un análisis de validación cruzada. En este caso, es fundamental evitar el sesgo a la hora de seleccionar los genes mediante la selección del conjunto de los genes del predictor en cada una de las rondas de validación cruzada (es decir, se debe hacer una completa validación cruzada). Por último, la selección del modelo óptimo entre los aparentemente buenos modelos (es decir, aquellos con similar y bajo índice de error) debe hacerse sobre la base de una medida de confianza en lugar de hacerse de forma arbitraria.

En la presente invención se construyó y se evaluó un predictor de la respuesta al tratamiento con infliximab teniendo en cuenta estos dos pasos anteriores. Los datos originales consistentes en 36 pacientes respondedores y 7 pacientes no respondedores fueron divididos al azar en un grupo de validación y en un grupo de prueba. El grupo de validación consistió en 1/3 de los datos (2 pacientes no respondedores y 12 pacientes respondedores) y el grupo de prueba consistió en los restantes 2/3 (5 pacientes no respondedores y 24 pacientes respondedores).

En la presente invención se utilizó la biblioteca "Bioconductor MLInterfaces" ([www.bioconductor.org](http://www.bioconductor.org)) que permite la rápida comparación de las distintas selecciones de genes y clasificadores. Los métodos explorados fueron "Diagonal Linear Discriminant Analysis", "Support Vector Machines", "k-Nearest Neighbours (kNN)" y "Random Forests". Los conjuntos de genes discriminados fueron seleccionados en función de su significación estadística en un test F-ratio. Para cada método se testaron diferentes números de genes discriminados y la predicción de error mediante LOOCV ("Leave One Out Cross Validation"). De todos los métodos probados, el método kNN mostró los menores índices de error para varias combinaciones de parámetros.

Con el objetivo de seleccionar entre predictores multigénicos con similar y buen rendimiento, se calculó el valor de permutación  $P$  para la significación de cada una de las modelos. Después de un número elevado de permutaciones ( $n = 10000$ ) se obtuvo un valor de permutación  $P$  mediante el cálculo de la proporción de los estadísticos de permutación que superiores al estadístico original.

Finalmente, después de elegir el modelo de predicción óptimo, éste se aplicó al conjunto de validación para estimar la fiabilidad del predictor. Se calcularon los intervalos al 95% de confianza para la predicción de la fiabilidad mediante el percentil "bootstrap" ( $n = 10.000$  repeticiones).

### Patrones de expresión génica

Una de las consecuencias más poderosas que se pueden extraer de los datos de expresión del genoma es su potencial para clasificar subtipos de una enfermedad (van't Veer et al. Nature 2002). Calculando distancias relativas entre los individuos y/o genes, se pueden construir agrupaciones que pueden mostrar mecanismos de la enfermedad nuevos e insospechados. En la presente invención se llevó a cabo un agrupamiento jerárquico de todos los transcritos (ARNm) en pacientes con AR activa con el objetivo de buscar patrones específicos. Sin embargo, un aspecto en general pasado por alto en el grupo de análisis es la determinación de una medida de la confianza acerca de la solidez de las agrupaciones. Se puede demostrar que incluso con datos aleatorios se pueden formar agrupaciones. Por lo tanto una medida de confianza que verifique la solidez del grupo es fundamental. Una posible manera de estimar dicha estabilidad es formar subconjuntos de genes repetidamente a partir de los datos originales y llevar a cabo el método de agrupamiento. Si los mismos grupos se identifican sistemáticamente, se puede tener confianza en que estos patrones son los reales. Con el fin de aplicar este análisis en la presente invención se utilizó el paquete de software "clusterStab", el cual da una estimación de la "estabilidad" estadística para cada uno de los clusters.

## RESULTADOS

### Evaluación de la respuesta clínica

Un total de 48 pacientes fueron estudiados desde la semana 0 del tratamiento con infliximab hasta la semana 14. De éstos, 4 personas no llegaron hasta la semana 14 por atriitis séptica, infarto de miocardio, reacción a la infusión y abandono voluntario, respectivamente. Se consideró que 7 pacientes presentaron "mala" respuesta EULAR, mientras que 10 pacientes fueron clasificados como respondedores "ACR0". Los 7 pacientes EULAR no respondedores son pacientes ACR0. La evaluación detallada del resto de los pacientes ACR0 mostró un resultado favorable para la mayoría de las variables clínicas. Por lo tanto, todos ellos se incluyeron en el grupo de respondedores.

### Construcción y chequeo del microarray predictor

El análisis del agrupamiento jerárquico de los datos normalizados reveló un microarray con una expresión génica fuera de dicho agrupamiento y fue eliminado de todos los análisis. Después de la filtración de genes, un total de  $n = 3364$  genes fueron seleccionados para la construcción del predictor. Los clasificadores kNN mostraron los porcentajes de

errores más bajos (sólo 1 mala clasificación de 29) para los ocho genes de la invención utilizando un modelo con 3, 4 ó 5 vecinos más cercanos. Las frecuencias de cada uno de los genes seleccionados a través de cada LOOCV se muestran en la Figura 1. Los análisis de permutación en cada uno de los modelos identificaron al modelo 3-NN como el más significativo ( $P = 0,0001$ ). La aplicación de este modelo de predicción para la validación de los datos, clasificó de forma correcta la respuesta al infliximab en 12 de los 14 individuos (85,7% de precisión de predicción, 75-100% a 95% IC).

#### **La agrupación jerárquica**

10 La evaluación de los diferentes números de agrupaciones utilizando los coeficientes de Jaccard mostró que la mejor agrupación se realizó cuando se utilizaron 2 grupos principales. En la presente invención se llevó a cabo un análisis de la estabilidad de los grupos y descubrió que la del grupo 1 fue relativamente buena (86% de estabilidad, 1000 permutaciones), frente a un grupo 2 no estable (40% la estabilidad, 1000 permutaciones).

#### **15 Principales resultados**

Mediante el uso de los perfiles de expresión génica en muestra de sangre, se construyó un predictor multigénico robusto para la estimación de la respuesta a infliximab por parte de pacientes con AR. El predictor mostró un muy buen porcentaje de precisión de predicción (86%) en un conjunto de datos independientes. También se han identificado 20 diferencias entre los niveles de linfocitos T CD4+ CD25+ en respondedores frente a no respondedores en la semana 0 de tratamiento, observándose un aumento estadísticamente significativo en este compartimento celular en los respondedores ( $P = 0.0009$ ).

25 Los pacientes no respondedores expresan mayores niveles de MXD4, HLA-DRB3, IL2RB, EAT2 y GNLY. Estos genes probablemente podrían sugerir un incremento en la actividad de las células CD4 + CD25 +, células NK, linfocitos T CD8 + ó APC. En cambio los genes sobreexpresados en pacientes respondedores antes de iniciar el tratamiento son: TLR5, CAMP, SLC2A3.

**LISTADO DE SECUENCIAS**

<110> Fundacio Institut de Recerca de l'Hospital Universitari Vall d'Hebron

5 <120> MÉTODO *IN VITRO* Y KIT PARA EL PRONÓSTICO O PREDICCIÓN DE LA RESPUESTA POR PARTE DE PACIENTES CON ARTRITIS REUMATOIDE AL TRATAMIENTO CON AGENTES BLOQUEANTES DEL FACTOR TNFalfa

<130> VALP-07030

10 <160> 8

<170> PatentIn version 3.3

15 <210> 1  
<211> 50  
<212> DNA  
<213> Secuencia artificial

20 <223> Sonda gen HLA-DRB3

<400> 1  
tcccgcctgg ctgttattct tccacgagag agggcttct caggacctag 50

25 <210> 2  
<211> 50  
<212> DNA  
<213> Secuencia artificial

30 <223> Sonda gen EAT2

<400> 2  
gtgagattg gggcacagc aatttatggc tactattccc tgggtctggt 50

35 <210> 3  
<211> 50  
<212> DNA  
<213> Secuencia artificial

40 <223> Sonda gen GNLY

<400> 3  
ctacaggfcc cctctgagcc ctctcacctt gtcctgtgga agaagcacag 50

45 <210> 4  
<211> 50  
<212> DNA  
<213> Secuencia artificial

50 <223> Sonda gen CAMP

<400> 4  
agagtcctag tgtgtgccct accctggctc aggcctctgg gctctgagaa 50

55 <210> 5  
<211> 50  
<212> DNA  
<213> Secuencia artificial

60 <223> Sonda gen SLC2A3

65



ES 2 530 743 T3

<400> 5  
atcatgtgaa cccgggacgc aggggttgca gtgagcggag atcgcacat 50

5 <210> 6  
<211> 50  
<212> DNA  
<213> Secuencia artificial  
<223> Sonda gen IL2RB

10

<400> 6  
ctgaatgttt cagaccacaa ggggtccac acctttgctg tgtgttctgg 50

15

<210> 7  
<211> 50  
<212> DNA  
<213> Secuencia artificial

20 <223> Sonda gen MXD4

<400> 7  
ggccacctgt ccaccgtgtg ggccgtgctg tgccttatg tcattgtaat 50

25

<210> 8  
<211> 50  
<212> DNA

30 <213> Secuencia artificial  
<223> Sonda gen TLR5

<400> 8  
cagaggcttt gtacagaaac agcagtatt gaggtggcct gaggatctcc 50

35

**REIVINDICACIONES**

- 5
1. Método *in vitro* para el pronóstico o predicción de la respuesta al tratamiento con agentes bloqueantes del factor TNF $\alpha$  por parte de pacientes con artritis reumatoide, que comprende la determinación en una muestra de sangre de dichos pacientes, el nivel de expresión del gen GNLY y la comparación de dicho nivel de expresión con respecto a los valores de expresión obtenidos de pacientes respondedores y no respondedores que mostraron, respectivamente, respuesta al tratamiento y aquellos que no mostraron respuesta al mismo.
- 10
2. Método *in vitro*, según la reivindicación 1, que además comprende la determinación del nivel de expresión conjunto del grupo de genes que consiste en: HLA-DRB3, EAT2, CAMP, SLC2A3, IL2RB, MXD4 y TLR5; o el nivel de expresión de cualquier combinación de los mismos.
- 15
3. Método *in vitro*, según la reivindicación 1, que comprende la determinación del nivel de expresión conjunto del grupo de seis genes que consiste en: GNLY, HLA-DRB3, EAT2, CAMP, IL2RB y TLR5.
- 20
4. Método *in vitro*, según la reivindicación 3, que además comprende la determinación del nivel de expresión del gen: SLC2A3.
- 25
5. Método *in vitro*, según cualquiera de las reivindicaciones 3 ó 4, que además comprende la determinación del nivel de expresión del gen: MXD4.
6. Método *in vitro*, según cualquiera de las reivindicaciones anteriores, caracterizado porque los genes GNLY, MXD4, HLA-DRB3, IL2RB y EAT2 están sobreexpresados en pacientes no respondedores antes de iniciar el tratamiento y los genes TLR5, CAMP y SLC2A3 están sobreexpresados en pacientes respondedores antes de iniciar el tratamiento.
- 30
7. Método *in vitro*, según cualquiera de las reivindicaciones anteriores, caracterizado porque el nivel de expresión de los genes es determinado mediante la cuantificación del nivel de ARNm en la muestra de sangre.
- 35
8. Método *in vitro*, según la reivindicación 1, caracterizado porque el agente bloqueante del factor TNF $\alpha$  es un anticuerpo anti-TNF $\alpha$ .
9. Método *in vitro*, según la reivindicación 8, caracterizado porque el agente bloqueante del factor TNF $\alpha$  es infliximab.
10. Método *in vitro*, según la reivindicación 8, caracterizado porque el agente bloqueante del factor TNF $\alpha$  es adalimumab.

FIG. 1

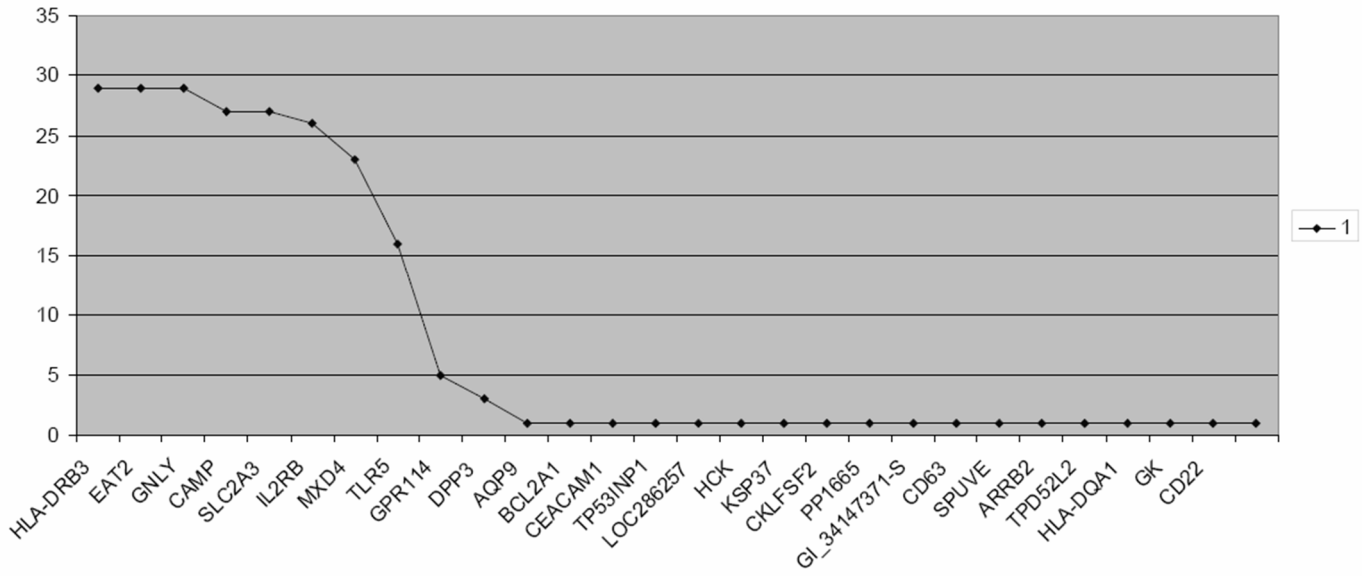


FIG. 2

MXD4

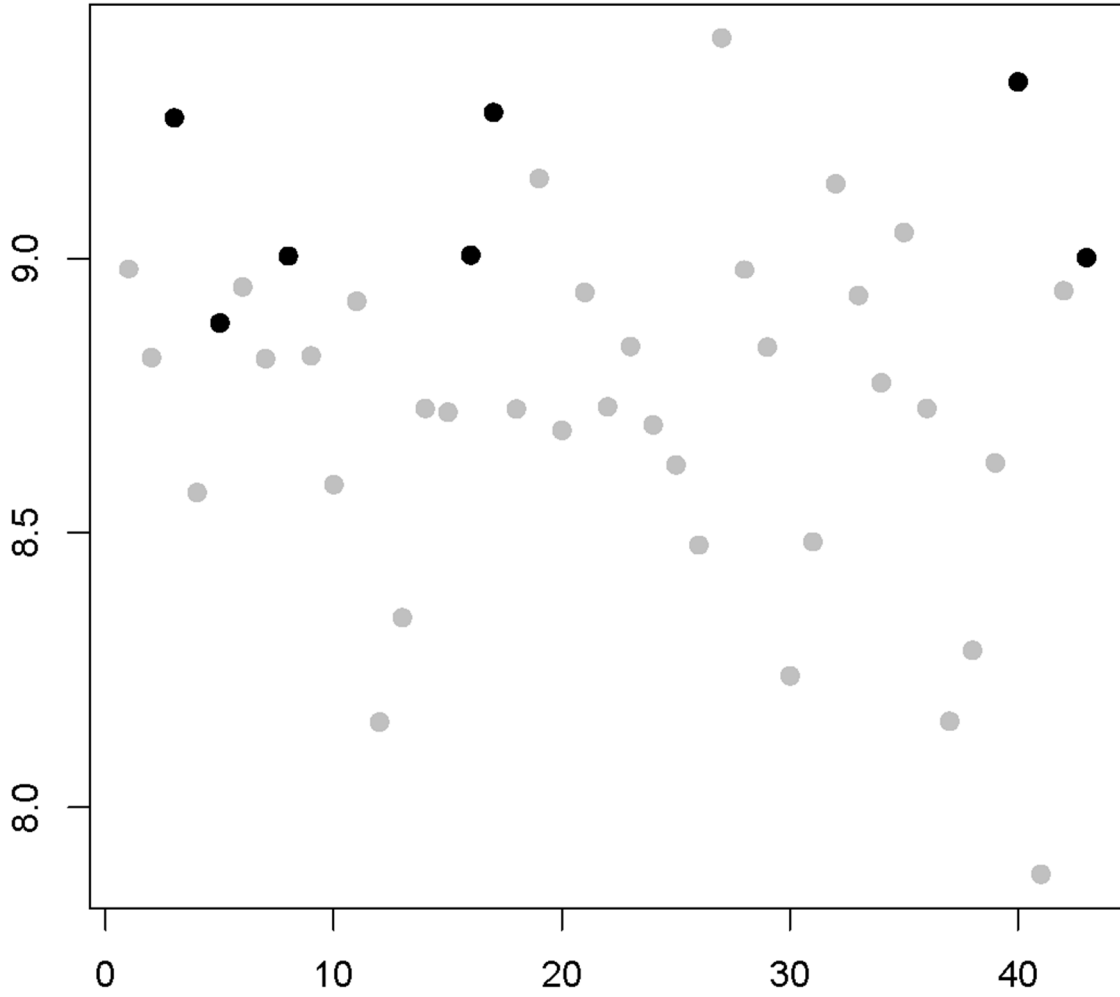


FIG. 3

HLA-DRB3

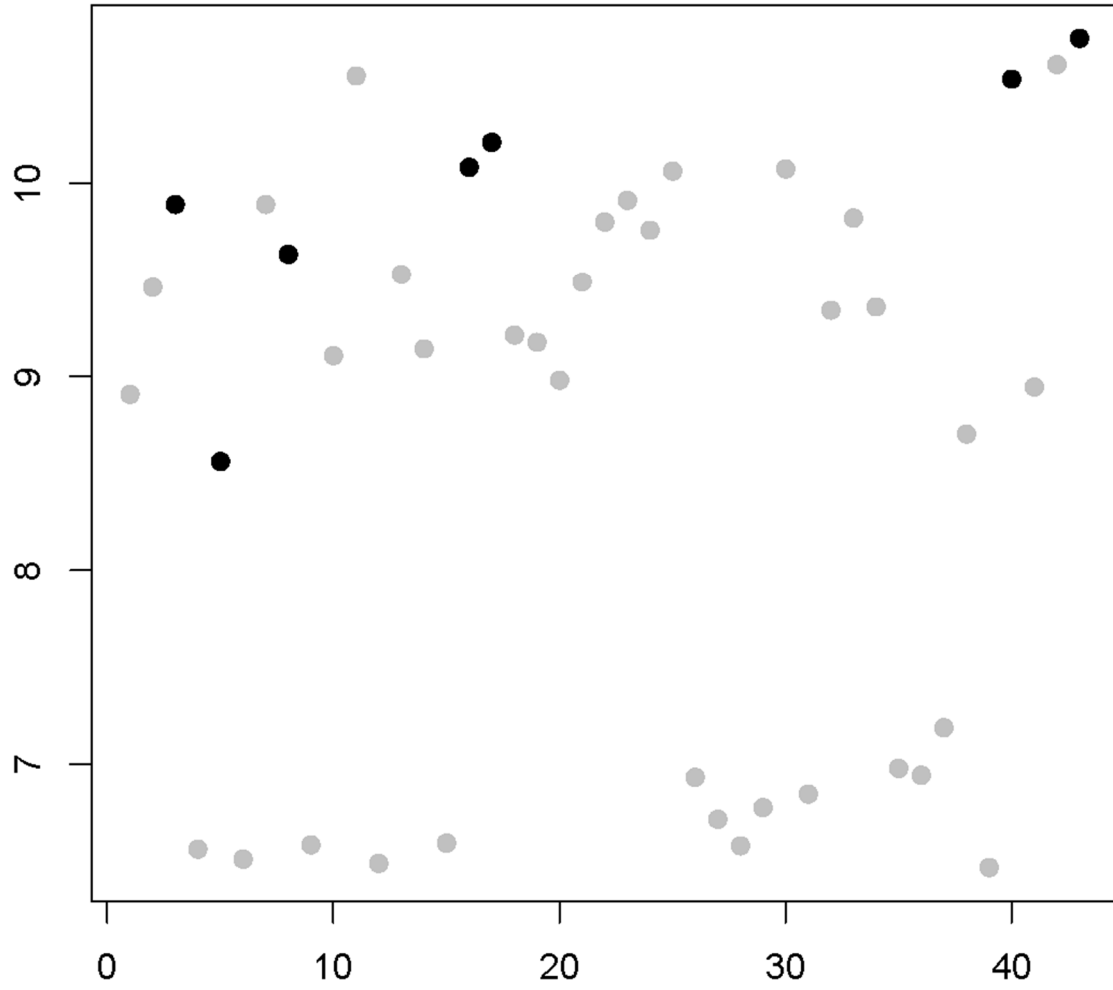


FIG. 4

TLR5

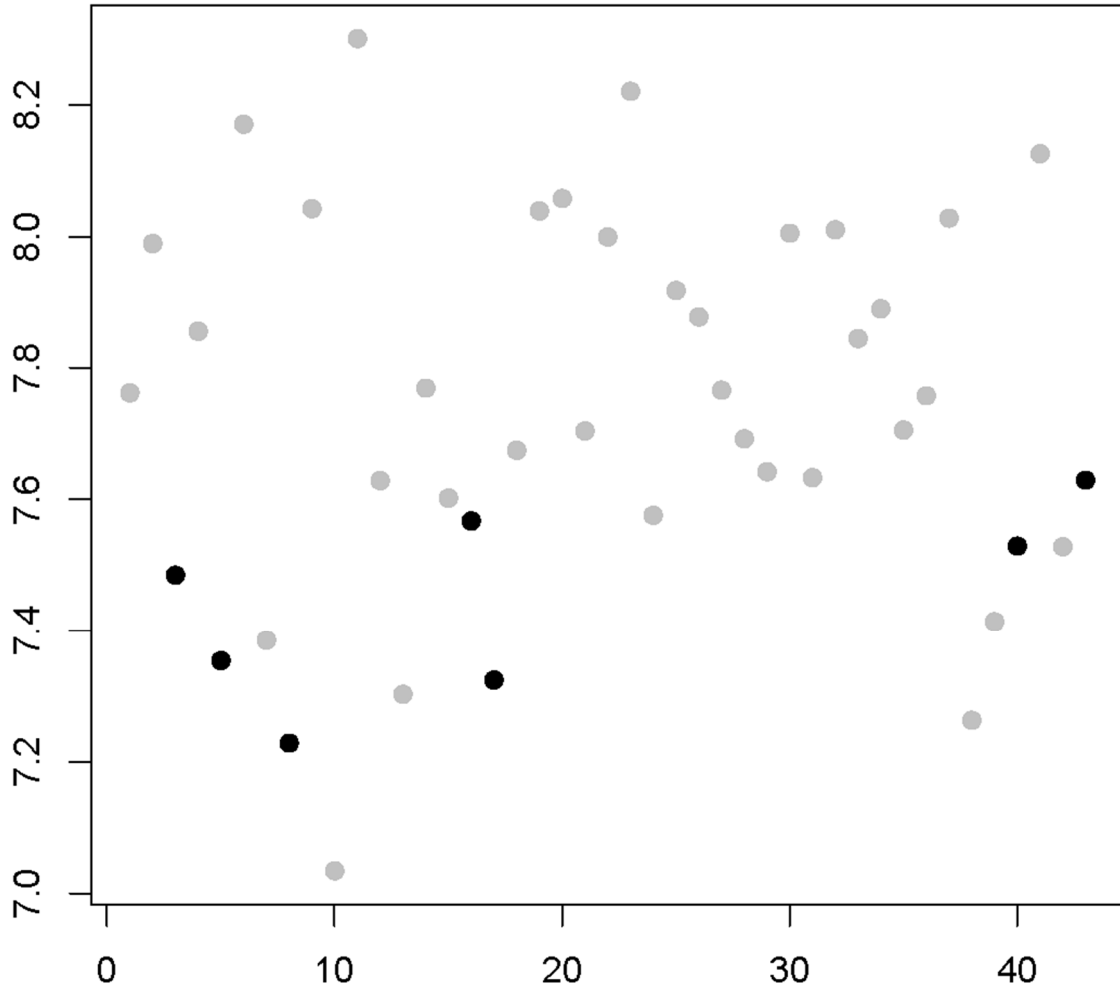


FIG. 5

IL2RB

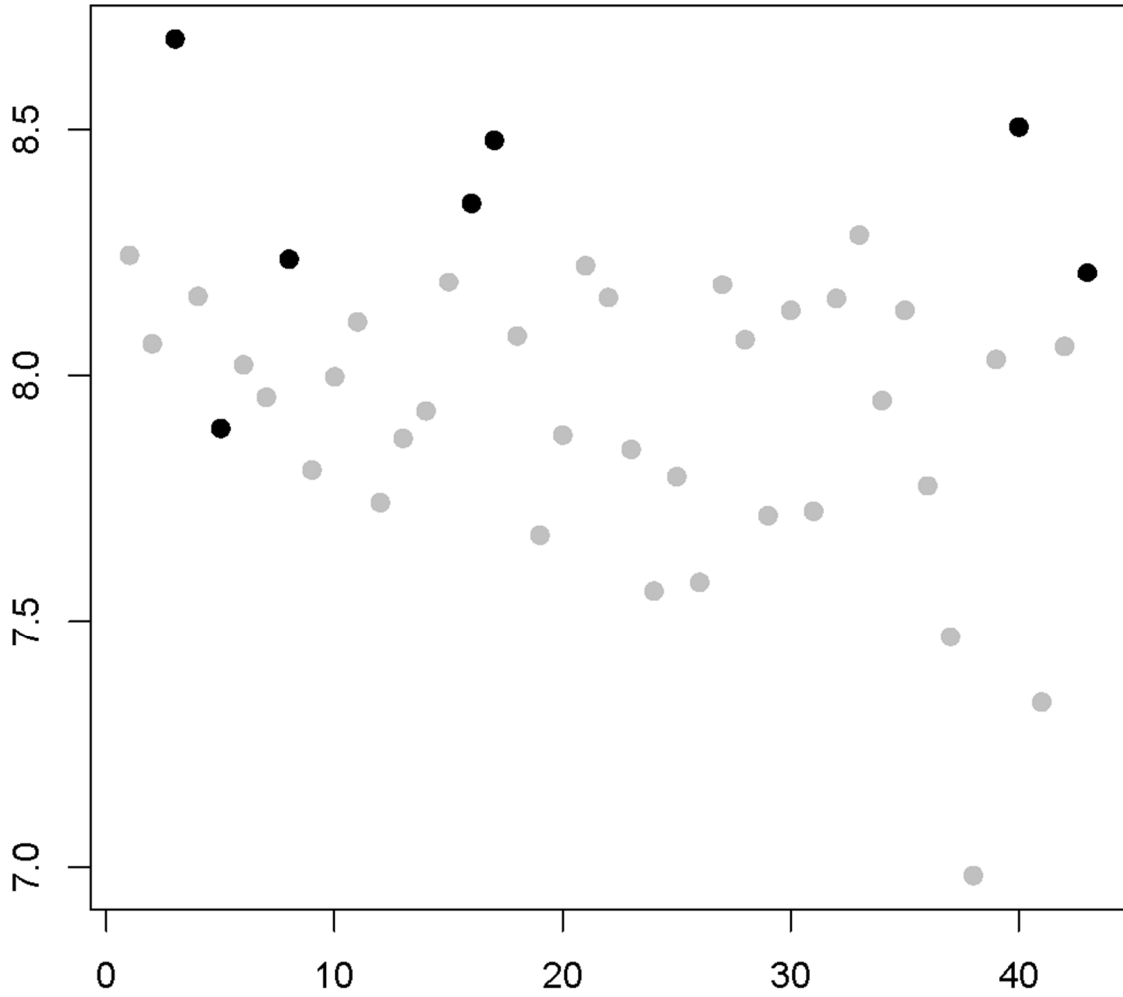


FIG. 6

EAT2

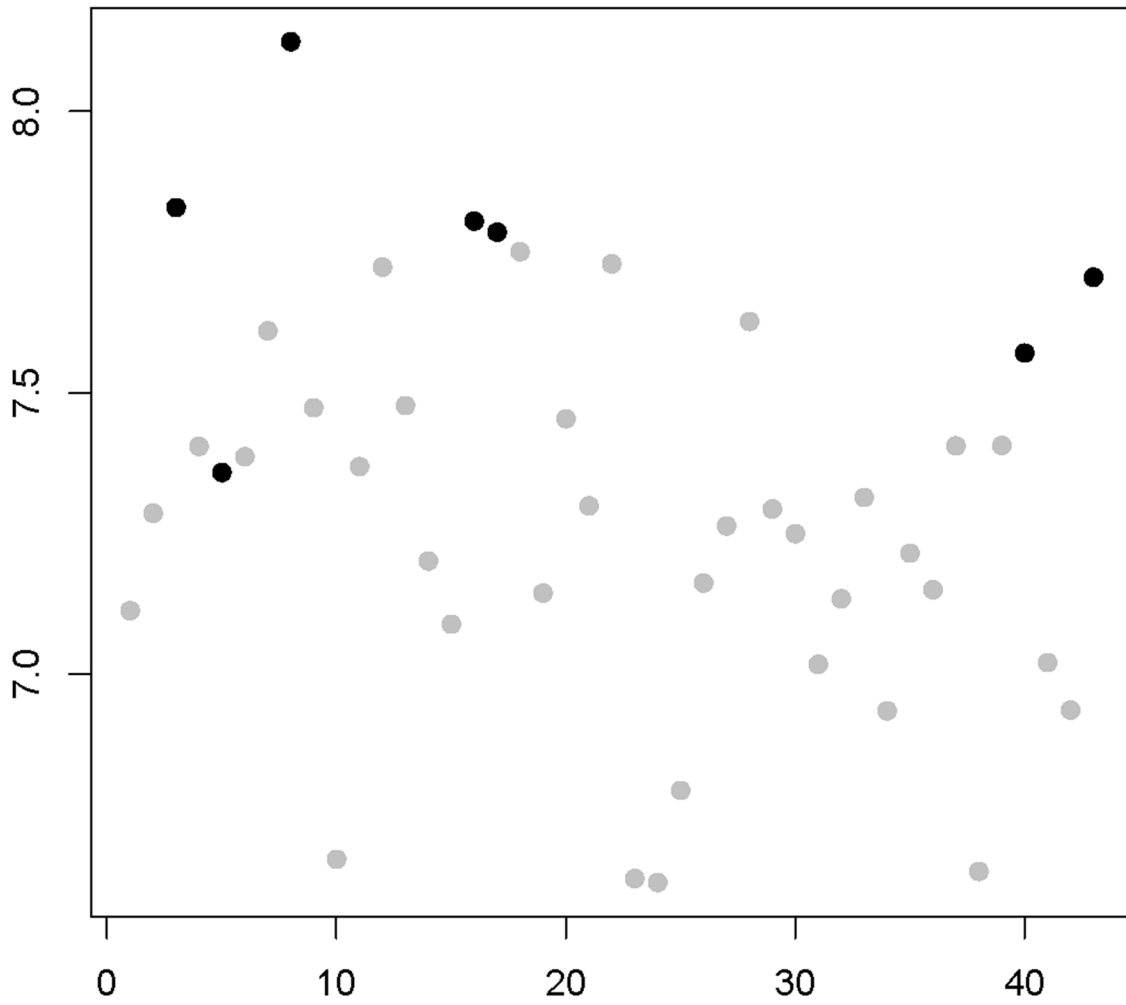




FIG. 7

CAMP

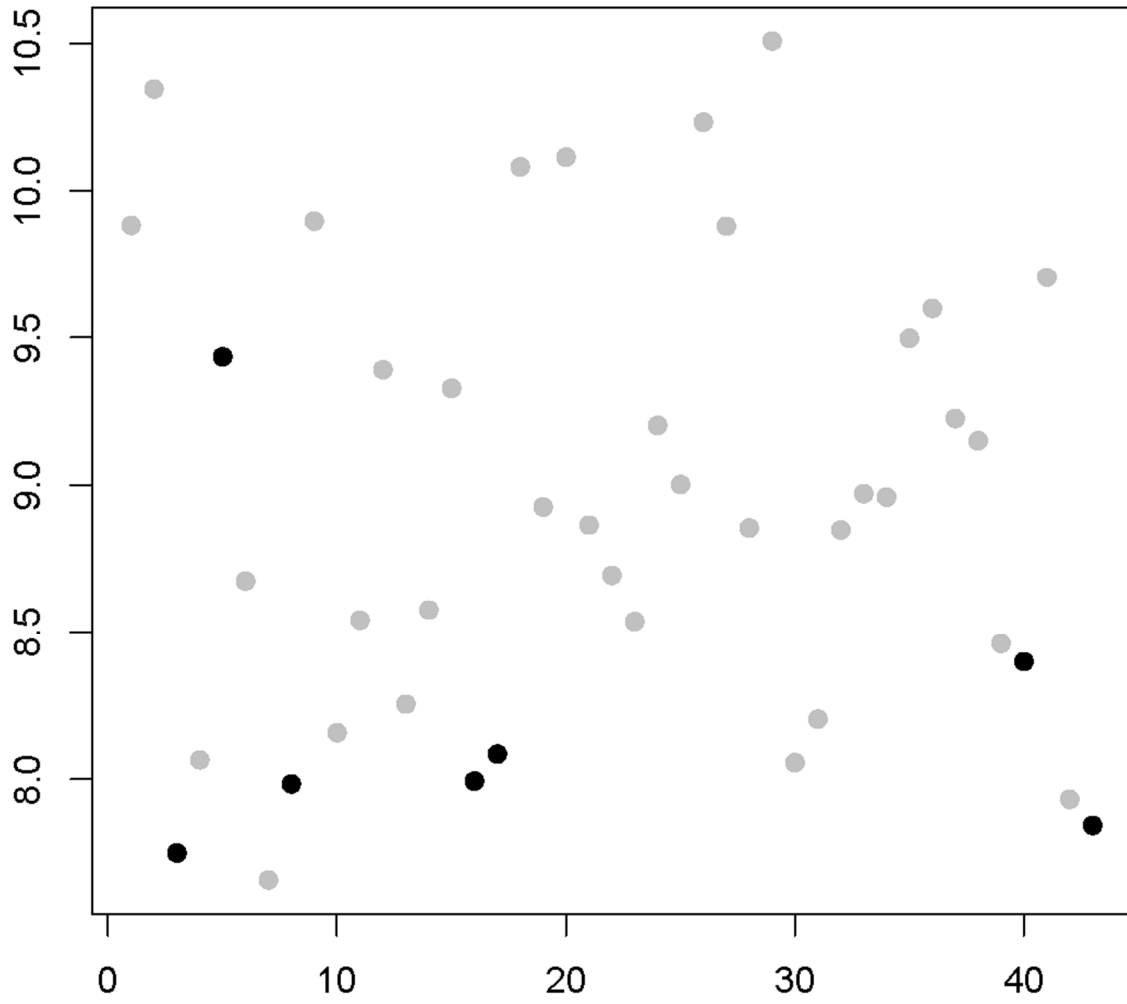


FIG. 8

SLC2A3

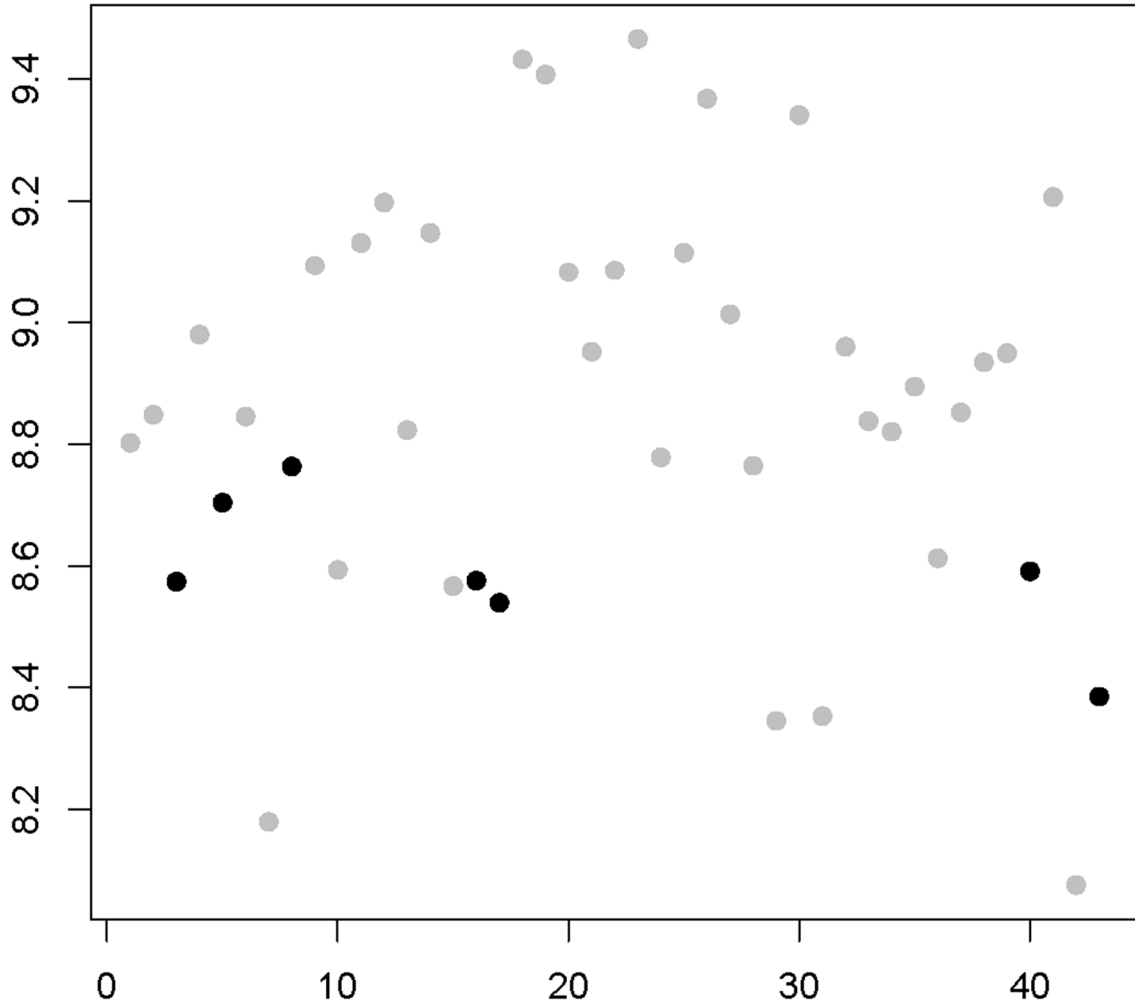
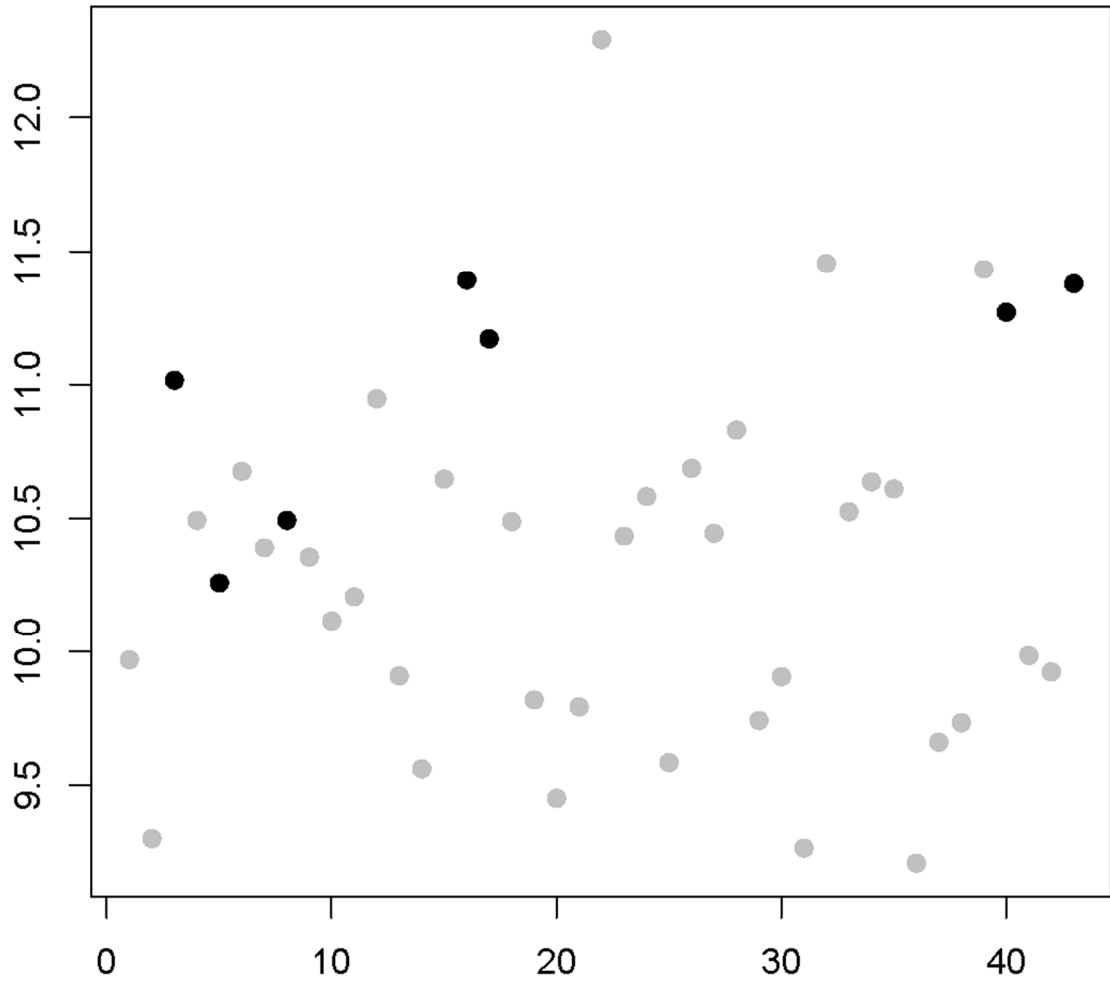


FIG. 9

GPLY



## REFERENCIAS CITADAS EN LA DESCRIPCIÓN

La lista de referencias citadas por el solicitante es, únicamente, para conveniencia del lector. No forma parte del documento de patente europea. Si bien se ha tenido gran cuidado al compilar las referencias, no pueden excluirse errores u omisiones y la OEP declina toda responsabilidad a este respecto.

### Documentos de patente citados en la descripción

- WO 2007135568 A [0008]
- WO 2007038501 A [0008]

### Literatura no patente citada en la descripción

- **THIERRY LEQUERRÉ.** profiling in white blood cells predicts infliximab responsiveness in rheumatoid arthritis. *Gene* [0006]
- **JOHAN LINDBERG.** Effect of infliximab on mRNA expression profiles in synovial tissue of rheumatoid arthritis patients. *Effect of infliximab on mRNA expression profiles in synovial tissue of rheumatoid arthritis patients* [0008]
- **LEDINGHAM et al.** *Rheumatology*, 2005 [0019]
- **VAN'T VEER et al.** *Nature*, 2002 [0040]