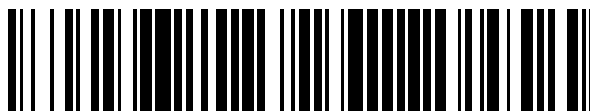


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 375 577**

51 Int. Cl.:
C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **07789645 .4**

96 Fecha de presentación: **12.07.2007**

97 Número de publicación de la solicitud: **2061910**

97 Fecha de publicación de la solicitud: **27.05.2009**

54 Título: **MÉTODO DE PRONÓSTICO.**

30 Prioridad:
12.07.2006 GB 0613844

45 Fecha de publicación de la mención BOPI:
02.03.2012

45 Fecha de la publicación del folleto de la patente:
02.03.2012

73 Titular/es:
**PROGENIKA BIOPHARMA, S.A.
PARQUE TECNOLÓGICO DE ZAMUDIO,
IBAZABAL BIDEA-EDIFICIO 801-A-2ª PLANTA
48160 DERIO, ES y
JUSTE S.A.Q.F.**

72 Inventor/es:
**EGUIARA, Arrate;
TEJEDOR HERNÁNDEZ, Diego;
MARTÍNEZ MARTÍNEZ, Antonio y
SIMÓN BUELA, Laureano**

74 Agente: **Zea Checa, Bernabé**

ES 2 375 577 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 de pronóstico.

5 **Solicitudes relacionadas**

[0001] Esta solicitud está relacionada con la Solicitud de Patente del Reino Unido N° 0613844.0, presentada el 12 de julio de 2006.

10 **Campo de la invención**

[0002] La invención se refiere a métodos y productos, en particular, a micromatrices para el genotipado *in vitro* de variaciones genéticas asociada a artritis reumatoide (RA). La invención se refiere además a métodos para el pronóstico y tratamiento de la RA, y a productos para su uso en los mismos.

15 **Antecedentes de la invención**

[0003] La artritis reumatoide (RA) es un trastorno inflamatorio crónico que característicamente tiene implicación articular. La prevalencia es de aproximadamente el 0,8 (aproximadamente 2,9 millones de personas en Europa), con algunas variaciones entre grupos étnicos. La RA es más común en mujeres que en hombres.

[0004] En el proceso patológico de la RA, la inflamación implica a la membrana sinovial y puede causar daños al cartílago y al propio hueso. Las áreas que pueden afectarse incluyen las articulaciones de las manos, muñecas, cuello, mandíbula, codos, pies y tobillos. Con frecuencia, la RA causa que las articulaciones del cuerpo se vean afectadas de una forma simétrica, lo que significa que la enfermedad ataca a la misma articulación en ambos lados del cuerpo, por ejemplo, ambas manos.

[0005] El diagnóstico de RA se realiza actualmente de acuerdo con los criterios expuestos por el Colegio Americano de Reumatología en 1987 (publicados en Arnett et al., *Arthritis and Rheumatism* 1988 Mar; 31(3): 315-24).

[0006] Se usan comúnmente varios agentes farmacológicos para gestionar los signos y síntomas de la RA. Estos se incluyen en cuatro clases principales: AINE (fármacos antiinflamatorios no esteroideos), que alivian el dolor y la inflamación asociados con la RA; DMARD (fármacos antirreumáticos modificadores de enfermedad), que tanto alivian los síntomas como ayudan a controlar la RA por modificación de su proceso patológico; glucocorticoides (un grupo de fármacos antiinflamatorios que están relacionados con el cortisol, una hormona esteroidea natural producida por el cuerpo); y agentes biológicos (fármacos como Kineret® (anakinra), que están diseñados para dirigirse a moléculas específicas en el sistema inmune que contribuyan al proceso patológico de la RA.

[0007] Los estudios epidemiológicos y genéticos han proporcionado pruebas de la presencia de factores de susceptibilidad genética para la RA, que representan aproximadamente el 60% de la variación en la predisposición a la enfermedad. La identificación de aquellos genes relacionados con la RA podría llevar a una mejor comprensión de la patogénesis, diagnóstico, localización y pronóstico y, en última instancia, a un tratamiento adecuado apropiado. Desde un punto de vista clínico, el diagnóstico, pronóstico y localización tempranos de la enfermedad significaría un cambio importante en las decisiones terapéuticas usadas para el tratamiento.

[0008] Los estudios de asociación han identificado recientemente varios genes en los que una o más variaciones genéticas dan como resultado un mayor o menor riesgo de contraer esta enfermedad, una mejor o peor respuesta a fármacos y/o un mejor o peor pronóstico. Un medio de caracterización de múltiples variaciones asociadas a RA que pudiera usarse clínicamente proporcionaría un gran avance en el diagnóstico y la terapia.

[0009] Con frecuencia se usan microplacas de ADN para discriminar entre alelos en loci genéticos.

[0010] En 2001, el Consorcio para el Proyecto del Genoma Humano y la compañía privada Celera presentaron el primer ejemplo completo del genoma humano con 30.000 genes. A partir de este momento, comenzó la posibilidad de estudiar el genoma completo o de estudios a gran escala (alto rendimiento). Las denominadas "microplacas de ADN", también denominadas "micromatrices", "matrices de ADN" o "biomicroplacas de ADN" son aparatos que pueden usar la genómica funcional para estudios a gran escala. La genómica funcional estudia cambios en la expresión de genes debidos a factores ambientales y a características genéticas de un individuo. Las secuencias genéticas presentan pequeñas variaciones interindividuales en un único nucleótido denominadas SNP ("polimorfismo de un solo nucleótido"), que en un pequeño porcentaje están implicadas en cambios en la expresión y/o función de genes que causan ciertas patologías. La mayoría de los estudios que aplican microplacas de ADN estudian la expresión génica, aunque las microplacas también se usan en la detección de SNP.

[0011] La primera microplaca de ADN fue la "transferencia de Southern", en la que se usaron moléculas de ácido

nucleico marcadas para examinar moléculas de ácido nucleico unidas a un soporte sólido. El soporte era típicamente una membrana de nylon.

[0012] Dos avances marcaron el comienzo definitivo de la microplaca de ADN. El uso de un soporte sólido no poroso, tal como vidrio, permitió la miniaturización de matrices, permitiendo de este modo que se incorporasen un gran número de características de sondas individuales sobre la superficie del soporte a una densidad de >1.000 sondas por cm². La adaptación de técnicas fotolitográficas semiconductoras permitió la producción de microplacas de ADN que contenían más de 400.000 oligonucleótidos diferentes en una región de aproximadamente 20 μm², denominadas microplacas de ADN de alta densidad.

[0013] En general, una microplaca de ADN comprende un soporte sólido que contiene cientos de fragmentos de secuencias de diferentes genes representados en forma de ADN, ADNc u oligonucleótidos fijos, unidos a la superficie sólida en posiciones fijas. Los soportes son generalmente portaobjetos de vidrio para el microscopio, membranas de nylon o "microplacas" de silicio. Es importante que las secuencias de nucleótidos o sondas se unan al soporte en posiciones fijas, ya que la localización robotizada de cada sonda determina el gen cuya expresión se está midiendo. Las microplacas de ADN pueden clasificarse como:

- microplacas de ADN de alta densidad: los oligonucleótidos que se encuentran en la superficie del soporte, por ejemplo, portaobjetos de vidrio, se han sintetizado "in situ", mediante un método denominado fotolitografía.
- microplacas de ADN de baja densidad: los oligonucleótidos, ADNc o fragmentos de amplificación por PCR se depositan en forma de nanogotas en la superficie del soporte, por ejemplo, vidrio, por medio de un robot que imprime esas secuencias de ADN en el soporte. Hay muy pocos ejemplos de microplacas de ADN de baja densidad que existen: una microplaca de ADN para detectar 5 mutaciones en el gen de la tirosinasa; una microplaca de ADN para detectar mutaciones en p54 y k-ras; una microplaca de ADN para detectar 12 mutaciones que causan cardiomiopatía hipertrófica; una microplaca de ADN para el genotipado de cepas de *Escherichia coli*; o microplacas de ADN para detectar patógenos tales como *Cryptosporidium parvum* o rotavirus.

[0014] Para estudios de expresión genética, las sondas depositadas en la superficie sólida, por ejemplo, vidrio, se hibridan con ADNc sintetizados a partir de ARNm extraídos de una muestra dada. En general, el ADNc se ha marcado con un fluoróforo. Cuanto mayor sea el número de moléculas de ADNc unidas a su secuencia complementaria en la microplaca de ADN, mayor será la intensidad de la señal fluorescente detectada, típicamente medida con un láser. Esta medida es por tanto un reflejo del número de moléculas de ARNm en la muestra analizada y, por consiguiente, un reflejo del nivel de expresión de cada gen representado en la microplaca de ADN.

[0015] Las microplacas de ADN de expresión génica también contienen típicamente sondas para la detección de la expresión de genes de control, con frecuencia denominados "genes constitutivos", que permiten normalizar los resultados experimentales y que se comparen múltiples experimentos de una forma cuantitativa. Con la microplaca de ADN, los niveles de expresión de cientos o miles de genes en una célula pueden determinarse en un solo experimento. El ADNc de una muestra de ensayo y el de una muestra de control pueden marcarse con dos fluoróforos diferentes, de modo que la misma microplaca de ADN puede usarse para estudiar diferencias en la expresión génica.

[0016] Las microplacas de ADN para la detección de polimorfismos genéticos, cambios o mutaciones (en general, variaciones genéticas) en la secuencia de ADN comprenden una superficie sólida, típicamente vidrio, sobre la que se depositan un gran número de secuencias genéticas (las sondas) complementarias a las variaciones genéticas a estudiar. Usando impresoras robóticas convencionales para aplicar sondas a la matriz, puede obtenerse una alta densidad de características de sondas individuales, por ejemplo, pueden conseguirse típicamente densidades de sondas de 600 características por cm² o más. La ubicación de las sondas en una matriz se controla con precisión mediante el dispositivo de impresión (robot, impresora de inyección, máscara fotolitográfica, etc.) y las sondas se alinean en una rejilla. La organización de las sondas en la matriz facilita la posterior identificación de interacciones de sonda-diana específicas. Además es común, pero no necesario, dividir las características de la matriz en sectores más pequeños, también en forma de rejilla, a los que posteriormente se hace referencia como submatrices. Las submatrices comprenden típicamente 32 características de sondas individuales, aunque cada submatriz puede comprender un número menor (por ejemplo 16) o mayor (por ejemplo 64 o más) de características.

[0017] Una estrategia usada para detectar variaciones genéticas implica la hibridación con secuencias que reconocen específicamente el alelo normal y el mutante en un fragmento de ADN obtenido en una muestra de ensayo. Típicamente, el fragmento se ha amplificado, por ejemplo, mediante el uso de la reacción en cadena de la polimerasa (PCR), y se ha marcado, por ejemplo, con una molécula fluorescente. Puede usarse un láser para detectar fragmentos marcados unidos en la microplaca y, por lo tanto, un individuo que sea homocigoto para el alelo normal puede distinguirse específicamente de individuos heterocigotos (en el caso de afecciones dominantes autosómicas entonces estos individuos se denominan portadores) o de los que son homocigotos para el alelo mutante.

[0018] Otra estrategia para detectar variaciones genéticas comprende llevar a cabo una reacción de amplificación

o una reacción de extensión en la propia microplaca de ADN.

[0019] Para métodos basados en hibridación diferencial, existen varios métodos para analizar los datos de hibridación para el genotipado:

- Aumento en el nivel de hibridación: Se compara el nivel de hibridación de sondas complementarias con los alelos normales y mutantes.
- Disminución en el nivel de hibridación: Pueden identificarse diferencias en la secuencia entre una muestra de control y una muestra de ensayo por una disminución en el nivel de hibridación de los oligonucleótidos totalmente complementarios con una secuencia de referencia. Se produce una pérdida completa en individuos homocigotos mutantes, mientras que se produce una pérdida de sólo el 50% en los heterocigotos. En microplacas de ADN para examinar todas las bases de una secuencia de "n" nucleótidos ("oligonucleótido") de longitud en ambas cadenas, son necesarios un mínimo de "2n" oligonucleótidos que solapen con el oligonucleótido previo en toda la secuencia excepto en el nucleótido. Típicamente, el tamaño de los oligonucleótidos es de aproximadamente 25 nucleótidos. El número aumentado de oligonucleótidos usado para reconstruir la secuencia reduce los errores derivados de fluctuaciones del nivel de hibridación. Sin embargo, el cambio exacto en la secuencia no puede identificarse con este método; posteriormente es necesaria una secuenciación para identificar la mutación.

[0020] Cuando se lleva a cabo la amplificación o extensión en la propia microplaca de ADN, se presentan tres métodos a modo de ejemplo:

[0021] En la estrategia de minisequenciación, un cebador específico de mutación se fija sobre el portaobjetos y después de una reacción de extensión con dideoxinucleótidos fluorescentes, se captura la imagen de la microplaca de ADN con un explorador.

[0022] En la estrategia de extensión de cebadores, se diseñan dos oligonucleótidos para la detección de las secuencias de tipo silvestre y mutante, respectivamente. La reacción de extensión se lleva a cabo posteriormente con un nucleótido marcado fluorescentemente y los nucleótidos restantes sin marcar. En cualquier caso el material de partida puede ser una muestra de ARN o un producto de ADN amplificado por PCR.

[0023] En la estrategia de matrices con etiqueta, se lleva a cabo una reacción de extensión en solución con cebadores específicos que llevan una secuencia 5' determinada o "etiqueta". El uso de microplacas de ADN con oligonucleótidos complementarios a estas secuencias o "etiquetas" permite la captura de los productos resultantes de la extensión. Los ejemplos de esto incluyen la microplaca de ADN de alta densidad "Flex-flex" (Affymetrix).

[0024] Para el diagnóstico genético, debe tenerse en cuenta la simplicidad. La necesidad de reacciones de amplificación y purificación presenta desventajas para los métodos de extensión/amplificación sobre la microplaca en comparación con métodos basados en hibridación diferencial.

[0025] Típicamente, el análisis de microplacas de ADN se lleva a cabo usando técnicas de hibridación diferencial. Sin embargo, la hibridación diferencial no produce una especificidad o sensibilidad tan elevada como los métodos asociados con amplificación sobre portaobjetos de vidrio. Por esta razón, es necesario el desarrollo de algoritmos matemáticos que aumenten la especificidad y sensibilidad de la metodología de hibridación (Cutler DJ, Zwick ME, Carrasquillo MN, Yohn CT, Tobi KP, Kashuk C, Mathews DJ, Shah N, Eichler EE, Warrington JA, Chakravarti A. Genome Research; 11: 1913-1925 (2001). La asociación del polimorfismo PTPN22 y de anticuerpos anti-péptido cíclico citrulinado con la aparición de RA en el futuro se describe en Johansson M. et al; Arthritis Research & Therapy 2006, 8: R19.

[0026] Por lo tanto, a pesar de los avances en tecnología, los problemas de los métodos existentes para analizar simultáneamente un gran número de variaciones genéticas de una forma sensible, específica y reproducible han impedido la aplicación de microplacas de ADN para su uso rutinario en el diagnóstico clínico.

Descripción resumida de la invención

[0027] Los inventores han identificado nuevos medios para pronosticar fenotipos de RA usando combinaciones de variables de SNP y variables clínicas informativas. Por consiguiente, la invención proporciona un método de pronóstico de un fenotipo de artritis reumatoide (RA) en un sujeto, que comprende:

- (I) obtener resultados para una o más variables de polimorfismos de un solo nucleótido y una o más variables clínicas enumeradas en la Tabla 11A para el sujeto; y
- (II) usar los resultados obtenidos en (I) para pronosticar el fenotipo;

en el que

- (i) un resultado para una variable de SNP es la identidad del nucleótido en el ADN genómico del sujeto en la

posición del polimorfismo de un solo nucleótido;

(ii) un resultado para la variable clínica ANTI-PCC es el nivel de anticuerpo anti-péptido cíclico citrulinado en el suero del sujeto;

5 (iii) un resultado para la variable clínica ANTI-PCC BI es una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado en el suero del sujeto;

(iv) un resultado para la variable clínica VSG es la velocidad de sedimentación eritrocitaria en la sangre del sujeto;

(v) un resultado para la variable clínica NoARTDOL es el número de articulaciones dolorosas mencionado por el sujeto;

10 (vi) un resultado para la variable clínica RF es el nivel de anticuerpo factor reumatoide en el suero del sujeto;

(vii) un resultado para la variable clínica PCR es el nivel de proteína C reactiva en la sangre; y

(viii) un resultado para la variable clínica edad de aparición es la edad en años a la que se le diagnosticó al sujeto RA;

15 y en el que:

(a) el fenotipo de RA es una puntuación de cuestionario de evaluación de la salud (HAQ) > 2 (HAQ>2) y las variables para las que se obtienen resultados en la etapa (I) comprenden las variables de SNP y clínicas de HAQ>2 en la Tabla 11A; y/o

20 (b) el fenotipo de RA es la presencia de múltiples erosiones en las manos y en los pies (RX2) y las variables para las que se obtienen resultados en la etapa (I) comprenden las variables de SNP y clínicas de RX2 en la Tabla 11A; y/o

(c) el fenotipo de RA es recibir 3 o más tratamientos diferentes (TTO_30M) y las variables para las que se obtienen resultados en la etapa (I) comprenden las variables de SNP y clínicas de TTO_30M en la Tabla 11A; y/o

25 (d) el fenotipo de RA es verse forzado a dejar el empleo debido a RA (dejar el trabajo) y las variables para las que se obtienen resultados en la etapa (I) comprenden las variables de SNP y clínicas de "dejar el trabajo" en la Tabla 11A; y/o

(e) el fenotipo de RA es recibir una prótesis articular (intervención quirúrgica) y las variables para las que se obtienen resultados en la etapa (I) comprenden las variables de SNP y clínicas de "intervención quirúrgica" en la Tabla 11A; y/o

30 (f) el fenotipo de RA es la intolerancia a metotrexato y las variables para las que obtienen resultados en la etapa (I) comprenden las variables de SNP y clínicas de "intolerancia a metotrexato" en la Tabla 11A; y/o

(g) el fenotipo de RA es la remisión y las variables para las que se obtienen resultados en la etapa (I) comprenden las variables de SNP y clínicas de "remisión" en la Tabla 11A.

35 **[0028]** La invención también proporciona un método de obtención de una función de probabilidad para su uso en el pronóstico de un fenotipo de RA, un método computacional de obtención de una función de probabilidad para su uso en el pronóstico de un fenotipo de RA, y un método para el pronóstico de un fenotipo de RA en un sujeto, que comprende el uso de una función de probabilidad obtenida usando los datos de cualquiera de las Tablas 4 a 10, como se expone en las reivindicaciones.

45 **[0029]** Los inventores también han identificado SNP que tienen una asociación alélica significativa con el recidivado del cáncer de próstata. Por consiguiente, la invención también proporciona un método de pronóstico de un fenotipo de RA en un sujeto, que comprende determinar el genotipo del sujeto en una o más posiciones de polimorfismo de un solo nucleótido seleccionadas de los SNP de la Tabla 12.

[0030] La invención también proporciona un método *in vitro* para genotipar variaciones genéticas asociadas a RA en un individuo, como se expone en las reivindicaciones.

50 **[0031]** Los aspectos adicionales de la invención incluyen un método computacional para obtener un genotipo a partir de datos de intensidad de hibridación de microplaca de ADN, un método de obtención de funciones lineales para su uso en un método de genotipado de la invención, un método computacional de obtención de funciones lineales para su uso en un método de genotipado de la invención, un método de diagnóstico de RA o susceptibilidad a RA en un individuo, que comprende genotipar a un individuo con respecto a una o más variaciones genéticas, métodos para seleccionar un tratamiento para RA en un sujeto y para tratar la RA en un sujeto, un método de identificación de variaciones genéticas predictivas de un fenotipo de RA particular, y un método de predicción del probable desarrollo de un fenotipo de RA en un individuo usando la variación o variaciones identificadas.

60 **[0032]** Todavía más aspectos incluyen un sistema informático que comprende un procesador y medios para controlar el procesador para llevar a cabo un método computacional de la invención, un programa informático que comprende un código de programa informático que, cuando se ejecuta en un ordenador o red de ordenadores, provoca que el ordenador o red de ordenadores lleven a cabo un método computacional de la invención.

65 **[0033]** La invención también proporciona una microplaca de ADN o micromatriz adecuada para su uso en los métodos de la invención, una sonda oligonucleotídica, pareja de sondas o conjunto de 4 sondas enumerado en la

Tabla 2 (Figura 2), un cebador oligonucleotídico o pareja de cebadores enumerado en la Tabla 3 (Figura 3), un kit de amplificación por PCR, que comprende al menos una pareja de los cebadores enumerados, un kit de diagnóstico para la detección de variaciones genéticas asociadas a RA, y un kit para el pronóstico de un fenotipo de RA en un sujeto.

- 5
- [0034]** En particular, se describen las realizaciones siguientes.
- 10
- 1. Un método de pronóstico de un fenotipo de remisión de la artritis reumatoide (RA) en un sujeto que tiene RA, comprendiendo dicho método determinar (i) el genotipo de un individuo al menos en la posición de polimorfismo de un solo nucleótido en rs2476601; y (ii) una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC BI) en el suero del sujeto.
 - 2. Un método de acuerdo con la realización 1, en el que el método comprende:
 - 15 (I) obtener resultados:
 - (i) de la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en el alelo nulo del gen de la Glutación S-transferasa M1 (GSTM1) y en rs2476601; y
 - 20 (ii) de una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC BI) en el suero del sujeto; y
 - (II) usar los resultados obtenidos en (I) para pronosticar el fenotipo de remisión de RA.
 - 3. Un método de acuerdo con la realización 1 o la realización 2, que comprende además pronosticar un fenotipo de artritis reumatoide (RA) agresivo en el sujeto, comprendiendo dicho fenotipo de RA agresivo uno o más de: una puntuación de cuestionario de evaluación de la salud (HAQ) de >2 (HAQ >2); la presencia de múltiples erosiones en las manos y en los pies (RX2); recibir 3 o más tratamientos diferentes (TTO_30M); verse forzado a dejar el empleo debido a la RA (dejar el trabajo); y recibir una prótesis articular (intervención quirúrgica), comprendiendo dicho método:
 - 25 (I) obtener resultados de acuerdo con el modelo 1, modelo 2, modelo 3, modelo 4 y/o modelo 5 para el sujeto; y
 - 30 (II) usar los resultados obtenidos en (I) para pronosticar el fenotipo de RA agresivo,
- 35 en el que:
- (a) cuando el fenotipo de RA agresivo comprende una puntuación de cuestionario de evaluación de la salud (HAQ) de > 2 (HAQ >2), dicha etapa (I) comprende los resultados de acuerdo con dicho modelo 1, comprendiendo dichos resultados:
 - 40 (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs1800790 y rs2070874M;
 - (ii) el nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC) en el suero del sujeto; y
 - 45 (iii) la velocidad de sedimentación eritrocitaria (VSG) en la sangre del sujeto,
 - (b) cuando el fenotipo de RA agresivo comprende la presencia de múltiples erosiones en las manos y en los pies (RX2), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 2, comprendiendo dichos resultados:
 - 50 (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs1544410, rs2071592, rs187238 y rs1801275;
 - (ii) una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC BI) en el suero del sujeto; y
 - 55 (iii) el número de articulaciones dolorosas (NoARTDOL) mencionadas por el sujeto,
 - (c) cuando el fenotipo de RA agresivo comprende recibir 3 o más tratamientos diferentes (TTO_30M), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 3, comprendiendo dichos resultados:
 - 60 (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs7531668, rs419598 y rs1024611;
 - (ii) el nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC) en el suero del sujeto;
 - (iii) el nivel de anticuerpo Factor Reumatoide (RF) en el suero del sujeto;
 - (iv) el nivel de proteína C reactiva (PCR) en la sangre del sujeto; y
 - 65 (v) la edad en años a la que se le diagnosticó al sujeto la RA,

(d) cuando el fenotipo de RA agresivo comprende verse forzado a dejar el empleo debido a la RA (dejar el trabajo), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 4, comprendiendo dichos resultados:

5 (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs1042522, el alelo nulo del gen de la Glutación S-transferasa T1 (GSTT1) y rs1801133,

10 (e) cuando el fenotipo de RA agresivo comprende recibir una prótesis articular (intervención quirúrgica), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 5, comprendiendo dichos resultados:

15 (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs2268277, rs1042522, rs10735810 y rs3025039;

(ii) una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC BI) en el suero del sujeto; y

(iii) la velocidad de sedimentación eritrocitaria (VSG) en la sangre del sujeto.

20 - 4. El método de acuerdo con la realización 1 o la realización 2, en el que el método comprende además pronosticar el fenotipo de RA de intolerancia a metotrexato, en el que:

dicha etapa (I) comprende además obtener resultados de acuerdo con el modelo 6, comprendiendo dichos resultados:

25 (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en rs1801133; y

(ii) la edad en años a la que se le diagnosticó al sujeto la RA.

30 - 5. El método de acuerdo con cualquiera de las realizaciones 1 a 4, en el que en la etapa (II), sólo se usan los resultados obtenidos en la etapa (I) para pronosticar el fenotipo de RA.

- 6. El método de acuerdo con cualquiera de las realizaciones anteriores, en el que la obtención de resultados para variables de SNP comprende el análisis de micromatrices o la secuenciación.

35 - 7. El método de acuerdo con cualquiera de las realizaciones 1 a 5, en el que la obtención de resultados para variables de SNP comprende el uso de una o más parejas de sondas oligonucleotídicas enumeradas en la Figura 2.

40 - 8. El método de acuerdo con cualquiera de las realizaciones 1 a 5, en el que la obtención de resultados para variables de SNP comprende la amplificación de ácido nucleico obtenido del sujeto, y el método comprende el uso de una o más parejas de cebadores oligonucleotídicos enumeradas en la Figura 3.

- 9. El método de acuerdo con cualquiera de las realizaciones anteriores, en el que la etapa (II) comprende:

45 (i) introducir los resultados determinados en la etapa (I) en una función de probabilidad, calculando de este modo un valor de función de probabilidad; y

(ii) comparar el valor de la función de probabilidad con los valores de la función de probabilidad calculados para individuos de fenotipo conocido.

50 - 10. El método de acuerdo con cualquiera de las realizaciones anteriores, que comprende el uso de una función de probabilidad obtenida usando los datos de cualquiera de las Tablas 4 a 10.

- 11. Un método de selección de un tratamiento adecuado para tratar la RA en un sujeto, comprendiendo el método:

55 (a) pronosticar un fenotipo de RA en el sujeto mediante un método de acuerdo con cualquiera de las realizaciones anteriores; y

(b) seleccionar un tratamiento que sea adecuado para el fenotipo de RA determinado.

[0035] Todos estos aspectos de la invención son como se exponen en las reivindicaciones.

60

Breve descripción de las secuencias

[0036]

65 Las SEC ID N°: 1-360 son sondas adecuadas para la detección de las variaciones genéticas asociadas a RA de

la Tabla 1A (o Tabla 1B). Las secuencias de sondas se muestran en la Tabla 2.

Las SEC ID N°: 361-540 son cebadores de PCR adecuados para amplificar regiones de ADN diana que comprenden variaciones genéticas asociadas a RA enumeradas en la Tabla 1A (o Tabla 1B). Las secuencias de cebadores se muestran en la Tabla 3.

La SEC ID N°: 541 es un ácido nucleico de control externo.

Las SEC ID N°: 542 y 543 son sondas adecuadas para la detección del ácido nucleico de control externo de la SEC ID N°: 541.

La SEC ID N°: 544 es una secuencia de TAG directa.

La SEC ID N°: 545 es una secuencia de TAG inversa.

Breve descripción de los dibujos

[0037]

Figura 1

Las Tablas 1A y 1B enumeran variaciones genéticas asociadas con RA que pueden analizarse de acuerdo con la invención. Las secuencias de todos los genes mencionados en las Tablas 1A y 1B se conocen y reconocen en el sitio web siguiente: GeneBank (NcRI), GeneCard (Weizmann Institute of Sciences) y snpper.chip.org (Innate Immunity PLA). La Tabla 1A de la Solicitud del Reino Unido N° 0613844.0 incluía un polimorfismo adicional, el polimorfismo IL6-174. En un aspecto, la Tabla 1A como se usa en la presente memoria puede incluir además este polimorfismo.

Figura 2

La Tabla 2 enumera sondas oligonucleotídicas para discriminar entre alelos en los SNP enumerados en las Tablas 1A y 1B. La Tabla enumera dos parejas de sondas para cada SNP (un conjunto de cuatro sondas).

Figura 3

La Tabla 3 enumera cebadores oligonucleotídicos para amplificación por PCR de cada uno de los SNP enumerados en las Tablas 1A y 1B.

Figura 4

(A) Tabla 4, que muestra los dos SNP (40 y 78) y las 2 variables clínicas (VSG y ANTI-PCC) junto con su significación (Sig.) y sus oportunidades relativas (Exp (B)) usados para calcular el modelo para la predicción del fenotipo de HAQ>2. Este modelo proporciona la probabilidad de desarrollar un fenotipo de HAQ>2 a partir de 0 (sin riesgo) a 1 (riesgo máximo).

(B) Curva de ROC (características operativas del receptor) obtenida para el modelo de HAQ>2 que permite la estimación de su potencia discriminatoria. La curva de ROC se ha calculado para maximizar la especificidad, reduciendo de este modo al mismo tiempo el índice de "falsos" positivos. Una especificidad del 95% con una sensibilidad del 41% es el punto de corte para este modelo respecto al fenotipo de HAQ>2. Este modelo muestra un valor de cociente de verosimilitud (LR) de 8.

Figura 5

(A) Tabla 5, que muestra los dos SNP (28 y 41) y la variable clínica (ANTI-PCC BI) junto con su significación (Sig.) y sus oportunidades relativas (Exp (B)) usados para calcular el modelo para la predicción del fenotipo de REMISIÓN (el paciente no precisó ningún tratamiento en los últimos 5 años).

(B) Curva de ROC obtenida para el modelo del fenotipo de REMISIÓN. Como se muestra por la curva, se obtuvo una sensibilidad del 71% y una especificidad del 89% con un LR de 6,5 para este modelo.

Figura 6

(A) Tabla 6, que muestra los cuatro SNP (14, 15, 20 y 26) y las dos variables clínicas (número de articulaciones implicadas y ANTI-PCC BI) junto con su significación (Sig.) y sus oportunidades relativas (Exp (B)) usados para calcular el modelo para la predicción del fenotipo de RX2.

(B) Curva de ROC obtenida para el modelo del fenotipo de RX2. Como se muestra por la curva, se obtuvo una sensibilidad del 41% y una especificidad del 95% con un LR de 6,5 para este modelo.

Figura 7

(A) Tabla 7, que muestra los tres SNP (16, 27 y 73) y las cuatro variables clínicas (y ANTI-PCC, FR, PCR y edad de aparición) junto con su significación (Sig.) y sus oportunidades relativas (Exp (B)) usados para calcular el modelo para la predicción del fenotipo de TTO_3OM.
 (B) Curva de ROC obtenida para el modelo del fenotipo de TTO_3OM. Como se muestra por la curva, se obtuvo una sensibilidad del 31% y una especificidad del 95% con un LR de 6,2 para este modelo.

Figura 8

(A) Tabla 8, que muestra los cuatro SNP (2, 12, 29 y 32) y las dos variables clínicas (y ANTI-PCC y VSG) junto con su significación (Sig.) y sus oportunidades relativas (Exp (B)) usados para calcular el modelo para la predicción del fenotipo de intervención quirúrgica.
 (B) Curva de ROC obtenida para el modelo del fenotipo de intervención quirúrgica. Como se muestra por la curva, se obtuvo una sensibilidad del 23% y una especificidad del 95% con un LR de 5,4 para este modelo.

Figura 9

(A) Tabla 9, que muestra los tres SNP (32, 42 y 50) junto con su significación (Sig.) y sus oportunidades relativas (Exp (B)) usados para calcular el modelo para la predicción del fenotipo de “dejar el trabajo”.
 (B) Curva de ROC obtenida para el modelo del fenotipo de “dejar el trabajo”. Como se muestra por la curva, se obtuvo una sensibilidad del 10% y una especificidad del 95% con un LR de 1,9 para este modelo.

Figura 10

(A) Tabla 10, que muestra el SNP (50) y la variable clínica (edad de aparición) junto con su significación (Sig.) y sus oportunidades relativas (Exp (B)) usados para calcular el modelo para la predicción del fenotipo de “respuesta a la terapia”.
 (B) Curva de ROC obtenida para el modelo del fenotipo de “respuesta a la terapia”. Como se muestra por la curva, se obtuvo una sensibilidad del 16% y una especificidad del 96% con un LR de 4,3 para este modelo.

Figura 11

(A) La Tabla 11 muestra las variables de SNP y las variables clínicas incluidas en cada uno de los modelos 1 a 7 descritos en la presente memoria. La Tabla indica qué variables de SNP y variables clínicas (de las enumeradas en la primera columna) son informativas para determinar cada fenotipo, y muestra qué resultado para cada variable está asociado con un mal pronóstico de fenotipo.
 (B) La Tabla 11B proporciona más información incluyendo identificadores rs para los SNP de la Tabla 11A.

Figura 12

La Tabla 12 enumera las asociaciones alélicas de SNP individuales con los fenotipos de RA descritos en la presente memoria.

Descripción detallada de la invención

[0038] La artritis reumatoide (RA) presenta varios fenotipos, más notablemente en términos de gravedad de la enfermedad. La enfermedad leve se distingue de la enfermedad grave y destructiva, así como la velocidad y naturaleza de la progresión de la enfermedad. Se cree que la heterogeneidad clínica está correlacionada con la heterogeneidad genética.

[0039] Usando la Arctip de la presente invención y la investigación clínica, los inventores han identificado varios perfiles (basándose en combinaciones de variables de SNP y clínicas) que son informativos para predecir fenotipos de RA. Los inventores han establecido por lo tanto modelos para predecir el curso de la RA en pacientes con RA. Por consiguiente, en un aspecto, la presente invención se refiere a métodos para el pronóstico de la RA.

[0040] Por ejemplo, la invención se refiere a métodos para predecir una RA agresiva, según esté representada por uno o más de los siguientes fenotipos:

1. tener un cuestionario de evaluación de la salud de más de 2;
2. tener múltiples erosiones en las manos y en los pies;
3. recibir 3 o más tratamientos diferentes;
4. verse obligado a dejar el trabajo; y
5. necesidad de prótesis articular (intervención quirúrgica).

[0041] La invención también se refiere a métodos para predecir la respuesta a la terapia, en particular para predecir la tolerancia a metotrexato. La invención se refiere además a métodos para predecir la remisión.

[0042] Los inventores seleccionaron una población de estudio de individuos españoles, como en el Ejemplo 2. Cada individuo se evaluó clínicamente para determinar la presencia (sí) o ausencia (no) de cada uno de siete fenotipos, en concreto los fenotipos 1 a 5 anteriores, el fenotipo de intolerancia a metotrexato y el fenotipo de remisión.

[0043] Cada individuo se ensayó también para las variables clínicas y analíticas basales, y se genotipó en varios loci genéticos usando la micromatriz de ADN Artchip de la presente invención (véase el Ejemplo 2). Los inventores usaron después análisis genético para seleccionar un subconjunto de los SNP más informativos para un modelado adicional. Se llevaron a cabo análisis estadísticos para establecer siete modelos (basándose en combinaciones de SNP informativos y variables clínicas informativas) que permitirán una discriminación fiable entre pacientes con formas alternativas (sí y no) de cada fenotipo en la población con alta especificidad, sensibilidad y precisión.

[0044] Las variables de SNP y clínicas que se seleccionaron para su inclusión en los modelos 1 a 7 se enumeran en la Tabla 11A (Figura 11A). La Tabla también muestra qué alelos de SNP están asociados con un mal pronóstico para cada fenotipo. Las variables de SNP incluidas en los modelos se enumeran en más detalle en la Figura 11B (Tabla 11B).

[0045] El modelo 1 permite la discriminación entre pacientes que tienen un cuestionario de evaluación de la salud de más de 2 y de menos de 2. El modelo 2 permite la discriminación entre pacientes que tienen o no múltiples erosiones en manos y pies. El modelo 3 permite la discriminación entre pacientes que reciben tres o más tratamientos diferentes y pacientes que reciben menos de tres tratamientos. El modelo 4 permite la discriminación entre pacientes cuya enfermedad les obliga a dejar su trabajo y pacientes que no se encuentran en este caso. El modelo 5 permite la discriminación entre pacientes que precisan y que no precisan una intervención quirúrgica (prótesis articular). El modelo 6 permite la discriminación de pacientes entre pacientes que tienen y que no tienen intolerancia al metotrexato. El modelo 7 permite la discriminación de pacientes entre pacientes que consiguen y no que consiguen una remisión.

[0046] Las Figuras 4 a 10 muestran las curvas de ROC, la sensibilidad, la especificidad y los cocientes de verosimilitud positiva (LR+) de todos los modelos desarrollados por los inventores.

[0047] Las Tablas 4 a 10 muestran el cálculo de funciones de probabilidad usando las variables de SNP y clínicas de discriminación para cada uno de los modelos. Las funciones de probabilidad de regresión se construyen usando el paquete estadístico para las ciencias sociales (SPSS Inc. Headquarters, Chicago, IL, Estados Unidos) Versión 14.0, SPSS Versión 14. B es el coeficiente asociado a cada genotipo en la función de probabilidad. ET es el error en el cálculo de B. Wald es la prueba estadística. GL son los grados de libertad. Sig .P es el valor de B para la prueba de Wald. EXP (B) es el riesgo relativo.

[0048] Las variables clínicas y de SNP identificadas y los modelos construidos usándolas proporcionan nuevos medios para predecir el desarrollo de cada uno de los fenotipos correspondientes en un sujeto. Por lo tanto, la invención proporciona métodos para el pronóstico de la RA y, en particular, para predecir el riesgo de desarrollar un fenotipo agresivo de RA, un fenotipo de intolerancia a metotrexato y un fenotipo de remisión, como se describen en la presente memoria.

[0049] La RA agresiva puede estar representada clínicamente por uno de varios fenotipos de presentación, y las referencias a la predicción de RA agresiva pueden referirse a predecir la verosimilitud de desarrollar uno o más de estos fenotipos.

[0050] Los fenotipos de presentación para la RA agresiva se enumeran a continuación.

HAQ>2

[0051] El cuestionario de evaluación de la salud (HAQ) (publicado en Sokka T, Krishnan E, Häkkinen A, Hannonen P, Arthritis Rheum. Ene 2003; 48(1): 59-63) mide la capacidad de los pacientes con RA para llevar a cabo tareas normales en la vida diaria. Una puntuación de HAQ de más de 2 (HAQ>2) indica que la enfermedad RA está avanzando relativamente rápidamente, y que la RA está causando una incapacidad funcional importante. Se conocen en la técnica métodos para llevar a cabo un HAQ y calcular la puntuación del HAQ.

Múltiples erosiones en manos y pies (RX2)

[0052] La presencia de múltiples erosiones en las manos y en los pies puede medirse por inspección visual de rayos x de pacientes. Cuanto mayor sea el número de erosiones, más agresiva se considera que es la enfermedad (Rossi F, Di Dia F, Galipò O, Pistorio A, Valle M, Magni-Manzoni S, Ruperto N, Tomà P, Martini A, Ravelli A., Arthritis

Rheum. Oct 2006 15; 55(5): 717-23). Típicamente, se considera que un paciente presenta el fenotipo de múltiples erosiones si las erosiones están afectando a múltiples articulaciones y se consideran graves.

Recibir ≥3 tratamientos diferentes (TT0_30M)

5 **[0053]** En general, se considera que un paciente presenta el fenotipo de recibir 3 o más tratamientos diferentes si el paciente no mejora después de dos tratamientos diferentes y tiene que prescribirse otro.

Obligación de dejar el trabajo (dejar el trabajo)

10 **[0054]** Los médicos en general consideran que una persona que padece una RA agresiva es más probable que tenga que dejar su trabajo. Se considera que un paciente presenta el fenotipo de dejar el trabajo si él o ella se ve forzado a dejar su empleo debido a la RA.

Recibir una prótesis articular (intervención quirúrgica)

15 **[0055]** La RA agresiva requiere normalmente una intervención quirúrgica en las articulaciones para insertar prótesis articulares (Verstappen SM, Hoes JN, Ter Borg EJ, Bijlsma JW, Blaauw AA, van Albada-Kuipers GA, van Booma-Frankfort C, Jacobs JW., Ann Rheum Dis. Nov 2006; 65(11): 1506-11). Se considera que un paciente presenta este fenotipo si recibe una prótesis articular.

20 **[0056]** Uno o más de cualquiera de los modelos para predecir estos fenotipos de presentación puede considerarse como un modelo para predecir la RA agresiva.

25 **[0057]** Otro fenotipo que puede predecirse de acuerdo con la invención es la probable respuesta a la terapia. En particular, la predicción de probable intolerancia a metotrexato. En general, la intolerancia a metotrexato se ensaya clínicamente ensayando los efectos tóxicos relacionados con la respuesta a la dosis, tales como intolerancia gastrointestinal, anomalías hematológicas, alopecia, hepatotoxicidad y toxicidad pulmonar. Aproximadamente del 30 al 90% de los pacientes a los que se administra metotrexato mostrarán intolerancia (McKendry RJ, Dale P., J Rheumatol. Nov 1993; 20(11): 1850-6). En general, se considera que un paciente presenta el fenotipo de metotrexato si muestra un efecto secundario relacionado con la administración del fármaco.

30 **[0058]** La invención también proporciona medios para predecir la verosimilitud de la remisión en pacientes con RA. En general, se considera que un sujeto tiene el fenotipo de remisión si hay una ausencia completa de dolor en las articulaciones en ausencia de terapia durante al menos cinco años.

35 **[0059]** En general, el sujeto en los presentes métodos es un ser humano. El sujeto puede ser, por ejemplo, chino, japonés o caucásico. Preferentemente, el sujeto es un caucásico tal como un individuo español. El sujeto puede ser masculino o femenino, incluso si para la artritis reumatoide la proporción de mujeres respecto a hombres es de 3:1.

40 **[0060]** Preferentemente, el sujeto cumple los criterios clínicos para el diagnóstico de artritis reumatoide, según se decidieron por el Colegio Americano de Reumatología en 1987 (criterios del ACR 1987). Estos criterios se han publicado por Arnette et al (Arthritis and Rheumatism, 1988). Al sujeto ya se le ha diagnosticado RA.

45 **[0061]** Típicamente, el presente método puede usarse para pronosticar la verosimilitud del desarrollo de uno o más de los fenotipos descritos en la presente memoria, por ejemplo, en cualquier momento durante los 5 años después del diagnóstico.

50 **[0062]** Los presentes métodos de pronóstico implican determinar un resultado para uno o más predictores o variables de polimorfismo de un solo nucleótido (SNP). Las variables de SNP se enumeran en las Tablas 11A y A11B. Los SNP incluidos en los modelos 1 a 7 se enumeran en la Tabla 11A. Los códigos de RefSNP (nº rs) para cada SNP se toman de la Base de Datos de Polimorfismos de un Solo Nucleótido (dbSNP) corregida por el National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=snp>, el 22 de junio de 2007). Obsérvese que no están asignados números rs a alelos NULOS.

55 **[0063]** Un resultado para una variable de SNP dada es la identidad del nucleótido en esa posición en la secuencia de ADN genómico de un sujeto, o el genotipo del sujeto en ese SNP. Por lo tanto, un resultado para un SNP dado puede ser A, T, C o G.

60 **[0064]** La Tabla 11B enumera los polimorfismos que existen en cada uno de los SNP incluidos en los presentes modelos, y la Tabla 11A enumera, para cada SNP, el polimorfismo o alelo que está asociado con un mal pronóstico para cada uno de los fenotipos.

65 **[0065]** Los inventores descubrieron que por determinación de resultados para estos SNP informativos (es decir, identidades de nucleótidos en los SNP), o combinaciones particulares de los mismos, es posible evaluar la

verosimilitud de un fenotipo particular, por ejemplo, RA agresiva, intolerancia a metotrexato o remisión, en un sujeto.

[0066] Los presentes métodos de pronóstico también pueden comprender determinar un resultado para una o más variables clínicas para un sujeto. Estas variables clínicas también se enumeran en la Tabla 11A, junto con el resultado clínico para cada variable que está asociado con un mal pronóstico.

[0067] El ANTI PCC es una variable clínica que se refiere al nivel de anticuerpo anti-péptido cíclico citrulinado en el suero del paciente. Esta es una medida en U/ml y está en el intervalo de 0-1700 U/ml. En los modelos, esta medida es una variable continua que se incluye en la función de regresión logística. Por lo tanto, un resultado para esta variable son las unidades por mililitro medidas en el suero del paciente.

[0068] El ANTI PCC BI es una representación binaria de la variable clínica ANTI PCC continua. Típicamente, los niveles en suero de anticuerpo se miden y se determina si el nivel de anticuerpo está por encima o por debajo de un valor umbral dado. Si el nivel está por encima del umbral, éste se clasifica como elevado (1). Si el nivel está por debajo del umbral, éste se clasifica como reducido o ausente (0). Por lo tanto, un resultado para esta variable es de 0 ó 1.

[0069] La sedimentación eritrocitaria (VSG) es una variable continua y se refiere al tiempo que le lleva a la parte sólida de la sangre sedimentar a partir de la parte líquida. Esto puede medirse (típicamente en mm/h) usando técnicas convencionales (Vives-Corrons J.L., Jou JM. Sangre (Barcelona). 1982; 27(4A): 573-8). En los modelos esta medida es una variable continua que se incluye en la función de regresión logística. Por lo tanto, un resultado para esta variable es un valor en mm/h.

[0070] El número de articulaciones implicadas (NoARTDOL) es una variable continua que se refiere al número de articulaciones que están causando dolor al paciente. Típicamente se evalúa por interrogatorio del paciente. En los modelos, esta medida es una variable continua que se incluye en la función de regresión logística. Por lo tanto, un resultado para esta variable es el número de articulaciones dolorosas mencionadas por los pacientes.

[0071] El ensayo de Factor Reumatoide (RF) se usa principalmente para ayudar a diagnosticar la RA y para distinguirla de otras formas de artritis y otras afecciones que causan síntomas similares de dolor, inflamación y rigidez articular. El ensayo comprende determinar el nivel de anticuerpo RF (típicamente en unidades/ml (U/ml)) en el suero de un paciente. Para los fines de los presentes modelos, esto es una variable continua (Greiner A, Plischke H, Kellner H, Gruber R. Ann N Y Acad Sci. Jun 2005; 1050: 295-303). En los modelos, esta medida es una variable continua que se incluye en la función de regresión logística. Por lo tanto, un resultado para esta variable es de U/ml de anticuerpos medidas en el suero del paciente.

[0072] La PCR es una medida de la cantidad de proteína C reactiva en la sangre. La concentración de proteína C reactiva en la sangre está relacionada con infecciones e inflamaciones. El ensayo de PCR de laboratorio mide la concentración de la proteína en el suero en miligramos por litro (mg/l) (Miller VM, Redfield MM, McConnell JP. Curr Vasc Pharmacol. Ene 2007; 5(1): 15-25). En los presentes modelos, esta medida se ha tratado como una variable continua, que se incluye en la función de regresión logística. Por lo tanto, un resultado para esta variable es la cantidad de proteína medida en la sangre del paciente.

[0073] La edad de aparición se refiere a la edad en años a la que se le diagnosticó al paciente RA de acuerdo con el ACR 1987. En los presentes modelos, esta medida se ha tratado como una variable continua, que se incluye en la función de regresión logística de los modelos. Por lo tanto, un resultado para esta variable es la edad del paciente cuando se le diagnosticó RA.

[0074] La Tabla 11A muestra qué variables de SNP y variables clínicas se incluyen en cada uno de los siete modelos para pronosticar fenotipos. Como se usan en la presente memoria, las variables de "(nombre de fenotipo)" para un fenotipo particular son las variables de SNP y variables clínicas seleccionadas de las de la primera columna de la Tabla 11A, que se incluyen en el modelo para pronosticar ese fenotipo, y que son informativas para pronosticar la verosimilitud de que se desarrolle el fenotipo. Por ejemplo, las "variables de HAQ>2" son las variables de SNP y variables clínicas seleccionadas de las de la primera columna de la Tabla 11A, que se incluyen en el modelo de HAQ>2 y que son informativas para pronosticar la verosimilitud del fenotipo de HAQ>2 (es decir, FGB rs1800790, IL4 rs2070874, ANTI-PCC y VSG).

[0075] Para cada una de las variables incluidas en cada modelo de fenotipo, la Tabla 11A indica qué resultado (alelo de SNP o resultado clínico) está asociado con o sugiere un mal pronóstico para ese fenotipo.

[0076] Por consiguiente, la invención proporciona en un aspecto un método para predecir el curso probable de la RA en un sujeto, que comprende la etapa de determinar u obtener, para ese sujeto, resultados para una o más variables de SNP y una o más variables clínicas enumeradas en las Tablas 11A y 11B. Predecir el curso de la RA puede referirse, en particular, a predecir uno o más de los fenotipos descritos en la presente memoria.

[0077] Los métodos incluyen un método para predecir la RA agresiva en un sujeto. Esto puede realizarse por determinación de la verosimilitud de que el sujeto desarrolle uno o más de los fenotipos de RA agresiva descritos en la presente memoria.

5 **[0078]** Por ejemplo, un método puede comprender determinar la verosimilitud de que un sujeto desarrolle el fenotipo de HAQ>2 descrito en la presente memoria. Típicamente, el método comprende determinar u obtener para el sujeto un resultado para cada una de las variables de HAQ>2 enumeradas en la Tabla 11A (variables del modelo 1). El uso de estas variables permite el pronóstico del fenotipo de HAQ>2 en una población española con un LR+ de 8 (véase el Ejemplo 2 y la Figura 4). Se proporcionan detalles para el cálculo de una función de probabilidad usando estas variables en la Tabla 4.

10 **[0079]** Por ejemplo, un método puede comprender determinar la verosimilitud de que un sujeto desarrolle el fenotipo de múltiples erosiones descrito en la presente memoria. Típicamente, el método comprende determinar u obtener para el sujeto un resultado para cada una de las variables de múltiples erosiones enumeradas en la Tabla 11A (variables del modelo 2). El uso de estas variables permite el pronóstico del fenotipo de múltiples erosiones en una población española con un LR+ de 6,5 (véase el Ejemplo 2 y la Figura 6). Se proporcionan detalles para el cálculo de una función de probabilidad usando estas variables en la Tabla 6.

15 **[0080]** Por ejemplo, un método puede comprender determinar la verosimilitud de que un sujeto desarrolle el fenotipo de >3 tratamientos diferentes descrito en la presente memoria. Típicamente, el método comprende determinar u obtener para el sujeto un resultado para cada una de las variables de >3 tratamientos diferentes enumeradas en la Tabla 11A (variables del modelo 3). El uso de estas variables permite el pronóstico del fenotipo de >3 tratamientos diferentes en una población española con un LR+ de 6,2 (véase el Ejemplo 2 y la Figura 7). Se proporcionan detalles para el cálculo de una función de probabilidad usando estas variables en la Tabla 7.

20 **[0081]** Por ejemplo, un método puede comprender determinar la verosimilitud de que un sujeto desarrolle el fenotipo de dejar el trabajo descrito en la presente memoria. Típicamente, el método comprende determinar u obtener para el sujeto un resultado para cada una de las variables de dejar el trabajo enumeradas en la Tabla 11A (variables del modelo 4). El uso de estas variables permite el pronóstico del fenotipo de dejar el trabajo en una población española con un LR+ de 1,9 (véase el Ejemplo 2 y la Figura 9). Se proporcionan detalles para el cálculo de una función de probabilidad usando estas variables en la Tabla 9.

25 **[0082]** Por ejemplo, un método puede comprender determinar la verosimilitud de que un sujeto desarrolle el fenotipo de intervención quirúrgica descrito en la presente memoria. Típicamente, el método comprende determinar u obtener para el sujeto un resultado para cada una de las variables de intervención quirúrgica enumeradas en la Tabla 11A (variables del modelo 5). El uso de estas variables permite el pronóstico del fenotipo de intervención quirúrgica en una población española con un LR+ de 5,4 (véase el Ejemplo 2 y la Figura 8). Se proporcionan detalles para el cálculo de una función de probabilidad usando estas variables en la Tabla 8.

30 **[0083]** Un método puede comprender llevar a cabo uno o más de los ensayos anteriores para el desarrollo de fenotipos agresivos. Por ejemplo, un método puede comprender llevar a cabo al menos 2, 3, 4 o los 5 métodos de ensayo anteriores, usando las variables de modelo correspondientes. Por ejemplo, un método puede comprender al menos evaluar la verosimilitud de los fenotipos de HAQ>2, múltiples erosiones, >3 tratamientos diferentes e intervención quirúrgica, o los fenotipos de HAQ>2, múltiples erosiones y >3 tratamientos diferentes, o los fenotipos de HAQ>2 y múltiples erosiones. En un ejemplo, el método comprende al menos predecir la verosimilitud del fenotipo de HAQ>2. Las variables ensayadas se seleccionarán por consiguiente en línea con los modelos anteriores y con la información de la Tabla 11A.

35 **[0084]** Los presentes métodos también incluyen un método para predecir la verosimilitud de la intolerancia a metotrexato en un paciente. Típicamente, el método comprende determinar u obtener para el sujeto un resultado para cada una de las variables de respuesta a la terapia enumeradas en la Tabla 11A (variables del modelo 6). El uso de estas variables permite el pronóstico del fenotipo de intolerancia a metotrexato en una población española con un LR+ de 4,3 (véase el Ejemplo 2 y la Figura 10). Se proporcionan detalles para el cálculo de una función de probabilidad usando estas variables en la Tabla 10.

40 **[0085]** Los presentes métodos también incluyen un método para predecir la verosimilitud de la remisión en un paciente. Típicamente, el método comprende determinar u obtener para el sujeto un resultado para cada una de las variables de remisión enumeradas en la Tabla 11A (variables del modelo 7). El uso de estas variables permite el pronóstico del fenotipo de remisión en una población española con un LR+ de 6,5 (véase el Ejemplo 2 y la Figura 5). Se proporcionan detalles para el cálculo de una función de probabilidad usando estas variables en la Tabla 5.

45 **[0086]** Un sujeto puede ensayarse para determinar el probable desarrollo de uno o más de los fenotipos anteriores. Por ejemplo, un sujeto puede ensayarse para determinar la verosimilitud de desarrollar una RA agresiva (mediante uno o más de cualquiera de los métodos descritos anteriormente) en combinación con un ensayo para determinar la verosimilitud de desarrollar intolerancia a metotrexato y/o remisión. De forma similar, un sujeto puede

ensayarse para determinar la probable remisión, así como la intolerancia a metotrexato. Puede usarse cualquier combinación de los métodos anteriores, de modo que un sujeto puede evaluarse simultáneamente para determinar el probable desarrollo de más de un fenotipo. Las variables ensayadas se seleccionarán por consiguiente en línea con los modelos anteriores y la información de la tabla 11A.

[0087] En un ejemplo, un método puede comprender al menos evaluar la verosimilitud de los fenotipos de HAQ>2, múltiples erosiones, remisión, >3 tratamientos diferentes, intervención quirúrgica y respuesta a la terapia, o los fenotipos de HAQ>2, múltiples erosiones, remisión, >3 tratamientos diferentes e intervención quirúrgica, o los fenotipos de HAQ>2, múltiples erosiones, remisión y >3 tratamientos diferentes, o los fenotipos de HAQ>2, múltiples erosiones y remisión. En un ejemplo, el método comprende al menos predecir la verosimilitud de los fenotipos de HAQ>2 y remisión.

[0088] En un ejemplo, un método de la invención comprende determinar la verosimilitud de desarrollar el fenotipo de HAQ>2, y determinar la verosimilitud de desarrollar el fenotipo de múltiples erosiones, y determinar la verosimilitud de desarrollar el fenotipo de >3 tratamientos diferentes, y determinar la verosimilitud de desarrollar el fenotipo de dejar el trabajo, y determinar la verosimilitud de desarrollar el fenotipo de intervención quirúrgica, y determinar la verosimilitud de desarrollar la intolerancia a metotrexato y determinar la verosimilitud de remisión en un sujeto. En ese caso típicamente, se determinan resultados para todas las variables de la Tabla 11A.

[0089] En algunos aspectos los presentes métodos pueden incluir determinar otros factores para un sujeto. Por ejemplo, el sujeto puede genotiparse para una o más variaciones genéticas distintas (tales como otros SNP no enumerados en la Tabla 11A u 11B). Estos pueden ser mutaciones asociadas con RA u otra afección. Por ejemplo, un sujeto puede genotiparse en uno o más de los SNP restantes enumerados en la Tabla 1A o 1B, por ejemplo, usando la micromatriz Artchip descrita en la presente memoria. También pueden determinarse otros marcadores (por ejemplo, SNP) asociados con otras enfermedades.

[0090] Los presentes métodos pueden usarse junto con o además de ensayos clínicos convencionales.

[0091] Los presentes métodos permiten una predicción con precisión de los fenotipos de RA basándose en un número relativamente pequeño de SNP informativos y variables clínicas. Esto puede ser ventajoso en el sentido de que permite el uso de técnicas de genotipado que no serían necesariamente adecuadas para una exploración de SNP a gran escala, así como de métodos de genotipado a mayor escala.

[0092] En general, incluso si se ensayan en los presentes métodos un mayor número de SNP o variaciones o factores genéticos, la predicción de un fenotipo de RA puede realizarse basándose solamente en los resultados para las variables del modelo de fenotipo correspondiente en la Tabla 11A. Estas variables son suficientes para la predicción. Por lo tanto, en un ejemplo, los presentes métodos permiten el pronóstico diferencial de los fenotipos de RA descritos en la presente memoria, basándose en (como máximo) los resultados para las variables de los modelos correspondientes en la Tabla 11A.

[0093] En algunos casos sin embargo, puede ser que se usen en la predicción algunas variables adicionales tales como SNP u otros factores. Por ejemplo, en los presentes métodos, el pronóstico puede realizarse basándose en los resultados de un máximo de 100, 90, 80, 70, 60, 50, 45, 40, 35, 30, 29, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3 ó 2 variables, tales como SNP o SNP asociados a RA. Los SNP pueden comprender (o consistir en) o seleccionarse de las variables de la Tabla 11A o de las variables de SNP.

[0094] En un aspecto, el método puede implicar genotipar un máximo de 100, 90, 80, 70, 60, 50, 40, 30, 25, 20, 19, 18, 17, 16, 15, 14, 13, 12, 11, 10, 9, 8, 7, 6, 5, 4, 3 ó 2 SNP o SNP asociados a RA. El método puede implicar genotipar un máximo de (no más de) todos los SNP de la Tabla 11A, la Tabla 1A o la Tabla 1B. En algunos casos, el método comprende genotipar como máximo las variables de SNP para uno o más de los modelos 1 a 7, seleccionados como se ha descrito anteriormente.

[0095] Preferentemente, el número y la combinación de variables tales como SNP usadas para construir un modelo para predecir un fenotipo de acuerdo con la invención es tal que el modelo permite que la predicción se realice con un valor de LR+ de al menos 1,5, tal como de al menos 2, 3, 4, 5, 6, 7, 8, 9 ó 10. El cálculo de valores de LR+ se describe en la presente memoria.

[0096] Una vez que se determina un resultado para cada una de las variables para la predicción de un fenotipo dado, estos resultados se usan en o se insertan en una función de probabilidad adecuada (para la predicción de ese fenotipo), como se describe en la presente memoria, y se calcula un valor de función de probabilidad. Los resultados pueden codificarse para su uso en la función de probabilidad y el cálculo del valor de la función de probabilidad. El valor de la función de probabilidad se compara después con los valores de función de probabilidad obtenidos para una población de individuos de fenotipo conocido (determinado clínicamente). El riesgo de que el sujeto tenga o desarrolle el fenotipo particular se determina de este modo.

- 5 **[0097]** Una función de probabilidad adecuada para determinar un fenotipo dado puede obtenerse por métodos como los expuestos en el Ejemplo 2 y descritos en la presente memoria. Típicamente, se proporciona una población de individuos de estudio. Estos individuos son de un fenotipo conocido (determinado clínicamente) con respecto al fenotipo para cuya determinación se usará la función de probabilidad. El diagnóstico clínico y de RA puede realizarse siguiendo los criterios del Colegio Americano de Reumatología (1987) (ACR 1987) como se describe en la presente memoria. Preferentemente, los individuos en la población de estudio cumplen la definición del ACR de 1987. Cada uno de los fenotipos de RA que pueden evaluarse de acuerdo con la invención puede diagnosticarse clínicamente como se describe en la presente memoria.
- 10 **[0098]** En un ejemplo, los individuos en la población de estudio pueden cumplir uno o más, por ejemplo, 2, 3, 4, 5 o los 6 criterios de inclusión completos para la población de estudio en el Ejemplo 2: cumplir los criterios del ACR de 1987 para el diagnóstico de RA; tener una aparición de RA después de la edad de 18; tener erosiones radiológicas (daños a nivel de las articulaciones afectadas por RA); ser caucásico; tener más de 5 años de evolución; haber sido diagnosticado después del 1 de enero de 1990.
- 15 **[0099]** La población puede ser, por ejemplo, una población china, japonesa o caucásica, tal como una población española. La población puede comprender mujeres y hombres. En un ejemplo, la proporción de mujeres respecto a hombres puede ser de aproximadamente 3:1. Preferentemente, la población usa para obtener una función de probabilidad comprende una muestra representativa de la población en la que se aplicará la función de probabilidad.
- 20 **[0100]** En general, se incluyen al menos n individuos en la población de estudio. Típicamente, n es 200-1000, por ejemplo, 300, 400, 500 ó 600. Cuando una función de probabilidad es para determinar entre fenotipos alternativos, preferentemente hay números aproximadamente equivalentes de individuos con cada uno de los fenotipos alternativos en la población. Por lo tanto, cuando hay dos fenotipos alternativos, A y B, la población es preferentemente aproximadamente el 50% de fenotipo A y el 50% de fenotipo B. Sin embargo, las proporciones pueden ser, por ejemplo, 60%/40%, 70%/30%, 80%/20%, 90%/10% o cualquier distribución estadísticamente aceptable. Por ejemplo, cuando la función de probabilidad es para el pronóstico de $H_{AQ} > 2$ frente a $H_{AQ} \leq 2$, preferentemente aproximadamente el 50% de la población es de fenotipo $H_{AQ} > 2$ determinado clínicamente y aproximadamente el 50% de la población es de fenotipo $H_{AQ} \leq 2$ determinado clínicamente. Cuando la función de probabilidad es para el pronóstico de múltiples erosiones, preferentemente aproximadamente el 50% de la población cumple los criterios clínicos para múltiples erosiones y aproximadamente el 50% de la población no los cumple.
- 25 **[0101]** Cada individuo en la población de estudio se ensaya después para determinar un resultado para cada una de las variables de discriminación para el fenotipo particular (véase la Tabla 11A). Esto proporciona varios resultados para cada individuo. El ensayo, por ejemplo, el genotipado, puede llevarse a cabo mediante cualquiera de los métodos descritos en la presente memoria, por ejemplo, mediante análisis de micromatrices como se describe en la presente memoria. El ensayo es típicamente *ex vivo*, llevado a cabo en una muestra adecuada obtenida de un individuo.
- 30 **[0102]** Después pueden analizarse múltiples asociaciones de genotipo-fenotipo usando un análisis de regresión logística multivariable por etapas, usando como variable dependiente el fenotipo de RA determinado clínicamente y como variables independientes los resultados de las variables informativas, por ejemplo, como se recomienda por Balding DJ. (2006³⁵). La calidad de ajuste de los modelos obtenidos puede evaluarse usando estadística de Hosmer-Lemeshow, y su precisión puede evaluarse calculando el área bajo la curva (AUC) de la curva de Características Operativas del Receptor (ROC) con intervalos de confianza del 95% (véase, por ejemplo, (Janssens ACJW et al., 2006³⁶). Se describen métodos adecuados en el Ejemplo 2.
- 35 **[0103]** La sensibilidad, la especificidad y el cociente de verosimilitud positiva ($LR+ = \text{sensibilidad}/(1-\text{especificidad})$) pueden calcularse por medio de curvas de ROC. Preferentemente, el modelo tiene un valor de $LR+$ de al menos 1,5, por ejemplo de al menos 2, 3, 4, 5, 6, 7, 8, 9 ó 10.
- 40 **[0104]** Los valores medios de la función de probabilidad para cada uno de los fenotipos alternativos en la población pueden compararse usando una prueba t . En general, las funciones de probabilidad son capaces de distinguir entre los diferentes fenotipos en la población de estudio de una forma estadísticamente significativa, por ejemplo, a $p \leq 0,05$ en una prueba t . Por lo tanto, las funciones de probabilidad producen una separación estadísticamente significativa entre individuos de diferente fenotipo en la población.
- 45 **[0105]** Pueden realizarse análisis estadísticos, por ejemplo, usando el Paquete Estadístico para las Ciencias Sociales (SPSS Inc. Headquarters, Chicago, IL, Estados Unidos) versión 14.0.
- 50 **[0106]** Los valores de función de probabilidad pueden calcularse para cada individuo de fenotipo conocido en la población de estudio y representarse en una gráfica adecuada.
- 55 **[0107]** Para llevar a cabo los presentes métodos de pronóstico, se calcula un valor de función de probabilidad para el individuo de ensayo, y éste se compara con los valores de función de probabilidad para los individuos de fenotipo

conocido en la población de estudio para determinar el riesgo de un fenotipo dado en ese individuo. La comparación puede realizarse por comparación con una gráfica o por cualquier otro medio adecuado conocido por los expertos en la materia.

5 **[0108]** Por lo tanto, por ejemplo, en la obtención de una función de probabilidad para su uso en el pronóstico diferencial de HAQ>2 frente a HAQ<2, se proporciona una población de individuos de estudio diagnosticados clínicamente como HAQ>2 e individuos diagnosticados clínicamente como HAQ≤2. Cada individuo puede ensayarse después para determinar un resultado para cada una de las variables de discriminación de HAQ>2 (modelo 1) en la Tabla 11A. Se realiza una regresión logística múltiple por etapas sobre los datos de “resultados” y “fenotipo” y se obtiene una función de probabilidad que es capaz de distinguir entre los dos grupos fenotípicos en la población de estudio de una forma estadísticamente significativa.

10 **[0109]** Por lo tanto, en un aspecto, la invención proporciona además un método de obtención de una función de probabilidad para su uso en la determinación de un fenotipo de RA como se describe en la presente memoria, que comprende:

- 15 (i) proporcionar una población de individuos de estudio, en la que cada individuo es de un fenotipo conocido determinado clínicamente con respecto al fenotipo de RA;
- 20 (ii) determinar u obtener para cada individuo un resultado para cada una de un conjunto de variables, obteniendo de este modo un conjunto de resultados para cada individuo;
- (iii) aplicar un análisis de regresión logística múltiple por etapas a los resultados obtenidos en (ii) y a los fenotipos conocidos mencionados en (i); y
- (iv) obtener de este modo una función de probabilidad que produce una separación estadísticamente significativa entre individuos de diferente fenotipo en la población;

25 en el que:

- 30 (a) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de HAQ>2 de acuerdo con la invención, y el conjunto de variables para las que se determinan u obtienen resultados en la etapa (ii) se selecciona de o consiste en las variables de HAQ>2 de la Tabla 11A;
- (b) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de múltiples erosiones de acuerdo con la invención, y el conjunto de variables para las que se determinan u obtienen resultados en la etapa (ii) se selecciona de o consiste en las variables de múltiples erosiones de la Tabla 11A;
- 35 (c) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de >3 tratamientos diferentes de acuerdo con la invención, y el conjunto de variables para las que se determinan u obtienen resultados en la etapa (ii) se selecciona de o consiste en las variables de >3 tratamientos diferentes en la Tabla 11A;
- (d) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de dejar el trabajo de acuerdo con la invención, y el conjunto de variables para las que se determinan u obtienen resultados en la etapa (ii) se selecciona de o consiste en las variables de dejar el trabajo de la Tabla 11A;
- 40 (e) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de intervención quirúrgica de acuerdo con la invención, y el conjunto de variables para las que se determinan u obtienen resultados en la etapa (ii) se selecciona de o consiste en las variables de intervención quirúrgica de la Tabla 11A;
- 45 (f) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de intolerancia a metotrexato de acuerdo con la invención, y el conjunto de variables para las que se determinan u obtienen resultados en la etapa (ii) se selecciona de o consiste en las variables de HAQ>2 de la Tabla 11A; y/o
- (g) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de HAQ>2 de acuerdo con la invención, y el conjunto de variables para las que se determinan u obtienen resultados en la etapa (ii) se selecciona de o consiste en las variables de HAQ>2 de la Tabla 11A.

50 **[0110]** La obtención de las funciones de probabilidad puede llevarse a cabo mediante un ordenador. Por lo tanto, en un aspecto, la invención también se refiere a un método computacional de obtención de una función de probabilidad para su uso en la determinación de un fenotipo de RA, comprendiendo dicho método aplicar un análisis de regresión logística múltiple por etapas a los datos de resultados y los datos de fenotipo obtenidos de una población de individuos de estudio adecuada, en el que cada individuo es de un fenotipo conocido determinado clínicamente con respecto al fenotipo de RA, obteniendo de este modo una función de probabilidad que produce una separación estadísticamente significativa entre individuos de diferente fenotipo en la población;

60 en el que:

- (i) los datos de fenotipo comprenden el fenotipo conocido determinado clínicamente de cada individuo;
- (ii) los datos de resultados para cada individuo comprenden resultados para una o más variables de polimorfismos de un solo nucleótido y una o más variables clínicas enumeradas en la columna 1 de la Tabla 11A;

65 y en el que:

(a) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de HAQ>2 de acuerdo con la invención, y las variables para las que se obtienen datos de resultados (y mencionadas en (ii)) comprenden o consisten en las variables de HAQ>2 de la Tabla 11A;

5 (b) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de múltiples erosiones de acuerdo con la invención, y las variables para las que se obtienen datos de resultados (y mencionadas en (ii)) comprenden o consisten en las variables de múltiples erosiones de la Tabla 11A;

(c) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de >3 tratamientos diferentes de acuerdo con la invención, y las variables para las que se obtienen datos de resultados (y mencionadas en (ii)) comprenden o consisten en las variables de >3 tratamientos diferentes de la Tabla 11A;

10 (d) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de dejar el trabajo de acuerdo con la invención, y las variables para las que se obtienen datos de resultados (y mencionadas en (ii)) comprenden o consisten en las variables de dejar el trabajo de la Tabla 11A;

15 (e) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de intervención quirúrgica de acuerdo con la invención, y las variables para las que se obtienen datos de resultados (y mencionadas en (ii)) comprenden o consisten en las variables de intervención quirúrgica de la Tabla 11A;

(f) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de intolerancia a metotrexato de acuerdo con la invención, y las variables para las que se obtienen datos de resultados (y mencionadas en (ii)) comprenden o consisten en las variables de HAQ>2 de la Tabla 11A; y/o

20 (g) la función de probabilidad es para pronosticar la verosimilitud de desarrollar el fenotipo de HAQ>2 de acuerdo con la invención, y las variables para las que se obtienen datos de resultados (y mencionadas en (ii)) comprenden o consisten en las variables de HAQ>2 de la Tabla 11A.

[0111] Se han descrito anteriormente poblaciones de estudio y métodos de análisis estadístico adecuados. También puede hacerse referencia a los presentes Ejemplos.

25 [0112] Se proporcionan detalles para el cálculo de una función de probabilidad a partir de las variables de SNP y clínicas enumeradas para cada fenotipo en las Tablas 4 a 10. Pueden realizarse análisis estadísticos, por ejemplo, usando el Paquete Estadístico para las Ciencias Sociales (SPSS Inc. Headquarters, Chicago, IL, Estados Unidos) versión 14.0. Estos pueden usarse para el cálculo de valores de función de probabilidad para su uso en los métodos de la presente memoria. Los datos de las Tablas pueden usarse para construir funciones de probabilidad para su uso en la invención. Las funciones de probabilidad, junto con la información de las Tablas 11A y 11B, pueden usarse para determinar un diagnóstico o pronóstico de acuerdo con la invención.

35 [0113] En un aspecto, la invención se refiere a funciones de probabilidad construidas u obtenidas usando los datos de cualquiera de las Tablas 4 a 10, y a su uso en un método, por ejemplo, un método computacional, para pronosticar un fenotipo de RA. La invención se refiere además a los programas informáticos y sistemas informáticos asociados, como se describen en la presente memoria. La invención también se refiere a las funciones de probabilidad obtenidas de acuerdo con los presentes métodos y a su uso en los métodos descritos en la presente memoria.

40 [0114] El procedimiento de cálculo de un valor de función de probabilidad para un sujeto de ensayo y de comparación del valor con los valores obtenidos a partir de una población de individuos de estudio de fenotipos conocidos para evaluar el riesgo de desarrollar un fenotipo en el sujeto de ensayo también puede llevarse a cabo usando un *software* apropiado.

45 [0115] Por lo tanto, en un aspecto, la invención se refiere a un método computacional para determinar un fenotipo de RA usando los resultados de variables de discriminación (“datos de resultados”) para ese fenotipo obtenido de acuerdo con los métodos descritos en la presente memoria. En el método computacional, los datos de resultados para las variables de discriminación para un fenotipo particular obtenidos a partir de un sujeto de ensayo (datos de resultados de ensayo) se introducen en una función de probabilidad adecuada para producir un valor de función de probabilidad para el sujeto de ensayo. El valor de la función de probabilidad de ensayo se compara después con los valores de función de probabilidad para individuos de fenotipo conocido, para diagnosticar o pronosticar el fenotipo del individuo de ensayo. La comparación puede realizarse usando los métodos descritos en la presente memoria.

55 [0116] La invención se refiere además a un sistema informático que comprende un procesador y medios para controlar el procesador, para llevar a cabo un método computacional descrito en la presente memoria, y a un programa informático que comprende un código de programa informático que cuando se ejecuta en un ordenador o red de ordenadores provoca que el ordenador o red de ordenadores lleven a cabo el método computacional. En un aspecto, el programa informático se almacena en un medio legible por ordenador.

60 [0117] Como se ha descrito anteriormente y en los Ejemplos, los presentes inventores también han identificado varios polimorfismos de un solo nucleótido (SNP) que muestran la asociación alélica de un solo locus con los fenotipos de RA descritos en la presente memoria. Los SNP que muestran una asociación estadísticamente significativa se enumeran en la Tabla 12 (Figura 12). Los valores P se calcularon mediante la prueba de Chi cuadrado usando los *software* SPSS y HelixTree.

65

[0118] Por identificación del nucleótido en el ADN genómico de un sujeto en uno (o más) de estos SNP, es posible evaluar el riesgo o la susceptibilidad de ese individuo al fenotipo dado.

5 **[0119]** En un aspecto, la invención se refiere al uso de uno más de los SNP de la Tabla 12 en un método para pronosticar un fenotipo de RA, en particular para determinar la verosimilitud del desarrollo de un fenotipo de RA con el que el o los SNP están estadísticamente asociados de forma significativa (Tabla 12), como se describe en la presente memoria.

10 **[0120]** Por lo tanto, la invención se refiere en un aspecto a un método para pronosticar el fenotipo de HAQ>2 (como se describe en la presente memoria), que comprende determinar el genotipo de un individuo en uno o más de los SNP de HAQ>2 de la Tabla 12.

15 **[0121]** La invención también se refiere a un método para pronosticar el fenotipo de RX2 (como se describe en la presente memoria), que comprende determinar el genotipo de un individuo en uno o más de los SNP de RX2 de la Tabla 12.

20 **[0122]** La invención también se refiere a un método para pronosticar el fenotipo de TT0_30M (como se describe en la presente memoria), que comprende determinar el genotipo de un individuo en uno o más de los SNP de TT0_3 de la Tabla 12.

[0123] La invención también se refiere a un método para pronosticar el fenotipo de intervención quirúrgica (como se describe en la presente memoria), que comprende determinar el genotipo de un individuo en el SNP de intervención quirúrgica de la Tabla 12.

25 **[0124]** La invención también se refiere a un método para pronosticar el fenotipo de remisión (como se describe en la presente memoria), que comprende determinar el genotipo de un individuo en uno o más de los SNP de remisión de la Tabla 12.

30 **[0125]** La invención también se refiere a un método para pronosticar el fenotipo de dejar el trabajo (como se describe en la presente memoria), que comprende determinar el genotipo de un individuo en uno o más de los SNP de dejar el trabajo de la Tabla 12.

35 **[0126]** La invención también se refiere a un método para pronosticar el fenotipo de intolerancia a metotrexato (como se describe en la presente memoria), que comprende determinar el genotipo de un individuo en uno o más de los SNP de intolerancia a metotrexato de la Tabla 12.

40 **[0127]** En general, los presentes métodos se llevan a cabo *ex vivo* o *in vitro*, por ejemplo, usando una muestra obtenida del individuo. Un método puede comprender el uso de los resultados de variables clínicas que se han obtenido mediante los métodos descritos en la presente memoria.

45 **[0128]** Se conocen en la técnica diversos métodos para determinar la presencia o ausencia en una muestra de ensayo de una secuencia de ácido nucleico particular, por ejemplo, una secuencia de ácido nucleico que tiene un nucleótido particular en una posición de polimorfismo de un solo nucleótido. Por ejemplo, el genotipo puede determinarse mediante análisis de micromatrices, secuenciación, extensión de cebadores, ligación de oligonucleótidos específicos de alelo, determinación de masa de productos de extensión de cebadores, análisis de polimorfismos en la longitud e restricción, análisis de polimorfismos conformacionales de cadenas sencillas, pirosecuenciación, HPLC o electroforesis en gel con gradiente desnaturalizante (DGGE). Además, habiendo secuenciado el ácido nucleico de un individuo o muestra, la información de secuencia puede conservarse y posteriormente buscarse sin recurrir al propio ácido nucleico original. Por lo tanto, por ejemplo, una alteración de secuencia o mutación puede identificarse por exploración de una base de datos de información de secuencia usando un ordenador u otro medio electrónico.

50 **[0129]** En general, se proporciona una muestra que contiene ácido nucleico, que comprende al menos una de las variaciones genéticas a ensayar. El ácido nucleico comprende una o más regiones diana que comprenden la variación o variaciones genéticas (SNP) que deben caracterizarse.

55 **[0130]** El ácido nucleico puede obtenerse a partir de cualquier muestra biológica apropiada que contenga ácido nucleico. La muestra puede tomarse de un fluido o tejido, secreción, célula o línea celular derivada del cuerpo humano.

60 **[0131]** Por ejemplo, pueden tomarse muestras de sangre, incluyendo suero, linfocitos, células linfoblastoides, fibroblastos, plaquetas, células mononucleares u otras células sanguíneas, a partir de saliva, hígado, riñón, páncreas o corazón, orina o de cualquier otro tejido, fluido, célula o línea celular derivada del cuerpo humano. Por ejemplo, una muestra adecuada puede ser una muestra de células de la cavidad bucal.

65

[0132] Preferentemente, el ácido nucleico se obtiene de una muestra de sangre.

[0133] En general, el ácido nucleico se extrae de la muestra biológica usando técnicas convencionales. El ácido nucleico que se extraerá de la muestra biológica puede ser ADN, o ARN, típicamente ARN total. Típicamente, se extrae ARN si la variación genética a estudiar se sitúa en la secuencia codificante de un gen. Cuando se extrae ARN de la muestra biológica, los métodos pueden comprender además una etapa de obtención de ADNc a partir del ARN. Esto puede llevarse a cabo usando métodos convencionales, tales como transcripción inversa usando cebadores adecuados. Los procedimientos posteriores se llevan a cabo después típicamente sobre el ADN extraído o el ADNc obtenido a partir del ARN extraído. El término ADN, como se usa en la presente memoria, puede incluir tanto ADN como ADNc.

[0134] En general, las variaciones genéticas a ensayar son conocidas y están caracterizadas, por ejemplo, en términos de secuencia. Por lo tanto, pueden obtenerse regiones de ácido nucleico que comprenden las variaciones genéticas usando métodos conocidos en la técnica.

[0135] En un aspecto, las regiones de ADN que contienen las variaciones genéticas (SNP) a identificar (regiones diana) se someten a una reacción de amplificación para obtener productos de amplificación que contienen las variaciones genéticas que se van a identificar. Cualquier técnica o método adecuado puede usarse para la amplificación.

[0136] Por ejemplo, puede usarse la reacción en cadena de la polimerasa (PCR) (revisada por ejemplo en "PCR protocols; A Guide to Methods and Applications", Eds. Innis et al, 1990, Academic Press, Nueva York, Mullis et al, Cold Spring Harbor Symp. Quant. Biol., 51: 263, (1987), Ehrlich (ed), PCR technology, Stockton Press, NY, 1989, y Ehrlich et al, Science, 252: 1643-1650, (1991)). El ácido nucleico usado como molde en la reacción de amplificación puede ser ADN genómico, ADNc o ARN.

[0137] Otras técnicas de amplificación de ácido nucleico específicas incluyen activación del desplazamiento de cadena, el sistema de replicasa QB, la reacción en cadena de reparación, la reacción en cadena de la ligasa, la amplificación por círculo rodante y la transcripción activada por ligación.

[0138] Pueden usarse oligonucleótidos específicos de alelo en PCR para amplificar específicamente secuencias particulares si están presentes en una muestra de ensayo. La evaluación de si una banda de PCR contiene una variante génica puede llevarse a cabo de varias formas familiares para los expertos en la materia. El producto de PCR puede tratarse, por ejemplo, de un modo que permita a los inventores visualizar el polimorfismo en un gel de secuenciación de ADN de poli(acrilamida) desnaturalizante, con bandas específicas que están vinculadas a las variantes génicas que se están seleccionando.

[0139] Los expertos en la materia tienen experiencia en el diseño de cebadores para su uso en procedimientos tales como PCR. Son bien conocidas en la materia diversas técnicas para sintetizar cebadores oligonucleotídicos, incluyendo métodos de síntesis de fosfotriéster y fosfodiéster.

[0140] Un aspecto adicional de la presente invención proporciona una pareja de cebadores de amplificación oligonucleotídicos adecuada para su uso en los métodos descritos en la presente memoria.

[0141] Los cebadores de PCR adecuados para la amplificación de regiones de ADN diana que comprenden los SNP de la Tabla 1 se enumeran en la Tabla 3A y en la Tabla 3B. Los presentes métodos pueden comprender el uso de uno o más de estos cebadores o una o más de las parejas de cebadores enumeradas, de acuerdo con los SNP a genotipar, en los que estos SNP se seleccionan como se describe en la presente memoria. En un aspecto, el método comprende el uso de todos los cebadores enumerados en las Tablas 3A y 3B. Las condiciones de reacción adecuadas pueden determinarse usando los conocimientos en la técnica.

[0142] El ácido nucleico amplificado puede secuenciarse después y/o ensayarse de cualquier otro modo para determinar la presencia o ausencia de una característica particular. El ácido nucleico para ensayo puede prepararse a partir de ácido nucleico extraído de células o en una biblioteca usando una diversidad de otras técnicas tales como digestión con enzimas de restricción y electroforesis.

[0143] Por ejemplo, el alelo del al menos un polimorfismo (es decir, la identidad del nucleótido en la posición de polimorfismo de un solo nucleótido) puede determinarse por determinación de la unión de una sonda oligonucleotídica a la región amplificada de la muestra de genómico. Una sonda oligonucleotídica adecuada comprende una secuencia de nucleótidos que se une específicamente a un alelo particular del al menos un polimorfismo y que no se une específicamente a otros alelos del al menos un polimorfismo. Dicha sonda puede corresponder en secuencia a una región de ácido nucleico genómico, o su complementaria, que contiene uno o más de los SNP descritos en la presente memoria. En condiciones convenientemente rigurosas, la hibridación específica de dicha sonda con ácido nucleico de ensayo es indicativa de la presencia de la alteración de secuencia en el ácido nucleico de ensayo. Con fines de exploración eficaces, puede usarse más de una sonda en la misma muestra de

ensayo.

[0144] Los expertos en la materia son bien capaces de emplear condiciones adecuadas de la rigurosidad deseada para una hibridación selectiva, teniendo en cuenta factores tales como la longitud de los oligonucleótidos y la composición de bases, la temperatura, etcétera.

[0145] Las condiciones de hibridación selectivas adecuadas para oligonucleótidos de 17 a 30 bases incluyen la hibridación durante una noche a 42°C en SSC 6X y el lavado en SSC 6X a una serie de temperaturas crecientes desde 42°C hasta 65°C.

[0146] Se describen otras condiciones y protocolos adecuados en Molecular Cloning: a Laboratory Manual: 2ª edición, Sambrook et al., 1989, Cold Spring Harbor Laboratory Press y Current Protocols in Molecular Biology, Ausubel et al. eds., John Wiley & Sons, 1992.

[0147] Un aspecto adicional de la presente invención proporciona un oligonucleótido que hibrida específicamente con una secuencia de ácido nucleico que comprende un alelo particular de un polimorfismo seleccionado del grupo que consiste en los polimorfismos de un solo nucleótido mostrados en la Tabla 1A, 1B o Tabla 18, y que no se unen específicamente a otros alelos del SNP. La hibridación puede determinarse en condiciones de hibridación selectiva adecuadas como se describe en la presente memoria.

[0148] Dichos oligonucleótidos pueden usarse en un método de exploración de ácido nucleico.

[0149] En algunas realizaciones preferidas, los oligonucleótidos de acuerdo con la presente invención tienen una longitud de al menos aproximadamente 10 nucleótidos, más preferentemente una longitud de al menos aproximadamente 15 nucleótidos, más preferentemente una longitud de al menos aproximadamente 20 nucleótidos. Los oligonucleótidos pueden ser de hasta 100 nucleótidos de longitud, más preferentemente de hasta aproximadamente 50 nucleótidos de longitud, más preferentemente de hasta aproximadamente 30 nucleótidos de longitud. El valor límite de "aproximadamente X nucleótidos", como se ha usado anteriormente, incluye el valor límite de "X nucleótidos". Los oligonucleótidos que hibridan específicamente con los particulares de los SNP enumerados en la Tabla 1A se enumeran en la Tabla 2 y se describen en la presente memoria.

[0150] Cuando el ácido nucleico es ADN bicatenario, la hibridación estará generalmente precedida por desnaturalización para producir ADN de cadena sencilla. La hibridación puede ser parte de una amplificación, por ejemplo, por procedimiento de PCR, o parte de un procedimiento de sondaje que no implique amplificación. Un procedimiento de ejemplo sería una combinación de PCR e hibridación de baja rigurosidad. Un procedimiento de exploración, seleccionado de los muchos disponibles para los expertos en la materia, se usa para identificar acontecimientos de hibridación con éxito y ácido nucleico hibridado aislado.

[0151] La unión de una sonda a ácido nucleico diana (por ejemplo, ADN) puede medirse usando cualquiera de una diversidad de técnicas a disposición de los expertos en la materia. Por ejemplo, las sondas pueden marcarse radiactivamente, fluorescentemente o enzimáticamente. Otros métodos que no emplean marcaje de sondas incluyen el examen de polimorfismos en la longitud de fragmentos de restricción, la amplificación usando PCR, la escisión con ARNasa y el sondaje con oligonucleótidos específicos de alelo. El sondaje puede emplear la técnica de transferencia de Southern convencional. Por ejemplo, puede extraerse ADN a partir de células y digerirse con diferentes enzimas de restricción. Después, los fragmentos de restricción pueden separarse por electroforesis en un gel de agarosa, antes de la desnaturalización y transferencia a un filtro de nitrocelulosa. La sonda marcada puede hibridarse con los fragmentos de ADN sobre el filtro y determinarse la unión. El ADN para el sondaje puede prepararse a partir de preparaciones de ARN de células.

[0152] Pueden emplearse estrategias que dependen de la hibridación entre una sonda y un ácido nucleico de ensayo, y la posterior detección de un emparejamiento erróneo. En condiciones apropiadas (temperatura, pH, etc.), una sonda oligonucleotídica hibridará con una secuencia que no sea totalmente complementaria. El grado de emparejamiento de bases entre las dos moléculas será suficiente para que hibriden a pesar de un emparejamiento erróneo. Son bien conocidas en la técnica diversas estrategias para detectar la presencia de un emparejamiento erróneo entre dos moléculas de ácido nucleico que hibridan.

[0153] Por ejemplo, la ARNasa A escinde en el sitio de un emparejamiento erróneo. La escisión puede detectarse sometiendo a electroforesis el ácido nucleico de ensayo con el que ha hibridado la sonda o sonda pertinente y buscando moléculas más pequeñas (es decir, moléculas con mayor movilidad electroforética) que el híbrido de sonda/ensayo de longitud completa.

[0154] El ácido nucleico en una muestra de ensayo, que puede ser una muestra de genómico o una región amplificada del mismo, puede secuenciarse para identificar o determinar la identidad de un alelo polimórfico. El alelo del SNP en el ácido nucleico de ensayo puede compararse por lo tanto con los alelos de susceptibilidad del SNP como se describen en la presente memoria, para determinar si el ácido nucleico de ensayo contiene uno o más

alelos que estén asociados con enfermedad.

5 **[0155]** Típicamente en la secuenciación, los cebadores complementarios a la secuencia diana se diseñan de modo que estén a una distancia adecuada (por ejemplo, 50-400 nucleótidos) del polimorfismo. Después, la secuenciación se lleva a cabo usando técnicas convencionales. Por ejemplo, los cebadores pueden diseñarse usando un *software* que tenga como objetivo seleccionar una secuencia o secuencias dentro de una ventana apropiada, que tengan valores de Tm adecuados y que no posean estructura secundaria o que hibriden con una secuencia no diana.

10 **[0156]** La secuenciación de un producto amplificado puede implicar la precipitación con isopropanol, la resuspensión y la secuenciación usando un kit de secuenciación con terminadores Dye TaqFS+. Los productos de extensión pueden someterse a electroforesis en un secuenciador de ADN ABI 377 y los datos analizarse usando el *software* Sequence Navigator.

15 **[0157]** El análisis de genotipo puede llevarse a cabo mediante análisis de micromatrices. Puede usarse cualquier tecnología de micromatrices adecuada. Puede usarse la metodología descrita en Tejedor et al 2005 (Clinical Chemistry 51: 1137-1144), incluyendo el *software* MG1.0 y en la Solicitud de Patente Internacional N° PCT/IB2006/00796, presentada el 12 de enero de 2006. Esta tecnología usa una matriz de ADN de baja densidad e hibridación con sondas oligonucleotídicas específicas de alelo para explorar para SNP. Por lo tanto, en un aspecto, la micromatriz Artchip y la tecnología de la presente invención pueden usarse para determinar el genotipo de los SNP informativos, como se describe en la presente memoria.

25 **[0158]** Una vez que un sujeto ha recibido un pronóstico de un fenotipo de RA particular (un riesgo significativo de desarrollar ese fenotipo de acuerdo con la invención), puede seleccionarse el tratamiento más apropiado para ese sujeto. De este modo, la invención permite un mejor direccionamiento de terapias a pacientes.

[0159] Por lo tanto, en un aspecto adicional, la invención proporciona un método de selección de un tratamiento adecuado para un sujeto al que se le ha diagnosticado que tiene RA, comprendiendo el método:

- 30 (a) determinar la verosimilitud de que se desarrolle un fenotipo de RA particular en el sujeto mediante un método descrito en la presente memoria; y
(b) seleccionar un tratamiento adecuado.

[0160] El tratamiento seleccionado puede administrarse después al sujeto. Por lo tanto, la invención también se refiere a un método de tratamiento de RA en un sujeto, que comprende:

- 35 (a) determinar la verosimilitud de que se desarrolle un fenotipo de RA particular en el sujeto mediante un método descrito en la presente memoria; y
(b) tratar al sujeto con un tratamiento adecuado.

40 **[0161]** Por ejemplo, si se determina que un sujeto presenta mayor riesgo de desarrollar un fenotipo de RA agresivo (por uno o más de los métodos descritos en la presente memoria), puede seleccionarse un tratamiento apropiado, por ejemplo, un tratamiento más intensivo. Si se evalúa que es probable que un sujeto sea intolerante a metotrexato, puede proporcionarse una terapia alternativa.

45 **[0162]** Pueden proporcionarse medios para llevar a cabo los presentes métodos de pronóstico en forma de kit, por ejemplo, en un recipiente adecuado tal como un vial en el que el contenido está protegido del entorno externo. Por lo tanto, en un aspecto, la invención se refiere además a kits de pronóstico adecuados para su uso en los métodos descritos en la presente memoria. Típicamente, un kit comprende:

- 50 (i) medios para determinar resultados para la variable o variables seleccionadas o variables de SNP; y
(ii) instrucciones para determinar el pronóstico basándose en los resultados de las variables.

55 **[0163]** Los medios (i) pueden comprender una o más sondas oligonucleotídicas adecuadas para la detección de una o más variables de SNP a determinar. Por ejemplo, los medios (i) pueden comprender una o más parejas de sondas o conjuntos de sondas enumeradas en la Tabla 2. En un caso, el kit puede comprender todos los conjuntos de sondas de la Tabla 2.

60 **[0164]** Los medios (i) pueden comprender una micromatriz adecuada, como se describe en la presente memoria. Los medios (i) pueden comprender una o más parejas de cebadores de secuenciación adecuadas para secuenciar una o más de las variables de SNP a determinar.

65 **[0165]** Las instrucciones (ii) comprenden típicamente instrucciones para usar los resultados determinados usando los medios (i) para el pronóstico. Las instrucciones pueden comprender una gráfica que muestre los riesgos de recaída de la RA. El kit puede incluir detalles de funciones de probabilidad que pueden usarse en el pronóstico, tales como las descritas en la presente memoria.

[0166] Un kit puede incluir en algunos casos un programa informático como se describe en la presente memoria.

[0167] Un kit puede incluir otros componentes adecuados para su uso en los presentes métodos. Por ejemplo, un kit puede incluir cebadores adecuados para la amplificación de regiones de ADN diana que contienen los SNP a determinar, tales como los descritos en la presente memoria. Por ejemplo, un kit puede contener una o más parejas de cebadores enumerados en las Tablas 3. Un kit también puede incluir medios de marcaje y detección adecuados, controles y/u otros reactivos tales como tampones, nucleótidos o enzimas, por ejemplo, polimerasa, nucleasa, transferasa.

[0168] Un ácido nucleico de acuerdo con la presente invención, tal como una sonda oligonucleotídica y/o pareja de cebadores de amplificación, puede proporcionarse como parte de un kit. El kit puede incluir instrucciones para el uso del ácido nucleico, por ejemplo, en PCR y/o un procedimiento para determinar la presencia de ácido nucleico de interés en una muestra de ensayo. Un kit en el que el ácido nucleico está destinado a su uso en PCR puede incluir uno o más de otros reactivos necesarios para la reacción, tales como polimerasa, nucleósidos, solución de tampón, etc. El ácido nucleico puede estar marcado.

[0169] Un kit para su uso en la determinación de la presencia o ausencia de ácido nucleico de interés puede incluir uno o más artículos y/o reactivos para la realización del método, tales como medios para proporcionar la propia muestra de ensayo, por ejemplo, un hisopo para retirar células de la cavidad bucal o una jeringa para extraer una muestra de sangre (siendo dichos componentes generalmente estériles).

[0170] En un aspecto adicional, la presente invención también se refiere a microplacas de ADN o micromatrices y a métodos para su uso, que permiten un genotipado fiable de individuos con respecto a múltiples variaciones genéticas asociadas a RA simultáneamente y con fines clínicos.

[0171] Por lo tanto, en un aspecto, la invención proporciona además un método de genotipado de variaciones genéticas asociadas a RA en un individuo, que es suficientemente sensible, específico y reproducible para uso clínico. Los inventores han desarrollado micromatrices de ADN de baja densidad con sondas específicamente diseñadas para su uso en el método, y un método computacional o algoritmo para interpretar y procesar los datos generados por las matrices.

[0172] En un aspecto, la invención se refiere a un método *in vitro* para genotipar variaciones genéticas asociadas a RA en un individuo. El método permite el genotipado simultáneo de múltiples variaciones genéticas humanas presentes en uno o más genes de un sujeto. El método de la invención permite la identificación de cambios de nucleótidos, tales como inserciones, duplicaciones y deleciones, y la determinación del genotipo de un sujeto para una variación genética dada.

[0173] Una variación genética o variante genética se refiere a mutaciones, polimorfismos o variantes alélicas. Una variación o variante genética se encuentra entre individuos dentro de la población y entre poblaciones dentro de la especie.

[0174] Una variación genética asociada a RA puede referirse a una variación genética que esté asociada con RA de una forma estadísticamente significativa, y que pueda usarse como ayuda en el diagnóstico, pronóstico o predicción de la respuesta a la terapia en un individuo.

[0175] Un polimorfismo se refiere a una variación en la secuencia de nucleótidos de ácido nucleico, en la que cada secuencia posible está presente en una proporción igual a o superior al 1% de una población; en un caso particular, cuando dicha variación se produce en sólo un nucleótido (A, C, T o G), se denomina polimorfismo de un solo nucleótido (SNP).

[0176] Una mutación genética se refiere a una variación en la secuencia de nucleótidos en un ácido nucleico en la que cada secuencia posible está presente en menos del 1% de una población.

[0177] Una variante alélica o alelo se refiere a un polimorfismo que aparece en el mismo locus en la misma población.

[0178] Por lo tanto, una variación genética puede comprender una deleción, sustitución o inserción de uno o más nucleótidos. En un aspecto, las variaciones genéticas a genotipar de acuerdo con los presentes métodos comprenden SNP.

[0179] Un gen dado puede comprender una o más variaciones genéticas. Por lo tanto, los presentes métodos pueden usarse para el genotipado de una o más variaciones genéticas en uno o más genes.

[0180] Típicamente el individuo es un ser humano.

[0181] Típicamente, para una variación genética dada, hay tres genotipos posibles:

- AA, el individuo es homocigoto para la variación genética A (por ejemplo, homocigoto para un alelo de tipo silvestre)
- BB, el individuo es homocigoto para la variación genética B (por ejemplo, homocigoto para un alelo mutante)
- AB, el individuo es heterocigoto para las variaciones genéticas A y B (por ejemplo, un alelo de tipo silvestre y uno mutante)

[0182] Las variaciones genéticas, tales como SNP, que se van a analizar de acuerdo con los presentes métodos, están asociadas con RA. Los ejemplos de variaciones genéticas asociadas con RA que pueden evaluarse mediante los presentes métodos incluyen las de la Tabla 1A y Tabla 1B (Figura 1).

[0183] Las secuencias de todos los genes mencionados en la Figura 1 se conocen y están reconocidas en los siguientes sitios web: GeneBank (NCBI), GeneCard (Weizmann Institute of Sciences) y Snpper.chip.org (Innate Immunity PGA). La Tabla 11 proporciona códigos de refSNP (n^o rs) para varios SNP. Estos se toman de la Base de Datos de Polimorfismos de un Solo Nucleótido (dbSNP) corregida por el National Center for Biotechnology Information (NCBI) (<http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?CMD=search&DB=snp>, el 22 de junio de 2007).

[0184] Al permitir el genotipado clínico de uno o más de las variaciones genéticas anteriores, el presente método tiene utilidad, por ejemplo, en el diagnóstico de la susceptibilidad a la presencia de RA en un sujeto. Los presentes métodos de genotipado también son útiles en el pronóstico de fenotipos de RA, como se describe en la presente memoria.

[0185] Al menos una variación genética asociada a RA, por ejemplo, SNP, se analiza en los presentes métodos de genotipado. Los presentes métodos permiten el genotipado simultáneo de múltiples variaciones en un individuo y, típicamente, se analizan múltiples variaciones, en general, al menos 10, 12, 14, 16, 18 ó 20 variaciones genéticas asociadas a RA. Por ejemplo, pueden ensayarse al menos 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105 ó 110 variaciones, o hasta 150, 200, 300, 400, 500 ó 600 variaciones, tal como 250, 350 ó 450 variaciones.

[0186] Por lo tanto, los métodos de genotipado pueden usarse para genotipar a un individuo con respecto a todas o a una selección de las variaciones de la Tabla 1A o 1B, como se describen en la presente memoria. Por ejemplo, puede genotiparse al menos 1, 2, 3, 4, 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 o todas las variaciones de la Tabla 1A o 1B. Las variaciones a detectar pueden incluir además otras variaciones genéticas asociadas a RA.

[0187] La presente invención también incluye métodos en los que se evalúan otras variaciones genéticas además de las variaciones genéticas asociadas a RA.

[0188] De acuerdo con los presentes métodos, se proporciona una muestra que contiene ácido nucleico, que comprende al menos una de las variaciones genéticas a ensayar (el ADN diana). Se describen en la presente memoria muestras adecuadas y métodos para obtener las muestras en relación con los métodos de pronóstico.

[0189] Como se describe, las regiones de ADN que contienen las variaciones genéticas a identificar (regiones de ADN diana) pueden someterse a una reacción de amplificación para obtener productos de amplificación que contienen las variaciones genéticas a identificar. Puede usarse cualquier técnica o método adecuado para la amplificación. En general, la técnica permite la amplificación (simultánea) de todas las secuencias de ADN que contienen las variaciones genéticas a identificar. En otras palabras, cuando se van a analizar múltiples variaciones genéticas, es preferible amplificar simultáneamente todas las regiones de ADN diana correspondientes (que comprenden las variaciones). El realizar la amplificación en una sola etapa (o en tan pocas etapas como sea posible) simplifica el método.

[0190] Por ejemplo, puede llevarse a cabo un PCR múltiple, usando parejas apropiadas de cebadores oligonucleotídicos de PCR que son capaces de amplificar las regiones diana que contienen las variaciones genéticas a identificar. Puede usarse cualquier pareja de cebadores adecuada que permita una amplificación específica de una región de ADN diana. En un aspecto, los cebadores permiten la amplificación en el menor número posible de reacciones de PCR. Por lo tanto, mediante el uso de parejas apropiadas de cebadores oligonucleotídicos y condiciones apropiadas, todas las regiones de ADN diana necesarias para genotipar las variaciones genéticas pueden amplificarse para el análisis de genotipado (por ejemplo, microplaca de ADN) con el mínimo número de reacciones. Los cebadores de PCR adecuados para la amplificación de regiones de ADN diana que comprenden las variaciones genéticas asociadas a RA en la Tabla 1A y 1B se enumeran en las Tablas 3. El presente método puede comprender el uso de uno o más de estos cebadores o una o más de las parejas de cebadores enumeradas. Por ejemplo, los presentes métodos pueden usarse para el genotipado de variaciones de la Tabla 1A seleccionadas como se ha descrito anteriormente. Los cebadores correspondientes de la Tabla 3 pueden seleccionarse para su uso en consecuencia.

[0191] En un caso, los productos de amplificación pueden marcarse durante la reacción de amplificación con un marcador detectable. El objetivo es ser capaz de detectar posteriormente la hibridación entre los fragmentos de ADN diana que contienen las variaciones genéticas que se están analizando y sondas fijadas a un soporte sólido. Cuanto mayor sea el grado de hibridación de un ADN diana marcado con una sonda, mayor será la intensidad de marcador detectable en esa posición de sonda.

[0192] Los productos de amplificación pueden marcarse por métodos convencionales. Por ejemplo, puede incorporarse un nucleótido marcado durante la reacción de amplificación o pueden usarse cebadores marcados para la amplificación.

[0193] El marcaje puede ser directo usando, por ejemplo, marcadores fluorescentes o radiactivos, o cualquier otro marcador conocido por los expertos en la materia. Los ejemplos de fluoróforos que pueden usarse incluyen, por ejemplo, Cy3 o Cy5. Como alternativa, pueden usarse enzimas para el marcaje de muestras, por ejemplo, fosfatasa alcalina o peroxidasa. Los ejemplos de isótopos radiactivos que pueden usarse incluyen, por ejemplo, ³³P, ¹²⁵I, o cualquier otro marcador conocido por los expertos en la materia. En un caso, el marcaje de productos de amplificación se lleva a cabo usando un nucleótido que se ha marcado directa o indirectamente con uno o más fluoróforos. En otro ejemplo, el marcaje de productos de amplificación se lleva a cabo usando cebadores marcados directa o indirectamente con uno o más fluoróforos.

[0194] El marcaje también puede ser indirecto, usando, por ejemplo, métodos químicos o enzimáticos. Por ejemplo, un producto de amplificación puede incorporar un miembro de una pareja de unión específica, por ejemplo, avidina o estreptavidina, conjugado con un marcador fluorescente, y la sonda con la que hibridará puede unirse al otro miembro de la pareja de unión específica, por ejemplo, biotina (indicador), permitiendo que la señal de unión de sonda/diana se mida por fluorimetría. En otro ejemplo, un producto de amplificación puede incorporar un miembro de una pareja de unión específica, por ejemplo, un anticuerpo anti-dioxigenina combinado con una enzima (marcador), y la sonda con la que hibridará puede unirse al otro miembro de la pareja de unión específica, por ejemplo, dioxigenina (indicador). En la hibridación del producto de amplificación a la sonda el sustrato enzimático se convierte en un producto luminoso o fluorescente y la señal puede leerse, por ejemplo, por quimioluminiscencia o fluorimetría.

[0195] El ácido nucleico que comprende la variación o variaciones genéticas a ensayar, por ejemplo, los productos de amplificación (opcionalmente marcados), puede someterse además a una reacción de fragmentación, obteniendo de este modo algunos productos de fragmentación que comprenden o contienen las variaciones genéticas a identificar o analizar. Típicamente, la fragmentación aumenta la eficacia de la reacción de hibridación. La fragmentación puede llevarse a cabo por cualquier método adecuado conocido en la técnica, por ejemplo, poniendo en contacto el ácido nucleico, por ejemplo, los productos de amplificación, con una enzima adecuada tal como ADNasa.

[0196] Si el ácido nucleico no se ha marcado previamente, por ejemplo, durante la reacción de amplificación (y, típicamente, cuando no se lleva a cabo una amplificación o ligación post-hibridación en el soporte sólido), entonces el marcaje con un marcador detectable puede llevarse a cabo pre-hibridación por marcaje de los productos de fragmentación. Se conocen en la materia técnicas de marcaje adecuadas, y pueden ser directas o indirectas, como se describe en la presente memoria. El marcaje directo puede comprender el uso de, por ejemplo, fluoróforos, enzimas o isótopos radiactivos. El marcaje indirecto puede comprender el uso de, por ejemplo, parejas de unión específicas que incorporan, por ejemplo, fluoróforos, enzimas, etc. Por ejemplo, si los productos de amplificación no se han marcado durante la reacción de amplificación, los productos de fragmentación pueden someterse a un marcaje directo o indirecto con uno o diversos marcadores, por ejemplo, uno o diversos fluoróforos, aunque pueden usarse otros marcadores conocidos por los expertos en la materia.

[0197] De acuerdo con los presentes métodos, el ácido nucleico, por ejemplo, los productos de amplificación o fragmentación, que comprende la variación o variaciones genéticas a detectar (ADN diana), se pone en contacto con sondas oligonucleotídicas que son capaces de detectar las variaciones genéticas correspondientes por hibridación en condiciones adecuadas.

[0198] Típicamente, las condiciones de hibridación permiten la hibridación específica entre sondas y los ácidos nucleicos diana correspondientes para formar complejos de hibridación de sonda/diana específicos, minimizando al mismo tiempo la hibridación entre sondas que llevan uno o más emparejamientos erróneos con el ADN. Dichas condiciones pueden determinarse empíricamente, por ejemplo, variando el tiempo y/o la temperatura de hibridación y/o el número y la rigurosidad de las etapas de lavado de la matriz que se realizan después de la hibridación, y que están diseñadas para eliminar todas las interacciones de sonda-ADN que sean inespecíficas.

[0199] En el método, las sondas se proporcionan depositadas en un soporte sólido o superficie. Las sondas se depositan en posiciones en el soporte sólido de acuerdo con un patrón predeterminado, formando una "microplaca de ADN". Se ha descubierto que las microplacas deberían cumplir varios requisitos para usarse en los presentes métodos, por ejemplo, en términos del diseño de las sondas, el número de sondas proporcionadas para cada variación genética a detectar y la distribución de las sondas en el soporte. Estos se describen en detalle en la

presente memoria. Los inventores han desarrollado microplacas de genotipado adecuadas para su uso en los presentes métodos y, por consiguiente, en un aspecto, la invención proporciona una microplaca de ADN o (micro)matriz que comprende una pluralidad de sondas depositadas o inmovilizadas en un soporte sólido como se describe en la presente memoria.

5 **[0200]** En general, el soporte o fase sólida comprende sondas oligonucleotídicas adecuadas para la detección de cada variación genética a ensayar en el presente método. El número y el tipo de variaciones genéticas a ensayar usando una microplaca puede seleccionarse como se describe en la presente memoria.

10 **[0201]** Típicamente, habrá al menos una sonda que sea capaz de hibridar específicamente con la variación genética A (por ejemplo, un alelo de tipo silvestre o normal) (sonda 1) y una sonda que sea capaz de hibridar específicamente con la variación genética B (por ejemplo, un alelo mutante) (sonda 2) en las condiciones de hibridación seleccionadas. Estas sondas forman una pareja de sondas. La sonda 1 es para la detección de la variación genética A y la sonda 2 es para la detección de la variación genética B. Típicamente, las sondas pueden usarse para discriminar entre A y B (por ejemplo, los alelos de tipo silvestre y mutante). Las sondas pueden examinar la cadena con sentido o la antisentido. Típicamente, las sondas 1 y 2 examinan la misma cadena de ácido nucleico (por ejemplo, la cadena con sentido o la cadena antisentido), aunque en algunos casos las sondas pueden examinar cadenas diferentes. En un aspecto, las sondas 1 y 2 tienen la misma secuencia excepto por el sitio de la variación genética.

20 **[0202]** En un caso, las sondas en una pareja de sondas tienen la misma longitud. En algunos aspectos, cuando se proporcionan dos o más parejas de sondas para el análisis de una variación genética, las sondas pueden tener todas la misma longitud.

25 **[0203]** Preferentemente, se proporciona más de una pareja de sondas para la detección de cada variación genética. Por lo tanto, pueden proporcionarse al menos 2, 3, 4, 5, 6, 7, 8, 9, 10 o más parejas de sondas por variación genética. En un aspecto, se proporcionan (al menos) 2 parejas de sondas. El objetivo es reducir el índice de falsos positivos y negativos en los presentes métodos.

30 **[0204]** Por ejemplo, para una variación genética dada, puede haber:

- Una sonda 1 que es capaz de hibridar con la variación genética A (por ejemplo, un alelo normal).
- Una sonda 2 que es capaz de hibridar con la variación genética B (por ejemplo, un alelo mutante).
- Una sonda 3 que es capaz de hibridar con la variación genética A (por ejemplo, un alelo normal).
- Una sonda 4 que es capaz de hibridar con la variación genética B (por ejemplo, un alelo mutante).

35 **[0205]** Las sondas pueden examinar la misma cadena o cadenas diferentes. Por lo tanto, en una realización, las sondas 3 y 4 son las sondas complementarias de las sondas 1 y 2, respectivamente, y están diseñadas para examinar la cadena complementaria. En un aspecto, se prefiere que las sondas proporcionadas para la detección de cada variación genética examinen ambas cadenas.

40 **[0206]** Pueden proporcionarse más de 2 parejas de sondas para el análisis de una variación genética como anteriormente. Por ejemplo, cuando existe una variación genética como cualquiera de 4 bases en la misma cadena (por ejemplo, hay tres posibilidades mutantes), puede proporcionarse al menos una pareja de sondas para detectar cada posibilidad. Preferentemente, se proporcionan al menos 2 parejas de sondas para cada posibilidad.

45 **[0207]** Por lo tanto, por ejemplo, para un SNP G2677T/A/C, puede proporcionarse al menos una pareja de sondas para la detección de G2677T, una pareja para la detección de G2677/A, y una pareja para la detección de G2677C. Preferentemente, se proporcionan al menos dos parejas de sondas para cada una de estas sustituciones.

50 **[0208]** Se conocen en la técnica varios métodos para diseñar sondas oligonucleotídicas adecuadas para su uso en microplacas de ADN.

55 **[0209]** Puede usarse un método de "embaldosado convencional". En este método, se diseñan 4 oligonucleótidos que son totalmente complementarios a la secuencia de referencia excepto en la posición central en la que, típicamente, se examina los 4 nucleótidos posibles A, C, G y T. Un ejemplo ilustrativo de esta estrategia es la microplaca de ADN para el genotipado de VIH-1 (Affymetrix).

60 **[0210]** En el "embaldosado alternativo" se diseñan 5 oligonucleótidos, de modo que el quinto examina una posible delección en la secuencia. Un ejemplo de esta estrategia es la microplaca de ADN para detectar mutaciones en p53 (Affymetrix).

65 **[0211]** En el "embaldosado en bloque" se diseñan 4 oligonucleótidos que son totalmente complementarios a la secuencia normal y otros 4 totalmente complementarios a la secuencia mutante. El nucleótido que cambia se pone en la posición central, pero se pone un emparejamiento erróneo de una de las 4 bases (A, C, T o G) 2 nucleótidos

antes o después de la posición de nucleótido que se desea cuantificar. Un ejemplo de esta estrategia es la microplaca de ADN para la detección de mutaciones en el citocromo p450 (Roche y Affymetrix).

[0212] Un ejemplo adicional es el “embaldosado en bloque alternativo”, en el que el “emparejamiento erróneo” se usa para aumentar la especificidad del híbrido no sólo en una posición, sino también en las posiciones -4, -1, 0, +1 y +4 para identificar el cambio producido en la posición central o 0. Un ejemplo es la microplaca de ADN para detectar 1.500 SNP (Affymetrix).

[0213] Pueden usarse una o más de cualquiera de estas estrategias para diseñar sondas para la presente invención. Preferentemente se usa el embaldosado convencional, en particular con 2 parejas de sondas, por ejemplo, 2 parejas de sondas complementarias como anteriormente. Por lo tanto, es preferible que la secuencia oligonucleotídica sea complementaria al ADN o secuencia diana en las regiones que flanquean el nucleótido o nucleótidos variables. Sin embargo, en algunos casos pueden introducirse uno o más emparejamientos erróneos, como se ha descrito anteriormente.

[0214] Las sondas oligonucleotídicas para su uso en la presente invención presentan típicamente la base que se va a examinar (el sitio de la variación genética) en el centro del oligonucleótido. Éste es particularmente el caso cuando se usan métodos de hibridación diferencial, ya que en general esto permite la mejor discriminación entre sondas emparejadas y emparejadas erróneamente. En estos métodos, típicamente hay formación de complejos de hibridación detectables específicos sin amplificación en la microplaca post-hibridación. Por ejemplo, para mutaciones precisas (una sola base), la base que difiere entre el alelo normal y el mutante se pone típicamente en la posición central de la sonda. En el caso de inserciones, deleciones y duplicaciones, el primer nucleótido que difiere entre la secuencia normal y la mutante se pone en la posición central. Se cree que la colocación de la mutación en el centro de la sonda maximiza la especificidad.

[0215] Cuando se emplea amplificación en la microplaca post-hibridación (por ejemplo, métodos de ligación o extensión de cebadores), las sondas oligonucleotídicas presentan típicamente la base o bases variables en el extremo 3' de la sonda. Cuando se usa metodología de OLA, también se diseñan oligonucleótidos (marcados directa o indirectamente) que hibridan con complejos de sonda-diana para permitir la ligación.

[0216] En general, las sondas para su uso en la presente invención comprenden, o en algunas realizaciones consisten (esencialmente) en 17 a 27 nucleótidos, por ejemplo, 19, 21, 23 ó 25 nucleótidos, o 18, 20, 22, 24 ó 26 nucleótidos.

[0217] Preferentemente, las sondas individuales proporcionadas para la detección de una variación genética son capaces de hibridar específicamente con los alelos normales y mutantes, respectivamente, en las condiciones de hibridación seleccionadas. Por ejemplo, la temperatura de fusión de los complejos de sonda/diana puede darse a 75-85°C, y la hibridación puede ser durante una hora, aunque también pueden ser suficientes temperatura superiores e inferiores e hibridaciones más largas y más cortas.

[0218] Las sondas proporcionadas para (adecuadas para) la detección de cada variación genética (como se ha descrito anteriormente) son típicamente capaces de discriminar entre la variación genética A y B (por ejemplo, los alelos normales y mutantes) en las condiciones de hibridación dadas como anteriormente. Preferentemente, la capacidad de discriminación de las sondas es sustancialmente del 100%. Si la capacidad de discriminación no es del 100%, preferentemente las sondas se rediseñan. Preferentemente, la temperatura de fusión de los complejos de sonda/diana se da a 75-85 grados C. Se describen en la presente memoria métodos para ensayar la capacidad de discriminación.

[0219] En un ejemplo, las sondas proporcionadas para la detección de una variación genética examinan ambas cadenas y tienen longitudes que varían de 19-27 nucleótidos. Preferentemente, las sondas tienen una capacidad de discriminación del 100%, y la temperatura de fusión de los complejos de sonda/diana es de 75-85 grados.

[0220] Típicamente, para obtener sondas para su uso en los presentes métodos, se diseñan y ensayan experimentalmente varias sondas, por ejemplo, para determinar su especificidad de hibridación y su capacidad para discriminar entre variantes genéticas (por ejemplo, un alelo normal y uno mutante). Pueden diseñarse secuencias de sondas oligonucleotídicas candidatas como se ha descrito anteriormente. Éstas pueden variar, por ejemplo, en longitud, especificidad de cadena, posición de la variación genética y grado de complementariedad con la secuencia que flanquea la variación genética en el ADN diana. Una vez que se han diseñado parejas de sondas, éstas pueden ensayarse para determinar su especificidad de hibridación y su capacidad de discriminación. La capacidad de sondas específicas para discriminar entre las variaciones genéticas A y B (por ejemplo, alelos normales y mutantes) depende de las condiciones de hibridación, de la secuencia que flanquea la mutación y de la estructura secundaria de la secuencia en la región de la mutación. Usando condiciones de hibridación estables pueden establecerse los parámetros apropiados, tales como las especificidades y longitudes de cadena, para maximizar la discriminación. Preferentemente, la variación genética se mantiene en la posición central en las sondas ensayadas.

- 5 **[0221]** En la presente memoria se describen métodos para ensayar la capacidad de discriminación de las sondas. Típicamente, se proporcionan varias parejas de sondas candidatas y se usan en un método de entrenamiento como se describe a continuación. En general, se ensayan dos parejas de sondas (sondas 1 y 2 y sondas 3 y 4) en el método. Por ejemplo, pueden ensayarse dos parejas de sondas que examinan ambas cadenas (complementarias entre sí). Si no es posible obtener una discriminación del 100% entre los tres grupos de genotipado usando las sondas, típicamente las sondas se rediseñan. Las condiciones de hibridación en el método de entrenamiento se mantienen generalmente estables. Típicamente, la temperatura de fusión de los complejos de sonda/diana es de 75-85°C.
- 10 **[0222]** Por ejemplo, partiendo de sondas de 25 nucleótidos que detectan una variación genética (por ejemplo, el alelo normal) y otra variación genética (por ejemplo, un alelo mutante) en ambas cadenas (con sentido y antisentido), en general puede ensayarse experimentalmente un promedio de 8 sondas para identificar dos parejas definitivas.
- 15 **[0223]** Las sondas se seleccionan para tener una especificidad de hibridación y una capacidad de discriminación máximas entre variantes genéticas (por ejemplo, un alelo normal y uno mutante) en condiciones de hibridación adecuadas. Por ejemplo, las sondas para la detección de una variación genética dada, por ejemplo, dos parejas de sondas, tienen típicamente una capacidad de discriminación de sustancialmente el 100%. Típicamente, la temperatura de fusión de complejos de sonda/diana es a 75-85°C.
- 20 **[0224]** Usando los métodos de la presente memoria los inventores han desarrollado sondas oligonucleotídicas adecuadas para la detección de las variaciones genéticas asociadas a RA en la Tabla 1A y 1B. Estas sondas se presentan como SEC ID N°: 1-360 (Tabla 2). Las sondas se enumeran en conjuntos de sondas (90 conjuntos en total), de acuerdo con la variación genética a detectar. Se enumeran al menos dos parejas de sondas en cada conjunto.
- 25 **[0225]** En un aspecto, la invención se refiere a una o más de cualquiera de las sondas oligonucleotídicas, parejas de sondas o conjuntos de sondas expuestos en las SEC ID N°: 1-360 (Tabla 2) y a su uso en el genotipado, diagnóstico o pronóstico o métodos terapéuticos de la invención. La invención se refiere además a una o más de cualquiera de las sondas oligonucleotídicas, parejas de sondas o conjuntos de sondas expuestos en las SEC ID N°: 30 1-360 para su uso en medicina, por ejemplo, en un método de diagnóstico o pronóstico o terapéutico descrito en la presente memoria. Una microplaca de la invención puede comprender una o más de las parejas o conjuntos de sondas enumerados como se describe en la presente memoria.
- 35 **[0226]** En general, las sondas se proporcionan en el soporte repetidas. Típicamente, se proporcionan al menos 4, 6, 8, 10, 12, 14, 16, 18 ó 20 repeticiones de cada sonda, en particular, 6, 8 ó 10 repeticiones. Por lo tanto, por ejemplo, el soporte (o microplaca de ADN) puede comprender o incluir 10 repeticiones para cada una de las (al menos) 4 sondas usadas para detectar cada variación genética (es decir, 40 sondas). Como alternativa, el soporte (o microplaca de ADN) puede comprender o incluir 8 repeticiones para cada una de las (al menos) 4 sondas usadas para detectar cada variación genética (es decir, 32 sondas). Todavía además el soporte (o microplaca de ADN) puede comprender o incluir 6 repeticiones para cada una de las (al menos) 4 sondas usadas para detectar cada variación genética (es decir, 24 sondas). El uso de repeticiones de sondas ayuda a minimizar las distorsiones en la interpretación de los datos de la microplaca y mejora la fiabilidad de los métodos.
- 40 **[0227]** En general, el soporte también comprende una o más sondas oligonucleotídicas de control. También se proporcionan repetidas como anteriormente. Por lo tanto, el soporte (o microplaca de ADN) puede comprender además uno o más oligonucleótidos depositados en el soporte que sean útiles como controles positivos y/o negativos de las reacciones de hibridación. Si las reacciones de amplificación o ligación post-hibridación se llevan a cabo en la microplaca, también puede haber uno o más controles positivos o negativos de estas reacciones.
- 45 **[0228]** Típicamente, la microplaca o matriz incluirá sondas de control positivo, por ejemplo, sondas que se sabe que son complementarias y que pueden hibridar con secuencias en las moléculas polinucleotídicas diana, sondas que se sabe que hibridan con un ADN de control externo, y sondas de control negativas, por ejemplo, sondas que se sabe que no son complementarias y que no pueden hibridar con secuencias en las moléculas polinucleotídicas diana. La microplaca puede tener uno o más controles específicos para cada diana, por ejemplo, 2, 3 o más controles. También puede haber al menos un control para la matriz.
- 50 **[0229]** Los controles positivos pueden sintetizarse, por ejemplo, a lo largo del perímetro de la matriz o en franjas en diagonal a través de la matriz. La complementaria inversa para cada sonda puede sintetizarse al lado de la posición de la sonda para servir como control negativo. En otro ejemplo más, pueden usarse secuencias de otras especies de organismos como controles negativos para contribuir a determinar la hibridación de fondo (inespecífica).
- 55 **[0230]** Como anteriormente, el soporte (o microplaca de ADN) puede incluir algunos (uno o más) oligonucleótidos depositados en el soporte que sean útiles como controles positivos y negativos de las reacciones de hibridación. En general, cada una de las submatrices, por ejemplo, 16, que típicamente constituyen una microplaca de ADN, está flanqueada por algunos controles de hibridación externos que sirven como puntos de referencia que permiten que
- 60
- 65

los puntos dentro de la rejilla se localicen más fácilmente.

[0231] En un caso, la secuencia de nucleótidos de un ADN de control externo es la siguiente (5'→3'):

5 CEH: GTCGTCAAGATGCTACCGTTCAGGAGTCGTCAAGATGCTACCGTTCAGGA (SEC ID N°: 541)

y las secuencias de los oligonucleótidos para su detección son las siguientes:

10 ON1: CTTGACGACTCCTGAACGG (SEC ID N°: 542)

ON2: CTTGACGACACCTGAACGG (SEC ID N°: 543)

[0232] Las sondas de control positivo están diseñadas generalmente para hibridar por igual con todas las muestras de ADN diana y proporcionar una intensidad de señal de referencia frente a la que puede compararse la hibridación del ADN diana (muestra) con las sondas de ensayo. Los controles negativos comprenden “blancos” en los que sólo se ha aplicado disolvente (DMSO) al soporte u oligonucleótidos de control que se han seleccionado para que no muestren hibridación o muestren una hibridación únicamente mínima con la diana, por ejemplo, ADN humano (el ADN de ensayo). La intensidad de cualquier señal detectada en características de oligonucleótidos de control negativo o blancos es un indicio de interacciones inespecíficas entre el ADN de muestra y la matriz y, por lo tanto, es una medida de la señal de fondo frente a la que debe discriminarse la señal de las interacciones de sonda-muestra reales.

[0233] De forma deseable, el número de secuencias en la matriz será tal que cuando el número de ácidos nucleicos adecuados para la detección de variaciones genéticas es n , el número de ácidos nucleicos de control positivos y negativos es n' , donde n' es típicamente 0,01 a 0,4 n .

[0234] En general, el soporte o microplaca es adecuado para genotipar variaciones genéticas asociadas a RA, en particular, el genotipado de acuerdo con los presentes métodos. La microplaca comprende típicamente sondas adecuadas para la detección de al menos una pero, preferentemente, múltiples variaciones genéticas asociadas a RA, típicamente al menos 10, 12, 14, 16, 18 ó 20 variaciones. Por ejemplo, pueden ensayarse al menos 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 105 ó 110 variaciones, o hasta 150, 200, 300, 400, 500 ó 600 variaciones, tal como 250, 350 ó 450 variaciones.

[0235] Las variaciones genéticas asociadas a RA pueden incluir cualquiera o todas las de la Tabla 1A o 1B. Por lo tanto, una matriz o microplaca puede comprender sondas adecuadas para genotipar a un individuo con respecto a todas las variaciones de la Tabla 1A o 1B, o una selección de las variaciones de la Tabla, como se describe en la presente memoria.

[0236] Una microplaca de ADN de acuerdo con la invención ('Artchip') permite el genotipado simultáneo, sensible, específico y reproducible de variaciones genéticas asociadas con RA. Se proporcionan ejemplos no limitantes de dichas variaciones en la Tabla 1A y 1B. No obstante, el número de variaciones genéticas contenidas en la Tabla puede aumentarse a medida que se identifiquen posteriormente otras variaciones genéticas y se asocien con RA. Por lo tanto, las variaciones genéticas detectables por la microplaca pueden comprender, o consistir (esencialmente) en las enumeradas en la Tabla 1A o Tabla 1B o en la Figura 11, o una selección de éstas, como se describen en relación con los presentes métodos. La microplaca comprenderá sondas adecuadas para la detección de estas variaciones genéticas como se describe en la presente memoria. Preferentemente, cuando una microplaca comprende sondas para la detección de una variación genética de la Tabla 1A, la microplaca comprende una o más de las sondas enumeradas en las SEC ID N°: 1-360 (Tabla 2) como adecuadas para la detección de esa variación genética, por ejemplo, los conjuntos de sondas enumerados en las SEC ID N°: 1-360 para la detección de esa variación. En un aspecto, la presente microplaca comprende una o más sondas seleccionadas de las SEC ID N°: 1-360. Las sondas se enumeran en conjuntos de sondas, de acuerdo con la variación genética a detectar. Se proporciona en cada conjunto al menos dos parejas de sondas. Una microplaca puede comprender al menos una pareja de sondas o al menos un conjunto de sondas, o una selección de los conjuntos de sondas, por ejemplo, una pareja de sondas o un conjunto de sondas de al menos 5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85 o los 90 conjuntos, de acuerdo con las variaciones genéticas que se estén ensayando. Una microplaca puede comprender otras sondas para la detección de variaciones de la Tabla 1A u otras variaciones asociadas con RA en lugar de o además de las específicamente enumeradas.

[0237] La Artchip puede comprender además sondas oligonucleotídicas para la detección de variaciones genéticas no asociadas con RA. Por ejemplo, las microplacas pueden comprender sondas para la detección de variaciones genéticas tales como SNP asociados con otra afección (relacionada) tal como cáncer de colon, rectal o de vejiga. Típicamente, en la Artchip, el número de ácidos nucleicos adecuados para la detección de variaciones genéticas asociadas con RA (por ejemplo, las de la Tabla 1A o Tabla 1B o Figura 11) representan al menos el 50%, 60%, 70%, 75%, 80%, 85%, 90%, 95%, 97%, 98%, 99% o más de los ácidos nucleicos en la matriz.

[0238] En general, el soporte o microplaca tiene de 300 a 40000 ácidos nucleicos (sondas), por ejemplo, de 400 a

30000 o de 400 a 20000. La microplaca puede tener de 1000 a 20000 sondas, tal como de 1000 a 15000 o de 1000 a 10000, o de 1000 de 5000. Una microplaca adecuada puede tener de 2000 a 20000, de 2000 a 10000 o de 2000 a 5000 sondas. Por ejemplo, una microplaca puede tener 1000, 2000, 3000, 4000, 5000, 6000, 7000, 8000, 9000, 10000, 12000, 14000, 16000, 18000 ó 20000 sondas. También se prevén microplacas más pequeñas de 400 a 1000 sondas, tales como 400, 500, 600, 700, 800, 900 ó 950 sondas.

[0239] En general, la matriz o microplaca de la invención comprende un soporte o superficie con una matriz ordenada de sitios de unión (por ejemplo, hibridación) o sondas. Por lo tanto, la disposición de las sondas en el soporte está predeterminada. Cada sonda (es decir, cada repetición de sonda) se localiza en una posición predeterminada conocida en el soporte sólido de modo que la identidad (es decir, la secuencia) de cada sonda puede determinarse a partir de su posición en la matriz. Típicamente, las sondas están uniformemente distribuidas en un patrón predeterminado.

[0240] Preferentemente, las sondas depositadas en el soporte, aunque mantienen una disposición predeterminada, no están agrupadas por variación genética sino que tienen una distribución aleatoria. Típicamente, tampoco están agrupadas dentro de la misma variación genética. Si se desea, esta distribución aleatoria puede ser siempre la misma. Por lo tanto, típicamente, las sondas se depositan en el soporte sólido (en una matriz) siguiendo un patrón predeterminado de modo que estén distribuidas uniformemente, por ejemplo, entre las dos áreas que pueden constituir una microplaca de ADN, pero no agrupadas de acuerdo con la variación genética a caracterizar. La distribución de repeticiones de sondas a través de la matriz de esta forma contribuye a reducir o eliminar cualquier distorsión de la señal y de la interpretación de datos, por ejemplo, que surja de una distribución no uniforme de interferencias de fondo a través de la matriz.

[0241] Como se ha explicado anteriormente, las sondas pueden disponerse en un soporte en submatrices.

[0242] El soporte, sobre el que se deposita la pluralidad de sondas, puede ser cualquier soporte sólido al que puedan unirse oligonucleótidos. Puede usarse en la invención prácticamente cualquier soporte al que pueda unirse o en el que pueda inmovilizarse un oligonucleótido, y que pueda usarse en la producción de microplacas de ADN. Por ejemplo, dicho soporte puede ser de un material no poroso, por ejemplo, vidrio, sílica, plástico o un material poroso tal como una membrana o filtro (por ejemplo, nylon, nitrocelulosa) o un gel. En una realización, dicho soporte es un soporte de vidrio, tal como un portaobjetos de vidrio.

[0243] Las micromatrices se preparan en general por selección de sondas que comprenden una secuencia polinucleotídica dada, y después inmovilización de dichas sondas en un soporte sólido o superficie. Las sondas pueden diseñarse, ensayarse y seleccionarse como se describe en la presente memoria. En general, las sondas pueden comprender secuencias de ADN. En algunas realizaciones, las sondas pueden comprender secuencias de ARN o secuencias de copolímeros de ADN y ARN. Las secuencias polinucleotídicas de las sondas también pueden comprender análogos de ADN y/o ARN, o combinaciones de los mismos. Por ejemplo, las secuencias polinucleotídicas de las sondas pueden ser fragmentos completos o parciales de ADN genómico. Las secuencias polinucleotídicas de las sondas también pueden ser secuencias de nucleótidos sintetizadas, tales como secuencias oligonucleotídicas sintéticas. Las secuencias de sondas pueden sintetizarse enzimáticamente *in vivo*, enzimáticamente *in vitro* (por ejemplo, mediante PCR) o no enzimáticamente *in vitro*.

[0244] Las micromatrices o microplacas pueden fabricarse de varias formas. Independientemente de cómo se produzcan, las micromatrices comparten típicamente ciertas características. Las matrices son reproducibles, permitiendo que se produzcan y comparen fácilmente entre sí múltiples copias de una matriz dada. Preferentemente, las micromatrices se fabrican a partir de materiales que son estables en condiciones de unión (por ejemplo, hibridación de ácido nucleico). Las micromatrices son preferentemente pequeñas, por ejemplo, de entre 0,25 y 25 o 0,5 y 20 cm², tal como de 0,5 a 20 cm² o de 0,5 a 15 cm², por ejemplo, de 1 a 15 cm² o de 1 a 10 cm², tal como de 2, 4, 6 u 8 cm².

[0245] Las sondas pueden unirse al presente soporte usando técnicas convencionales para la inmovilización de oligonucleótidos en la superficie de los soportes. Las técnicas usadas dependen, entre otros factores, de la naturaleza del soporte usado [porosa (membranas, micropartículas, etc.) o no porosa (vidrio, plástico, sílica, etc.)]. En general, las sondas pueden inmovilizarse en el soporte usando técnicas de inmovilización no covalentes o usando técnicas de inmovilización basadas en la unión covalente de las sondas al soporte por procedimientos químicos.

[0246] La preparación de soportes no porosos (por ejemplo, vidrio, sílica, plástico) requiere, en general, el pretratamiento con grupos reactivos (por ejemplo, amino, aldehído) o el recubrimiento de la superficie del soporte con un miembro de una pareja de unión específica (por ejemplo, avidina, estreptavidina). Asimismo, en general, es aconsejable preactivar las sondas a inmovilizar por medio de grupos correspondientes tales como tiol, amino o biotina para conseguir una inmovilización específica de las sondas en el soporte.

[0247] La inmovilización de las sondas en el soporte puede llevarse a cabo por métodos convencionales, por

ejemplo, por medio de técnicas basadas en la síntesis *in situ* de sondas en el soporte (por ejemplo, fotolitografía, síntesis química directa, etc.) o mediante técnicas basadas en, por ejemplo, brazos robóticos que depositan la sonda presintetizada correspondiente (por ejemplo, impresión sin contacto, impresión por contacto).

5 **[0248]** En una realización, el soporte es un portaobjetos de vidrio y, en este caso, las sondas, en el número de repeticiones establecido (por ejemplo, 6, 8 ó 10) se imprimen sobre portaobjetos de vidrio pretratados, por ejemplo, recubiertos con amino silanos, usando un equipo para la producción automática de microplacas de ADN por deposición de los oligonucleótidos en los portaobjetos de vidrio ("dispositivo de generación de micromatrices"). La deposición se lleva a cabo en condiciones apropiadas, por ejemplo, por medio de entrecruzamiento con radiación ultravioleta y calentamiento (80°C), manteniendo la humedad y controlando la temperatura durante el procedimiento de deposición, típicamente una humedad relativa de entre el 40-50%, y típicamente a una temperatura de 20°C.

10 **[0249]** Las sondas repetidas se distribuyen uniformemente entre las áreas o sectores (submatrices) que constituyen típicamente una microplaca de ADN. El número de repeticiones y su distribución uniforme a través de la microplaca de ADN minimiza la variabilidad que surge del procedimiento de impresión que puede afectar a los resultados experimentales. Asimismo, pueden imprimirse controles de hibridación positivos y negativos (como se describe en la presente memoria).

15 **[0250]** Para controlar la calidad del procedimiento de fabricación de la microplaca de ADN, en términos de señal de la hibridación, interferencia de fondo, especificidad, sensibilidad y reproducibilidad de cada repetición, así como las diferencias causadas por variaciones en la morfología de las características de sondas aplicadas puntualmente después de la impresión, puede usarse un ADN comercial. Por ejemplo, como control de calidad de la impresión de las microplacas de ADN puede llevarse a cabo la hibridación con un ADN comercial (por ejemplo, k562 DNA High Molecular Weight, Promega).

20 **[0251]** En primer lugar, se analiza la morfología y el tamaño de las manchas impresas. En la hibridación con ADN de control, se cumplen los parámetros descritos a continuación para determinar la fiabilidad de la determinación de genotipo; en concreto, el cociente entre la intensidad de señal y la interferencia de fondo, el promedio de especificidad y sensibilidad y la reproducibilidad entre copias repetidas de la misma sonda. Este método permite que se determine el genotipo correcto del ADN de control.

25 **[0252]** Como anteriormente, de acuerdo con el presente método, una muestra de ácido nucleico, por ejemplo, productos de amplificación o fragmentación, que comprende la variación o variaciones genéticas a detectar (ADN diana), se pone en contacto con una matriz de sondas como se describe en la presente memoria, en condiciones que permitan que se produzca la hibridación entre ADN diana y las sondas correspondientes. De este modo se forman complejos de hibridación específicos entre el ácido nucleico diana y las sondas correspondientes.

30 **[0253]** La hibridación de, por ejemplo, productos de fragmentación, con sondas capaces de detectar las variaciones genéticas correspondientes depositadas en un soporte puede llevarse a cabo usando métodos y dispositivos convencionales. En un caso, la hibridación se lleva a cabo usando una estación de hibridación automática. Para que se produzca la hibridación, los, por ejemplo, productos de fragmentación, se ponen en contacto con las sondas en condiciones que permitan que tenga lugar la hibridación. El uso de condiciones de hibridación estables permite que se optimice la longitud y la secuencia de las sondas para maximizar la discriminación entre las variaciones genéticas A y B, por ejemplo, entre secuencias de tipo silvestre y mutantes, como se describe en la presente memoria.

35 **[0254]** En un caso, el método depende de la hibridación diferencial, en particular, un aumento en la señal de hibridación. El método implica la formación de complejos de hibridación específicos entre ADN diana y las sondas correspondientes. Por lo tanto, el ADN diana que lleve la secuencia de tipo silvestre hibridará con las sondas diseñadas para detectar la secuencia de tipo silvestre, mientras que el ADN diana que lleve una secuencia mutante hibridará con las sondas diseñadas para detectar esa secuencia mutante. Los complejos de hibridación se marcan de forma detectable por medios descritos en la presente memoria (por ejemplo, el ADN diana se marca directamente, o tanto la diana como la sonda se marcan de tal modo que el marcador sólo sea detectable en la hibridación). Por detección de la intensidad de marcador detectable (si existe) en las posiciones de sondas predeterminadas es posible determinar la naturaleza del ADN diana en la muestra. En este caso, las sondas (también denominadas oligonucleótidos específicos de alelo, ASO) tienen preferentemente el nucleótido o nucleótidos variables en la posición central, como se describe en la presente memoria.

40 **[0255]** En otro caso, la hibridación de ADN diana con sondas en el soporte sólido (microplaca) puede estar seguida de amplificación en la microplaca, por ejemplo, usando extensión de cebadores o ligación, por ejemplo, tecnologías de ensayo de ligación de oligonucleótidos (OLA) (Eggerding FA, Iovannisci DM, Brinson E., Grossman P., Winn-Deen E. S. 1995 Human Mutation, 5: 153-65). En este caso, las sondas en el soporte comprenden típicamente el nucleótido o nucleótidos variables en el extremo 3' de la sonda.

45 **[0256]** El marcaje puede llevarse a cabo durante la amplificación post-hibridación. El marcaje puede ser por

marcaje directo usando, por ejemplo, fluoróforos, enzimas, isótopos radiactivos, etc., o por marcaje indirecto usando, por ejemplo, parejas de unión específicas que incorporan fluoróforos, enzimas, etc., mediante el uso de métodos convencionales, tales como los mencionados previamente en relación con el marcaje de productos de amplificación o fragmentación.

5 **[0257]** La amplificación post-hibridación puede llevarse a cabo, por ejemplo, usando la metodología de “extensión de cebadores”. Típicamente, después de la hibridación, se lleva a cabo una reacción de extensión de los oligonucleótidos híbridos en el soporte (por ejemplo, un portaobjetos de vidrio). La extensión puede llevarse a cabo con nucleótidos marcados directa o indirectamente y sólo tendrá lugar si el extremo 3' del oligonucleótido hibrida perfectamente con el producto de amplificación.

10 **[0258]** La extensión de cebadores es un método conocido para la discriminación de genotipos (Pastinen T, Raitio M, Lindroos K, Tainola P, Peltonen L, Syvanen AC. 2000 Genome Research 10: 1031-42.) y puede realizarse de varias formas diferentes. En una estrategia usada comúnmente, se diseñan un conjunto de sondas oligonucleótidas específicas de alelo para hibridar con las secuencias diana. Las sondas difieren entre sí en su nucleótido del extremo 3', que para cada sonda está diseñado para que sea complementario a uno de los posibles nucleótidos polimórficos en una posición dada.

15 **[0259]** Cuando el nucleótido 3' de la sonda es complementario a la secuencia bajo ensayo, entonces el emparejamiento de bases consiguiente permite que una ADN polimerasa extienda el cebador oligonucleotídico por incorporación de nucleótidos adicionales que pueden estar marcados directa o indirectamente, permitiendo de este modo la identificación posterior de las sondas que se han extendido y las que no lo han hecho. Las sondas que se extienden con éxito llevan el nucleótido complementario al SNP en su extremo 3' permitiendo de este modo que se determine el genotipo de la muestra de ensayo. También se han desarrollado estrategias similares, por ejemplo, el sistema de mutaciones refractarias a la amplificación (ARMS).

20 **[0260]** Como alternativa, puede llevarse a cabo una reacción de ligación post-hibridación, por ejemplo, usando metodología de OLA. Después de la hibridación, se lleva a cabo una reacción de ligación de los oligonucleótidos hibridados en el soporte (por ejemplo, portaobjetos de vidrio) con oligonucleótidos marcados. Una ligación tendrá lugar solamente si el extremo 3' de la sonda depositada en el soporte hibrida perfectamente con el ADN diana (por ejemplo, producto de amplificación).

25 **[0261]** El ensayo de ligación de oligonucleótidos (OLA) es otro método para cuestionar SNP (Eggerding FA, Iovannisci DM, Brinson E., Grossman P., Winn-Deen E.S. 1995 Human Mutation, 5: 153-65). El OLA usa una pareja de sondas oligonucleotídicas que hibridan con segmentos adyacentes de ADN diana incluyendo la base variable. La sonda diseñada para hibridar con el extremo 5' del nucleótido polimórfico es un oligonucleótido específico de alelo (ASO) para uno de los alelos diana. La última base en el extremo 3' de este ASO se sitúa en el sitio del polimorfismo del ADN diana; el ASO también tiene típicamente una molécula de biotina en su extremo 5' que funciona como un “gancho” que puede usarse posteriormente para recuperar el oligonucleótido en virtud de la interacción altamente específica que experimenta la biotina con la estreptavidina. El oligómero en el lado a la derecha o 3' de la pareja es el oligómero común (la secuencia es la misma para los dos o más alelos diferentes que se desee ensayar). El oligómero común se sitúa en un sitio invariable al lado del polimorfismo del ADN diana y está marcado fluorescentemente en su extremo 3'.

30 **[0262]** Si el ASO es perfectamente complementario a la secuencia diana, el ASO hibrida completamente cuando hibrida y se situará plano contra esa diana permitiendo que la ADN ligasa una covalentemente el ASO al oligómero común. Después de la reacción de ligación, el gancho de biotina se usa para eliminar el ASO y, por ejemplo, también se eliminará el oligómero común marcado fluorescentemente, produciendo una fluorescencia detectable.

35 **[0263]** Cuando el ASO no es un emparejamiento perfecto con la secuencia diana, la hibridación es incompleta y la base 3' del oligómero no presentará emparejamiento de bases con el ADN diana, impidiendo de este modo la ligación. En estas circunstancias, cuando se usa el gancho de biotina para eliminar el ASO, el oligonucleótido común no se eliminará y, por lo tanto, no hay marcador detectable, por ejemplo fluorescencia, en la molécula eliminada.

40 **[0264]** Para distinguir entre dos alelos conocidos que difieran en una sola base, son necesarios tres oligonucleótidos: dos son oligonucleótidos específicos de alelo (ASO) que difieren entre sí solamente en la base terminal 3'; el primero es complementario a un alelo y el segundo es complementario al segundo alelo. El tercer oligonucleótido es complementario a la secuencia invariable adyacente a la base variante.

45 **[0265]** Una vez que ha tenido lugar la hibridación (y opcionalmente la amplificación post-hibridación), puede determinarse la intensidad de marcador detectable en cada posición de sonda (incluyendo sondas de control). La intensidad de la señal (el valor de intensidad bruto) es una medida de la hibridación en cada sonda.

50 **[0266]** La intensidad de marcador detectable en cada posición de sonda (cada repetición de sonda) puede determinarse usando cualquier medio adecuado. Los medios seleccionados dependerán de la naturaleza de l

marcador. En general, un dispositivo apropiado, por ejemplo, un explorador, recoge la imagen de la microplaca de ADN hibridada y revelada. Se captura una imagen y se cuantifica.

5 **[0267]** En un caso, por ejemplo, cuando se usa un marcaje fluorescente, después de la hibridación (opcionalmente después de la amplificación o ligación post-hibridación), la microplaca de ADN hibridada y revelada se pone en un explorador para cuantificar la intensidad de marcaje en los puntos en los que ha tenido lugar la hibridación. Aunque puede usarse prácticamente cualquier explorador, en una realización se usa un explorador confocal de fluorescencia. En este caso, la microplaca de ADN se pone en dicho aparato y la señal emitida por el fluoróforo debido a la excitación por un láser se explora para cuantificar la intensidad de señal en los puntos en los que ha
10 tenido lugar la hibridación. Los ejemplos no limitantes de exploradores que pueden usarse de acuerdo con la presente invención incluyen exploradores comercializados por las compañías siguientes: Axon, Agilent, Perkin Elmer, etc.

15 **[0268]** Típicamente, en la determinación de la intensidad de un marcador detectable en cada posición de sonda (es decir, para cada repetición de sonda) se tiene en cuenta la interferencia de fondo, que se elimina. La interferencia de fondo surge debido a la unión inespecífica a la matriz de sondas y puede determinarse por medio de controles incluidos en la matriz. Una vez que se ha determinado la intensidad de la señal de fondo, ésta puede restarse del valor de intensidad bruto para cada repetición de sonda para obtener un valor de intensidad puro. Típicamente, el fondo local, basado en la intensidad de señal detectada en las proximidades de cada característica individual, se resta del valor de intensidad de señal bruto. Este fondo se determina a partir de la intensidad de señal en un área predeterminada que rodea a cada característica (por ejemplo, un área de X, Y o Z μm^2 centrada en la posición de la sonda).

20 **[0269]** La señal de fondo se determina típicamente a partir de la señal local de controles "blancos" (solamente disolvente). En muchos casos, el dispositivo, por ejemplo, explorador, que se usa para determinar intensidades de señal proporcionará medios para determinar la señal de fondo.

25 **[0270]** Por lo tanto, por ejemplo, cuando el marcador es un marcador fluorescente, los valores de fluorescencia absolutos (valores de intensidad bruta) pueden recopilarse para cada repetición de sonda y las interferencias de fondo asociadas con cada repetición de sonda también pueden evaluarse para producir valores "puros" para la intensidad de señal en cada posición de sonda.

30 **[0271]** Una vez que se ha hibridado el ADN diana a la microplaca y que se ha determinado la intensidad de marcador detectable en las posiciones de repeticiones de sondas en la microplaca (los valores de intensidad brutos), es necesario proporcionar un método (modelo) que pueda relacionar los datos de intensidad de la microplaca con el genotipo del individuo.

35 **[0272]** Los inventores han descubierto que esto puede realizarse aplicando un algoritmo adecuado a los datos de intensidad. El algoritmo y el *software* informático desarrollados por los inventores permiten el análisis de las variaciones genéticas con una sensibilidad y reproducibilidad suficientes como para permitir su uso en un entorno clínico. El algoritmo usa tres funciones lineales que caracterizan a cada uno de los tres genotipos AA, AB y BB para una variación genética dada. El método implica generalmente cotejar los valores de intensidad para todas las repeticiones de cada sonda, para calcular un valor de intensidad promedio para cada sonda. Opcionalmente, los valores de intensidad brutos para cada repetición pueden corregirse para tener en cuenta la interferencia de fondo (para obtener un valor de intensidad puro) antes de que se cotejen los valores de intensidad para cada una de las repeticiones.

40 **[0273]** En general, para una variación genética dada, el análisis e interpretación de una microplaca comprende las etapas siguientes:

50 (a) proporcionar la intensidad de marcador detectable en cada repetición para cada una al menos cuatro sondas (sondas 1, 2, 3 y 4) proporcionadas para la detección de la variación genética (el valor de intensidad bruto), en la que:

- 55
- la sonda 1 detecta (es capaz de hibridar específicamente con) la variación genética A (por ejemplo, un alelo normal) y la sonda 2 detecta (es capaz de hibridar específicamente con) la variación genética B (por ejemplo, un alelo mutante);
 - la sonda 3 detecta (es capaz de hibridar específicamente con) la variación genética A (por ejemplo, un alelo normal) y la sonda 4 detecta (es capaz de hibridar específicamente con) la variación genética B (por ejemplo, un alelo mutante); y
 - las sondas 1 y 2 forman una primera pareja de sondas y las sondas 3 y 4 forman una segunda pareja de sondas;
- 60

65 (b) opcionalmente corregir el valor de intensidad bruto para cada repetición para tener en cuenta la interferencia de fondo, obteniendo de este modo un valor de intensidad puro;

- (c) cotejar los valores de intensidad (opcionalmente puros) para cada una de las repeticiones de cada sonda y determinar un valor de intensidad promedio para cada sonda;
 (d) calcular los cocientes 1 y 2, en los que:

5
$$\text{Cociente 1} = \frac{\text{valor de intensidad promedio para sonda 1}}{\text{valor de intensidad promedio para sonda 1} + \text{valor de intensidad promedio para sonda 2}}$$

y

10
$$\text{Cociente 2} = \frac{\text{valor de intensidad promedio para sonda 3}}{\text{valor de intensidad promedio para sonda 3} + \text{valor de intensidad promedio para sonda 4}}$$

- 15 (e) introducir los cocientes 1 y 2 en cada una de tres funciones lineales que caracterizan a cada uno de los tres genotipos posibles, AA, AB y BB, en la que:

20 la Función 1 es la función lineal que caracteriza a individuos con el genotipo AA y consiste en una combinación lineal de los cocientes 1 y 2;

la Función 2 es la función lineal que caracteriza a individuos con el genotipo AB y consiste en una combinación lineal de los cocientes 1 y 2;

25 la Función 3 es la función lineal que caracteriza a individuos con el genotipo BB y consiste en una combinación lineal de los cocientes 1 y 2;

las funciones lineales están formadas por coeficientes que acompañan a las variables cociente 1 y 2;

- 30 (f) determinar cuál de las tres funciones lineales tiene el mayor valor; y
 (g) determinar de este modo el genotipo del individuo para la variación genética.

[0274] Por lo tanto, la función lineal correspondiente al genotipo de ese individuo tendrá el mayor valor absoluto.

35 **[0275]** Los inventores han descubierto que el uso de repeticiones y promedios calculados a partir de repeticiones es importante para el funcionamiento fiable de la invención. El uso de las funciones acelera el análisis y permite una mejor discriminación.

[0276] Preferentemente, la capacidad de discriminación entre los tres genotipos es de (aproximadamente) el 100%. Si la discriminación es inferior al 100%, preferentemente las sondas se rediseñan.

40 **[0277]** El valor de intensidad bruto para cada repetición de sonda puede determinarse de acuerdo con los métodos descritos anteriormente. Por lo tanto, pueden seleccionarse secuencias de sondas y repeticiones como se describe en la presente memoria. En un ejemplo, se usan 4 sondas por variación genética y se usan 6, 8 ó 10 repeticiones por sonda.

45 **[0278]** Típicamente, la corrección del valor de intensidad bruto para obtener el valor de intensidad puro para cada repetición de sonda comprende estar la interferencia de fondo del valor bruto. La interferencia de fondo se determina típicamente usando controles apropiados como se describe en la presente memoria.

50 **[0279]** Típicamente, el cálculo del valor de intensidad promedio comprende eliminar valores extremos o atípicos. Por lo tanto, cuando los valores de intensidad (opcionalmente puros) de cada una de las repeticiones de sondas se cotejan, los valores atípicos pueden identificarse y excluirse de consideraciones posteriores. En una realización, los valores atípicos componen entre el 10% y el 50%, por ejemplo, el 15, 20, 25, 30, 35, 40 ó 45% de los valores obtenidos. En una realización, se eliminan el 40% de los valores. En una realización, se usan 4 sondas con 6, 8 ó 10 repeticiones por sonda y los valores extremos o atípicos componen entre el 10% y el 50% de los valores obtenidos.

55 **[0280]** Se conocen en la técnica varias funciones lineales adecuadas. Estas funciones pueden usarse en un análisis de discriminación lineal para los fines de la presente invención.

60 **[0281]** En un aspecto, la invención se refiere por lo tanto a un método o modelo computacional (algoritmo) para determinar el genotipo con respecto a una variación genética dada usando los cocientes 1 y 2 en las tres funciones lineales que se han definido anteriormente (etapas e y f). El método puede producir por lo tanto en una realización un resultado de genotipo (AA, AB o BB) a partir de una entrada de datos de cocientes 1 y 2. El método también puede incluir calcular uno o ambos cocientes 1 y 2 (etapa d). En algunas realizaciones, el método comprende además calcular un valor de intensidad promedio para cada sonda (etapa c) y/o calcular un valor de intensidad puro

para cada repetición de sonda (etapa b). Por lo tanto, la introducción de datos en el modelo puede comprender uno o más de los valores de intensidad promedios, valores de intensidad de repeticiones puros o valores de intensidad de repeticiones brutos. El método puede comprender además determinar el valor de intensidad bruto para cada repetición de sonda (etapa a). El método puede comprender una o más de las etapas anteriores.

[0282] Para llevar a cabo los métodos anteriores, los coeficientes para las funciones lineales deben determinarse en un procedimiento de entrenamiento usando datos de individuos de control cuyo genotipo para la variación genética ya se conoce. Se conocen en la técnica métodos de entrenamiento. Típicamente, en dichos métodos se usan datos de entrada (en este caso, típicamente los cocientes 1 y 2) para los que ya se conoce el resultado (en el presente caso, el genotipo). Los coeficientes se sustituyen en las tres ecuaciones lineales aleatoriamente y se calcula el resultado. Basándose en ese resultado, se alteran uno o más coeficientes y los datos introducidos se introducen de nuevo para producir otro resultado. El procedimiento continúa hasta que se obtienen coeficientes que optimizan el resultado deseado. Estos coeficientes optimizados se usan después en las funciones lineales cuando el método se aplica a datos de ensayo (cuando el resultado se desconoce todavía).

[0283] Para entrenar el presente modelo, se obtienen los cocientes 1 y 2 para n individuos de control que tienen genotipo AA (por ejemplo, tipo silvestre homocigoto), n individuos de control que tienen genotipo AB (heterocigoto) y n individuos de control que tienen genotipo BB (por ejemplo, mutante homocigoto). Los cocientes pueden obtenerse usando los métodos descritos anteriormente. Los cocientes se introducen como anteriormente y los coeficientes se alteran en un análisis discriminatorio hasta que se obtienen tres funciones lineales que maximizan la discriminación entre los grupos AA, AB y BB. Estos coeficientes se usan después en las tres funciones cuando el modelo se usa en muestras de ensayo desconocidas (cuando el genotipo no está predeterminado).

[0284] Por lo tanto, en un aspecto la invención proporciona un método de obtención de funciones lineales para su uso en los presentes métodos de genotipado. El método comprende típicamente llevar a cabo las etapas de los métodos de genotipado que se describen, para n individuos de control que tienen genotipo AA (por ejemplo, tipo silvestre homocigoto), n individuos de control que tienen genotipo AB (heterocigoto) y n individuos de control que tienen genotipo BB (por ejemplo, mutante homocigoto) con respecto a una variación genética. Los valores de intensidad obtenidos para cada una de las repeticiones de sondas se recopilan como se ha descrito y se aplica un algoritmo.

[0285] Como se describe para los métodos de genotipado, la aplicación del algoritmo comprende calcular un valor de intensidad promedio para cada sonda y el algoritmo usa tres funciones lineales destinadas a caracterizar cada uno de los tres genotipos posibles, AA, AB y BB para la variación genética dada. Los coeficientes se insertan en las funciones de una forma repetitiva hasta que se obtienen funciones que maximizan la discriminación entre los genotipos en un análisis discriminatorio. Esto proporciona los coeficientes para su uso en las funciones lineales cuando el método o algoritmo está en uso operativo (es decir, para determinar el genotipo de individuos de ensayo).

[0286] El algoritmo o método que usa las tres funciones lineales para analizar los datos de intensidad puede ser como se ha descrito anteriormente.

[0287] En algunos casos, el método de entrenamiento permite una optimización por retroalimentación. Por lo tanto, a medida que se obtienen valores de intensidad y cocientes para individuos de ensayo y estos se genotipan, los datos de intensidad, por ejemplo, los cocientes, y el genotipo se introducen y se vuelven a calcular los coeficientes para las funciones lineales.

[0288] En un aspecto, la invención se refiere a un método computacional para entrenamiento. El método puede usarse para obtener funciones lineales para su uso en los presentes métodos de genotipado usando los cocientes 1 y 2 obtenidos para cada uno de n individuos que tienen genotipo AA, n individuos que tienen genotipo AB y n individuos que tienen genotipo BB con respecto a una variación genética. Los cocientes pueden obtenerse mediante los métodos descritos anteriormente. El método comprende típicamente aplicar el algoritmo que usa las tres funciones lineales (Funciones 1, 2 y 3) destinadas a caracterizar cada uno de los tres genotipos posibles AA, AB o BB para la variación genética de modo que:

la Función 1 es la función lineal que caracteriza a individuos con el genotipo AA y consiste en una combinación lineal de los cocientes 1 y 2;

la Función 2 es la función lineal que caracteriza a individuos con el genotipo AB y consiste en una combinación lineal de los cocientes 1 y 2;

la Función 3 es la función lineal que caracteriza a individuos con el genotipo BB y consiste en una combinación lineal de los cocientes 1 y 2; y

las funciones lineales están formadas por coeficientes que acompañan a las variables cociente 1 y 2;

y obtener funciones lineales que maximicen la discriminación entre los tres grupos genotípicos AA, AB y BB en un análisis discriminatorio, para obtener los coeficientes que pueden usarse en las funciones lineales cuando el algoritmo se usa en un método de ensayo (es decir, está en uso operativo para determinar el genotipo).

5 **[0289]** El algoritmo o método que usa las tres funciones lineales para analizar los datos de intensidad puede ser como se ha descrito anteriormente.

10 **[0290]** El método de entrenamiento computacional puede implicar además calcular los cocientes 1 y 2 a partir del valor de intensidad promedio proporcionado para cada una de las sondas y/o cotejar los valores de intensidad de repeticiones de sondas para determinar un valor de intensidad promedio para cada sonda y/o corregir un valor de intensidad bruto para una repetición de sonda para tener en cuenta la interferencia de fondo, obteniendo de este modo valores de intensidad puros para las repeticiones.

15 **[0291]** En algunos casos, el método computacional también permite una etapa de optimización por retroalimentación como se describe.

[0292] Típicamente, en el entrenamiento $n \geq$, por ejemplo, 3, 4, 5, 6, 7, 8, 9 ó 10. En un aspecto, $n \geq 5$. En algunos casos, n puede ser de 10 a 50 o más, por ejemplo, de 15 a 40 o de 25 a 35, tal como 20 ó 30.

20 **[0293]** Las sondas y repeticiones de sondas para el método de entrenamiento se seleccionan como se describe en la presente memoria. En una realización, se usan 4 sondas para cada variación genética, con 6, 8 ó 10 repeticiones de cada sonda. Una vez seleccionadas, las sondas usadas en el entrenamiento se usan también cuando el modelo está en uso operativo (para determinar un genotipo desconocido). Si las sondas se alteran, típicamente el modelo debe reentrenarse para optimizar la discriminación con las nuevas sondas.

25 **[0294]** Preferentemente, los coeficientes son tales que la discriminación entre los tres grupos genotípicos (tanto en uso de entrenamiento como operativo) es sustancialmente del 100%. Si la discriminación no es del 100%, preferentemente las sondas se rediseñan.

30 **[0295]** Como anteriormente, el modelo también puede experimentar una optimización por retroalimentación cuando está en uso operativo. En ese caso, el modelo se usa primero para determinar el genotipo de un individuo (AA, AB o BB). Los cocientes 1 y 2 para ese individuo se introducen después en el modelo y los coeficientes en las funciones lineales se alteran según sea necesario para optimizar la discriminación entre los tres grupos genotípicos. De este modo, los datos adicionales recopilados a medida que se usa el modelo pueden usarse para optimizar la capacidad de discriminación de las funciones lineales.

35 **[0296]** Existen varios parámetros que pueden determinarse y optimizarse para optimizar el rendimiento y la fiabilidad del modelo o método analítico.

40 (i) En un aspecto, los cocientes 1 y 2 determinadas para un individuo se incluyen dentro del intervalo de cocientes 1 y 2 usado para entrenar el modelo (es decir, para optimizar las tres funciones lineales). Si se desea esto puede proporcionar por lo tanto un doble ensayo para el genotipo de un individuo.

45 (ii) En un aspecto, la intensidad de fluorescencia promedio de $4n$ repeticiones (donde "n" es el número de repeticiones para cada sonda, por ejemplo, 6, 8 ó 10), por ejemplo, 40 repeticiones, con respecto a la interferencia de fondo es superior a 5.

50 (iii) En un aspecto, la variación entre valores de intensidad (brutos o puros) para repeticiones de la misma sonda es de un mínimo. Por ejemplo, el coeficiente de variación entre los valores de intensidad para las repeticiones de una sonda dada es preferentemente inferior a 0,25.

55 (iv) En un aspecto, el cociente de la suma de los valores de intensidad brutos para todas las repeticiones de sondas en una microplaca respecto a la intensidad de la interferencia de fondo es superior a 15 cuando se usa un explorador de fluorescencia.

60 (v) En un aspecto, el valor de intensidad de señal bruto obtenido para los controles negativos es ≤ 3 veces superior al valor de intensidad de la interferencia de fondo. Por ejemplo, los controles negativos pueden incluir el "blanco" de DMSO y los oligonucleótidos que no hibridan mencionados anteriormente. La interferencia de fondo es la señal obtenida de las regiones de la matriz en las que no se ha aplicado puntualmente sonda, y pueden determinarse como anteriormente.

[0297] Preferentemente, uno o más de cualquiera de (i) a (v) es aplicable cuando la intensidad es la intensidad de fluorescencia de un marcador fluorescente, en particular cuando la intensidad se determina por medio de un explorador fluorescente confocal.

65

- 5 [0298] Asegurar que el modelo cumple uno o más de los anteriores contribuye a proporcionar fiabilidad y reproducibilidad. Uno o más de cualquiera de (i) a (v) puede ser cierto para el modelo. Preferentemente, el modelo cumple el (i) anterior. En un ejemplo, (i), (i i) y (i ii) son ciertos. En otro ejemplo, (i ii), (i v), (v) son ciertos. Preferentemente, todos los anteriores son ciertos para el modelo. Esto es aplicable tanto al uso de entrenamiento como operativo.
- 10 [0299] Como anteriormente, los cocientes obtenidos experimentalmente conseguidos para una muestra de ensayo pueden compararse con los cocientes obtenidos previamente para las (n) muestras de control obtenidas de individuos de genotipo conocido, donde n es como anteriormente, habitualmente >5, o >10 o >20. Los cocientes de referencia obtenidos del análisis de las muestras de control permite que se asigne un genotipo a la muestra de ensayo. Por lo tanto, éste puede ser un ensayo doble.
- 15 [0300] En un caso, el método o algoritmo analítico de la invención comprende una secuencia de las etapas siguientes: usar 4 sondas (2 parejas de sondas) repetidas (6, 8 ó 10 repeticiones), calcular la intensidad promedio de cada sonda a partir de las intensidades co-tejadas de las repeticiones; calcular los cocientes 1 y 2 como anteriormente para las 2 parejas de sondas (para detectar las variaciones genéticas A y B); sustituir los cocientes 1 y 2 obtenidos en tres ecuaciones lineales que se han obtenido en un análisis discriminatorio usando los cocientes 1 y 2 calculados para "n" pacientes de control con genotipo AA, "n" pacientes de control con genotipo AB y "n" pacientes de control con genotipo BB (con respecto a la variación genética) (en un experimento "n" es 5); y determinar el genotipo de un paciente para la variación genética (para cada variación genética incluida en la microplaca de ADN) basándose en qué función lineal tiene el mayor valor absoluto. Los cocientes de ensayo también pueden compararse con los cocientes de los "n" pacientes de control para determinar cada genotipo.
- 20 [0301] El análisis y la interpretación anteriores se han descrito con respecto a una variación genética. Sin embargo, debe entenderse que la presente microplaca incluye generalmente sondas para la detección de múltiples variaciones genéticas que pueden analizarse al mismo tiempo. Por lo tanto, los presentes métodos incluyen el análisis de múltiples variaciones genéticas, como se describen en la presente memoria, en paralelo.
- 25 [0302] En un aspecto adicional, la invención se refiere a un sistema informático que comprende un procesador y medios para controlar el procesador para llevar a cabo un método computacional de la invención.
- 30 [0303] La invención se refiere además a un programa informático que comprende un código de programa informático que cuando se ejecuta en un ordenador o red de ordenadores provoca que el ordenador o red de ordenadores lleven a cabo un método computacional de la invención. El programa informático puede almacenarse en un medio legible por ordenador.
- 35 [0304] Además de las sondas y microplacas descritas en la presente memoria, los inventores también han diseñado y validado cebadores oligonucleotídicos que son capaces de amplificar, por ejemplo, por medio de PCR múltiple, regiones de ADN diana que contienen las variaciones genéticas humanas asociadas con RA de la Tabla 1A. Estos cebadores son útiles en la preparación de un ácido nucleico para su uso en los presentes métodos de genotipado, pronóstico y terapéuticos.
- 40 [0305] La Tabla 3 enumera parejas de cebadores que amplifican regiones de ADN diana que contienen las variaciones genéticas asociadas a RA de la Tabla 1A (SEC ID N°: 361-540) junto con la variación genética correspondiente.
- 45 [0306] Los cebadores oligonucleotídicos enumerados tienen la ventaja de permitir la amplificación específica de dichas regiones de ADN diana en un número muy pequeño de reacciones de PCR. Los cebadores enumerados permiten, en un número mínimo de reacciones de PCR múltiple, la amplificación de todos los fragmentos necesarios para genotipar las variaciones genéticas de la Tabla 1A, y que pueden analizarse en el Artchip.
- 50 [0307] En un aspecto adicional, la presente invención se refiere a cada uno de los cebadores de PCR enumerados en la Tabla 3, y en particular a cada una de las parejas de cebadores de PCR enumeradas y a su uso en la amplificación por PCR, por ejemplo, en una reacción de PCR múltiple, de una región de ADN diana que contiene la variación genética correspondiente. La invención proporciona en un aspecto cualquiera de estos cebadores o parejas de cebadores para su uso en medicina, en particular para su uso en los presentes métodos de genotipado, pronóstico o terapéuticos.
- 55 [0308] La invención se refiere además a un kit de amplificación por PCR que comprende al menos una pareja de cebadores de PCR enumerados. El kit puede incluir además, por ejemplo, una polimerasa (termoestable), dNTP, un tampón adecuado, cebadores adicionales y/o instrucciones para su uso, por ejemplo, para amplificar una región de ADN diana que contiene la variación genética correspondiente. El kit puede usarse para la amplificación de regiones de ADN diana a partir de muestras de ácido nucleico, para su uso en los presentes métodos.
- 60 [0309] En otro aspecto, la presente invención se refiere a un kit de genotipado o diagnóstico (preferentemente *in*
- 65

vitro) para genotipar variaciones genéticas asociadas a RA y/o para diagnosticar RA o susceptibilidad a RA. El kit comprende una microplaca o matriz de ADN de acuerdo con la invención. El kit puede comprender además instrucciones para el uso de la microplaca en un método de genotipado de la invención, por ejemplo, instrucciones para su uso en el presente método o algoritmo analítico. Los componentes adicionales de un kit pueden incluir:

- *software* informático, un programa informático o un sistema informático de acuerdo con la invención;
- uno o más cebadores de PCR o parejas de cebadores de PCR de acuerdo con la invención; y/o
- un kit de amplificación por PCR de acuerdo con la invención.

[0310] Las sondas para la microplaca o cebadores de PCR pueden seleccionarse como anteriormente dependiendo de las variaciones genéticas a detectar o del fin diagnóstico del kit.

[0311] El kit puede contener uno o más controles positivos y/o negativos de la reacción de hibridación.

[0312] La invención se refiere además al uso del kit en un método de genotipado, pronóstico o terapéutico de la invención.

[0313] Como se describe en la presente memoria, los presentes métodos de genotipado son útiles para diagnosticar RA o susceptibilidad a RA en un sujeto. Los resultados de genotipado obtenidos en los métodos pueden usarse para determinar el pronóstico y pueden ser útiles en la determinación del tratamiento apropiado para la RA (por ejemplo, por predicción de la respuesta a la terapia).

[0314] La RA presenta varios fenotipos, más particularmente en términos de gravedad de la enfermedad. La enfermedad leve se distingue de la enfermedad grave y destructiva, así como la velocidad y la naturaleza de la progresión de la enfermedad, y esta heterogeneidad clínica está correlacionada con una heterogeneidad genética.

[0315] Las variaciones genéticas particulares asociadas con RA pueden ser predictivas de fenotipos particulares o del desarrollo de fenotipos particulares y, por lo tanto, de la progresión de la enfermedad. En otras palabras, puede ser que haya una asociación estadísticamente significativa entre, por ejemplo, el alelo mutante B de una variación genética dada y la aparición/desarrollo de un fenotipo particular.

[0316] Puesto que los presentes métodos de genotipado permiten un genotipado fiable de múltiples variaciones genéticas en un entorno clínico, estos pueden usarse para genotipar individuos de fenotipo de RA conocido y, de este modo, identificar variaciones genéticas predictivas de fenotipos de RA particulares.

[0317] En un aspecto, la invención se refiere por lo tanto a un método de identificación de variaciones genéticas predictivas de un fenotipo de RA particular, tal como los fenotipos enumerados anteriormente. El método implica genotipar una pluralidad de individuos con respecto a una o más variaciones genéticas asociadas con RA usando un método de la invención. En dicho estudio retrospectivo, se genotipan típicamente 300-1000 individuos, por ejemplo, pueden genotiparse 400, 500 ó 600 individuos. El fenotipo de cada individuo ya se conoce basándose en procedimientos clínicos convencionales.

[0318] Una vez que se obtienen los genotipos, estos datos se comparan con los datos de fenotipo y se identifican asociaciones estadísticamente significativas entre genotipos particulares y fenotipos particulares. Se conocen en la técnica métodos para determinar la significación estadística.

[0319] Las variaciones genéticas identificadas como predictivas de fenotipos/cursos de enfermedad particulares pueden usarse después para diagnosticar estos fenotipos/cursos de enfermedad en individuos de ensayo, por genotipado de los individuos con respecto a la variación o variaciones genéticas predictivas. Por lo tanto, es posible determinar el curso probable de progresión de la enfermedad en el individuo. El genotipado puede realizarse por cualquier método apropiado, dependiendo del número de variaciones a ensayar. Por ejemplo, puede usarse un método de genotipado de la invención. Como alternativa, pueden ser apropiados métodos basados en secuencias u otros métodos basados en microplacas.

[0320] Por lo tanto, en un aspecto, la invención se refiere además a un método de diagnóstico del fenotipo de RA o de predicción del curso probable de progresión de la enfermedad en un individuo por determinación del genotipo del individuo con respecto a una o más variaciones genéticas que se han identificado como predictivas (del fenotipo o curso patológico de RA particular) mediante los métodos descritos en la presente memoria.

[0321] Una vez que se ha realizado la predicción, entonces será posible seleccionar la estrategia terapéutica más adecuada, por ejemplo, determinar la necesidad de intervención quirúrgica.

[0322] Las presentes matrices y métodos proporcionan por lo tanto un medio para que los médicos puedan predecir el curso probable de la progresión de la enfermedad en pacientes con RA y contribuir también a la selección del régimen de tratamiento más adecuado. Por lo tanto, son herramientas de pronóstico útiles. La información de

genotipo obtenido e acuerdo con la presente invención puede contribuir a la toma de decisiones clínicas o al diagnóstico en casos en los que los síntomas (fenotipo patológico) se amaniguan. La información genética proporcionada por el Artchip u otros métodos también podría contribuir a determinar la verosimilitud del desarrollo de la enfermedad en individuos asintomáticos (por ejemplo, miembros familiares inmediatos de los que padecen RA) permitiendo, por ejemplo, que se proporcionen directrices sobre el estilo de vida y la dieta y que se indique la necesidad de un control continuo de individuos que tengan una constitución genética que indique una posible susceptibilidad al desarrollo de la enfermedad.

[0323] En un aspecto, la invención se refiere por lo tanto a un método de diagnóstico de RA o susceptibilidad a RA en un individuo, o determinación del curso probable de la progresión de la enfermedad en un individuo como anteriormente. Preferentemente, el método es *in vitro*. La invención se refiere además a un método de selección de un tratamiento, para un individuo que tiene RA, en algunos casos cuando al individuo se le ha diagnosticado o se ha ensayado de acuerdo con los métodos de la invención. Todavía además la invención se refiere en algunos aspectos a métodos de tratamiento de un individuo que padece RA en los que, después de seleccionarse el tratamiento, el tratamiento se administra al individuo.

[0324] Los métodos de diagnóstico, predictivos y terapéuticos pueden comprender llevar a cabo un método de genotipado de la invención como se describe en la presente memoria. Cualquiera de los métodos puede implicar llevar a cabo un método de entrenamiento de la invención como se describe en la presente memoria para obtener funciones lineales para su uso en la determinación del genotipo. Los métodos adicionales pueden comprender el uso de una microplaca, sistema informático, programa informático, sondas oligonucleotídicas o parejas o conjuntos de sondas, cebadores oligonucleotídicos o parejas de cebadores, kit de amplificación por PCR o kit de diagnóstico de la invención, como se describen en la presente memoria.

[0325] Aparte de la contribución al diagnóstico y tratamiento de la RA y al desarrollo de nuevas estrategias terapéuticas para esta enfermedad, la presente invención es útil para elucidar la fisiopatología de la reacción inflamatoria en la RA, que también será de gran interés para el estudio de otras enfermedades, de base autoinmune, y que pertenecen a campos tan diversos tales como la neumología, la dermatología, etc.

[0326] En un aspecto, la presente invención se refiere a una micromatriz adaptada para su uso en los presentes métodos como se describe en la presente memoria.

[0327] La invención se refiere además al uso de una o más sondas oligonucleotídicas y/o uno o más cebadores o parejas de cebadores de la invención en un método para pronosticar la RA, tal como un método descrito en la presente memoria.

[0328] A continuación se ilustrarán aspectos adicionales de la invención con referencia a las Figuras y a la ejemplificación experimental adjuntas, a modo de ejemplo y no como limitación. Serán evidentes aspectos y realizaciones adicionales para los expertos en la materia.

EJEMPLOS

[0329] Aunque, en general, las técnicas mencionadas en la presente memoria son bien conocidas en la materia, puede hacerse referencia en particular a Sambrook et al, 1989, Molecular Cloning: a laboratory manual.

Ejemplo 1 - Detección de variaciones genéticas humanas asociadas a RA usando una microplaca de ADN de acuerdo con la invención (Artchip)

1.1 Diseño de la microplaca de ADN para genotipar variaciones genéticas asociadas a RA

[0330] Se diseñó y fabricó una microplaca de ADN para detectar variaciones genéticas humanas asociadas con RA que permite una detección simultánea, sensible, específica y reproducible. Dichas variaciones genéticas están relacionadas con un mayor o menor riesgo de padecer RA, una mejor o peor respuesta al tratamiento y también un mejor o peor pronóstico de la enfermedad. Se muestran en la Tabla 1A y 1B ejemplos ilustrativos de variaciones genéticas humanas asociadas con antígenos relacionados con la RA que pueden determinarse usando esta microplaca de ADN.

[0331] La microplaca de ADN diseñada y fabricada consiste en un soporte (portaobjetos de vidrio), que muestra una pluralidad de sondas en su superficie, que permite la detección de variaciones genéticas previamente mencionadas. Estas sondas son capaces de hibridar con las secuencias amplificadas de los genes relacionados con RA. Las secuencias de ADN de cada una de las sondas usadas se enumeran en la Tabla 2 (Figura 2). En general, se indica el nombre del gen y la mutación (cambio de nucleótido, "ins" inserción, "del" delección o cambio de aminoácido). Todas las sondas enumeradas se han validado técnicamente.

[0332] Las sondas se enumeran a continuación en conjuntos de sondas, consistiendo cada conjunto en 4 o más

sondas. Las sondas enumeradas corresponden a las SEC ID N°: 1-360, dándose a las sondas consecutivas SEC ID N° consecutivos. Por ejemplo, en las sondas enumeradas a continuación, el conjunto para el análisis del polimorfismo C1672T corresponde a las SEC ID N°: 1-4, respectivamente, con la primera sonda enumerada SEC ID N°: 1 y la cuarta sonda enumerada en el conjunto, SEC ID N°: 4. De forma similar, por ejemplo, las sondas enumeradas a continuación en el conjunto para el análisis del polimorfismo G1154A (VEGF) corresponden a las SEC ID N°: 37-40, con la primera sonda enumerada SEC ID N°: 37 y la cuarta sonda enumerada en el conjunto SEC ID N°: 40.

1.2 Producción de la microplaca de ADN

Impresión y procesamiento de los portaobjetos de vidrio

[0333] Las sondas capaces de detectar las variaciones genéticas previamente identificadas se imprimen sobre soportes (portaobjetos de vidrio) recubiertos con aminosilano usando DMSO como disolvente. La impresión se lleva a cabo usando una impresora o aplicador puntual de oligonucleótidos (sondas) al tiempo que controlan la temperatura y la humedad relativa.

[0334] La unión de las sondas al soporte (portaobjetos de vidrio) se lleva a cabo por medio de entrecruzamiento con radiación ultravioleta y calentamiento como se describe en la documentación proporcionada por el fabricante (por ejemplo, Corning LifeSciences <http://www.corning.com>). La humedad relativa durante el procedimiento de deposición se mantiene entre el 40-50% y la temperatura alrededor de 20°C.

1.3 Validación de la utilidad clínica de la microplaca de ADN

1.3.1 Preparación de la muestra a hibridar

[0335] El ADN del individuo se extrae de una muestra de sangre mediante un protocolo de filtración convencional. (Por ejemplo, kits comerciales de Macherey Nagel, Qiagene etc.).

[0336] Las regiones de ADN diana que contienen las variaciones genéticas de interés se amplifican mediante PCR múltiple usando parejas apropiadas de cebadores oligonucleotídicos. Puede usarse cualquier pareja de oligonucleótidos adecuada que permita una amplificación específica de fragmentos genéticos en los que pueda existir una variación genética a detectar. Ventajosamente, se usan aquellas parejas de cebadores oligonucleotídicos que permitan que dichas amplificaciones se realicen en el menor número posible de reacciones de PCR.

[0337] Los cebadores oligonucleotídicos usados para amplificar por PCR regiones diana que contienen las variaciones genéticas de la Figura 1 se enumeran en la Tabla 3 (Figura 3). Estos cebadores representan un aspecto adicional de la invención.

[0338] Las reacciones de PCR múltiple se llevan a cabo simultáneamente en las mismas condiciones de tiempo y temperatura que permiten la amplificación específica de los fragmentos genéticos en los que podrían existir las variaciones genéticas a detectar. Una vez que ha terminado la PCR múltiple, se usa análisis en gel de agarosa para comprobar que ha tenido lugar la reacción de amplificación.

[0339] A continuación, la muestra a hibridar (productos de amplificación) se somete a fragmentación con una ADNasa y los productos de fragmentación resultantes se someten a marcaje indirecto. Una transferasa terminal añade un nucleótido, unido covalentemente a un miembro de una pareja de moléculas que se unen específicamente entre sí (por ejemplo, biotina que permite la unión específica posterior a estreptavidina) en los extremos de estos fragmentos de ADN pequeños.

[0340] Antes de aplicar la muestra a la microplaca de ADN, la muestra se desnatura por calentamiento a 95°C durante 5 minutos, y después se añade el "tampón de hibridación del kit ChipMap" (Ventana Medical System).

1.3.2 Hibridación

[0341] La hibridación se lleva a cabo automáticamente en una estación de hibridación tal como la Ventana Discovery (Ventana Medical Systems) que se ha desarrollado específicamente para tal uso. Como alternativa, la hibridación puede realizarse manualmente.

[0342] La prehibridación y el bloqueo de los portaobjetos se llevan a cabo con BSA. A continuación, se aplica la solución de hibridación (tampón de hibridación del kit ChipMap, Ventana Medical System) a la superficie de la microplaca de ADN, que se mantiene a 45°C durante 1 hora siguiendo el protocolo de Ventana 9.0 Europe (Ventana Medical System). Por último, los portaobjetos se someten a diferentes soluciones de limpieza (tampones del kit de hibridación ChipMap, Ventana Medical System). Una vez que ha terminado el procedimiento de hibridación, comienza la limpieza y el secado final de los portaobjetos.

[0343] Cuando ha tenido lugar la hibridación, la microplaca de ADN se revela por incubación con una molécula marcada fluorescentemente que es capaz de unirse específicamente a la molécula incorporada en el producto de amplificación por transferasa terminal (por ejemplo, en el caso de incorporación de biotina puede usarse un fluoróforo acoplado a estreptavidina tal como estreptavidina-Cy3) para marcar las posiciones de sondas en las que se ha producido la hibridación.

1.3.3 Exploración de los portaobjetos

[0344] Los portaobjetos se ponen en un explorador confocal fluorescente, por ejemplo Axon 4100^a, y la señal emitida por el fluoróforo se explora cuando se estimula por el láser.

1.3.4 Cuantificación de la imagen

[0345] El propio software del explorador permite la cuantificación de la imagen obtenida a partir de la señal en los puntos en los que ha tenido lugar la hibridación.

1.3.5 Interpretación de los resultados

[0346] A partir de la señal obtenida con las sondas que detectan las diferentes variaciones genéticas, se establece el genotipo del individuo. En el primer caso, el software del explorador ejecuta una función para restar la interferencia de fondo local del valor de intensidad de señal absoluto obtenido para cada sonda. A continuación, las repeticiones para cada una de las 4 sondas que se usan para caracterizar cada variación genética se agrupan. Se calcula el valor de intensidad promedio para cada una de las 4 sondas usando el promedio cotejado a partir de las repeticiones para identificar los valores anormales (atípicos) que pueden excluirse de consideraciones adicionales. Una vez que se conoce el valor de intensidad promedio para cada una de las sondas, entonces se calculan dos cocientes (cociente 1 y cociente 2):

$$\text{Cociente 1} = \frac{\text{Intensidad promedio para sonda 1}}{\text{Intensidad promedio para sonda 1} + \text{Intensidad promedio para sonda 2}}$$

$$\text{Cociente 2} = \frac{\text{Intensidad promedio para sonda 3}}{\text{Intensidad promedio para sonda 3} + \text{Intensidad promedio para sonda 4}}$$

en los que la sonda 1 detecta (es capaz de hibridar específicamente con) la variación genética A (por ejemplo, un alelo normal), la sonda 2 detecta (es capaz de hibridar específicamente con) la variación genética B (por ejemplo, un alelo mutante), la sonda 3 detecta (es capaz de hibridar específicamente con) la variación genética A (por ejemplo, un alelo normal) y la sonda 4 detecta (es capaz de hibridar específicamente con) la variación genética B (por ejemplo, un alelo mutante).

[0347] Estos cocientes se sustituyen en tres funciones lineales que caracterizan a cada uno de los tres genotipos posibles.

AA	Función 1
AB	Función 2
BB	Función 3

[0348] La función que presenta el mayor valor absoluto determina el genotipo del paciente.

[0349] En este caso, las funciones lineales se obtienen por análisis de 5 sujetos para cada uno de los tres genotipos posibles de la variación genética (AA, AB, BB). Con los resultados, se calculan los cocientes 1 y 2 para los 15 sujetos. Estos cocientes son variables de clasificación para los tres grupos para crear las funciones lineales, con las que se evalúa la capacidad discriminativa de las dos parejas de sondas diseñadas. Si la capacidad discriminativa no es del 100%, las sondas se rediseñan. Nuevos sujetos caracterizados para cada uno de los tres genotipos componen nuevos cocientes 1 y 2 para perfeccionar las funciones lineales y en breve, mejorar la capacidad discriminativa del algoritmo basado en estas tres funciones.

[0350] Cuando se usa un explorador fluorescente confocal, para obtener resultados fiables es preferible que los cocientes 1 y 2 estén dentro del intervalo de los cocientes usados para construir los grupos, la intensidad de fluorescencia promedio de las 4n (por ejemplo 40) repeticiones con respecto a la interferencia de fondo sea superior a 5 y el coeficiente de variación de todas las repeticiones de la microplaca de ADN esté por debajo de 0,25.

[0351] De nuevo, cuando se usa un explorador confocal fluorescente en el experimento, para que una hibridación completa se considere fiable, preferentemente el cociente de la intensidad de fluorescencia de sonda respecto a la

interferencia de fondo de todas las sondas de la microplaca de ADN está por encima de 15. Asimismo, el promedio de todos los cocientes está preferentemente por encima de 0,6 y el control negativo es preferentemente inferior a o igual a 3 veces la interferencia de fondo.

5 **[0352]** Resumiendo, en este caso se presentan 4 sondas (repetidas 10 veces) en el portaobjetos para la detección de cada mutación. Dos de las sondas detectan una variación genética (A) y las otras dos la otra variación genética (B). La base examinada se localiza en la posición central de las sondas.

10 **[0353]** Un sujeto homocigoto para la variación genética A no mostrará la variación genética B. Por consiguiente, en la imagen obtenida a partir del soporte de vidrio las sondas que detectan la variación genética B mostrarán una señal de hibridación significativamente inferior que la mostrada por la variación A y viceversa. En este caso, los cocientes 1 y 2 resultarán 1 y a los sujetos se les asignará como homocigotos AA por el análisis de *software*.

15 **[0354]** Por otro lado, un sujeto heterocigoto para la variación genética determinada muestra ambas variaciones genéticas. Por lo tanto, las sondas que las detectan muestran una señal de hibridación equivalente. Los cocientes 1 y 2 resultarán 0,5 y al sujeto se le asignará como heterocigoto AB por el análisis de *software*.

20 **[0355]** Los cebadores oligonucleotídicos usados para amplificar (por medio de PCR múltiple) regiones diana que contienen variaciones genéticas asociadas a RA en la Figura 1 se enumeran en la Tabla 3. Dichos cebadores oligonucleotídicos representan un aspecto adicional de la presente invención.

25 **[0356]** Los cebadores oligonucleotídicos usados para amplificaciones por PCR se enumeran en la Tabla 3. Estos corresponden con las SEC ID N°: 361-540, estando los cebadores numerados consecutivamente según se enumeran.

Ejemplo 2 - Establecimiento de modelos para predecir fenotipos de RA

Métodos

30 Diseño de estudio

[0357] Se inscribieron pacientes con RA que cumplieran los criterios del ACR de 1987 de cinco Departamentos de Reumatología de hospitales universitarios en España: Hospital Puerta de Hierro (Madrid), Hospital Virgen de las Nieves (Granada), Hospital Clinic (Barcelona), Hospital La Paz (Madrid) y Hospital Juan Canalejo (A Coruña). Para tener una gestión terapéutica homogénea, los pacientes incluidos en el estudio tenían que haber sido diagnosticados y tratados en un Servicio de Reumatología después del 1 de enero de 1990, es decir, las muestras debían tener al menos cinco años de seguimiento. Los criterios de inclusión completos eran los siguientes: 1) pacientes descritos de acuerdo con los criterios del Colegio Americano de Reumatología (ACR), 2) aparición de RA después de los 18 años de edad, 3) con erosiones radiológicas, 4) caucásicos, 5) más de 5 años de evolución y 6) diagnosticados después del 1 de enero de 1990. Se incluyeron en el estudio un total de 375 individuos que se ajustaban a esos criterios.

45 **[0358]** Se registraron las variables clínicas y analíticas basales de cada paciente: edad de aparición, número de articulaciones implicadas, anti-péptido cíclico citrulinado, velocidad de sedimentación eritrocitaria, factor reumatoide, proteína C reactiva, género, edad y estado de tabaquismo. El estudio cumplía las recomendaciones de la Declaración de Helsinki (World Medical Association) y la EMEA (European Medicines Agency).

Fenotipos estudiados

50 **[0359]** Después de cinco años de seguimiento se consideraron los siete resultados siguientes: 1) Agresividad; la predicción de RA agresiva se analizó estudiando diferentes subfenotipos: 1.1) Pacientes con un cuestionario de evaluación de la salud (HAQ) superior a 2, en los que indica que la RA está causando una incapacidad funcional importante. 1.2) Pacientes que muestran múltiples erosiones en manos y pies. 1.3) Pacientes que reciben más de tres tratamientos diferentes. 1.4) Pacientes cuya enfermedad les obligó a dejar sus trabajos. 1.5) Pacientes que recibieron una prótesis articular debido a la RA. 2) Respuesta a la terapia; la predicción de pacientes que muestran intolerancia a metotrexato. 3) Remisión; definida como ausencia completa de síntomas articulares sin terapia durante al menos 5 años. Cada paciente se clasificó como No/Sí (0/1) respecto a esos siete fenotipos.

Genotipado y selección de polimorfismos de un solo nucleótido (SNP)

60 **[0360]** Se obtuvo sangre periférica (2 ml) de cada paciente, se puso en un tubo tratado con EDTA. Se extrajo el ADN plasmático con el minikit sanguíneo de ADN QIAamp (Qiagen) siguiendo las especificaciones del fabricante. El genotipado se llevó a cabo usando la microplaca de ADN ARTchip. Se genotiparon para cada paciente varios SNP que pertenecían a genes que codifican proteínas implicadas en la inmunidad innata y adaptativa y en el metabolismo del hueso y del cartílago. La selección de SNP estaba basada en los datos previos publicados, las terapias farmacológicas emergentes y la propia experiencia en investigación de los inventores.

[0361] La selección de SNP en esos genes estaba basada en una frecuencia alélica minoritaria de 0,1. Sólo se tuvieron en cuenta "TagSNP" ($R^2 < 0,8$) ya que estos proporcionaban más potencia estadística por reducción de los grados de libertad (df) de los ensayos.

TAG Directo	SEC ID N°: 554	GCTAGATGAAGAGCAAGCGC
TAG Inverso	SEC ID N°: 555	TACAACCGACAGATGTATGT

[0362] De todos los SNP genotipados, sólo se incluyó una fracción en el análisis de regresión logística por etapas para limitar el índice de falsos positivos global. En primer lugar, se realizaron pruebas de chi cuadrado (X^2) para ensayar la conformidad con las expectativas de Hardy-Weinberg (HWE) de los polimorfismos genéticos bajo análisis. Sólo se incluyeron en este estudio los SNP que concordaban con la ley de HWE en ambos grupos separados bajo análisis. Los SNP con desviaciones extremadamente elevadas de las predicciones de HWE (valores p inferiores a 0,01) se excluyeron del análisis ya que las desviaciones podían indicar problemas tales como errores de genotipado. Además, se llevaron a cabo ensayos de asociación de un solo locus entre la frecuencia alélica de SNP (asociaciones alélicas) y el fenotipo del paciente usando la prueba de X^2 de contingencia convencional, y se determinaron los valores p, incluyendo la corrección de Bonferroni para ensayos múltiples. La posibilidad de que las desviaciones de las HWE en la población global (ambos fenotipos bajo análisis juntos) pudiera ser importante para causar la enfermedad también se investigó por combinación del efecto de la asociación alélica y las HWE totales. El producto del valor p de HWE y el valor p de la asociación alélica se usó para clasificar los SNP por orden de importancia. Aquellos con los menores valores p se incluyeron en el análisis de regresión. Todos los análisis genéticos se llevaron a cabo usando el *software* HelixTree® (Golden Helix, Inc., Bozeman, MT, Estados Unidos).

Modelado estadístico

[0363] El análisis estadístico se llevó a cabo entre los grupos de pacientes (No-0/Sí-1) para intentar la discriminación entre ambos grupos para cada uno de los siete fenotipos diferentes estudiados. Se evaluaron siete modelos diferentes: un modelo para distinguir entre HAQ<2 y HAQ>2 (modelo 1); un modelo para distinguir pacientes con o sin múltiples erosiones en manos y pies (modelo 2); un modelo para distinguir pacientes que recibieron un número igual a o superior a o inferior a tres tratamientos diferentes (modelo 3); un modelo para distinguir pacientes cuya enfermedad les obligó a dejar sus trabajos (Sí o No) (modelo 4); un modelo para distinguir pacientes que recibieron una prótesis articular debido a RA (Sí o No) (modelo 5); un modelo para distinguir pacientes que mostraban intolerancia a metotrexato (Sí o No) (modelo 6) y un modelo para distinguir pacientes que consiguieron una remisión (Sí o No) (modelo 7).

[0364] Los análisis estadísticos se realizaron usando el Paquete Estadístico para las Ciencias Sociales (SPSS Inc. Headquarters, Chicago, IL, Estados Unidos) versión 11.0. Se analizaron múltiples asociaciones de genotipo-fenotipo por medio de regresión logística multivariable (LR directa) con fenotipos patológicos determinados clínicamente como variables dependientes y los loci individuales y datos clínicos y analíticos como variables independientes. La calidad de ajuste de los modelos se evaluó usando estadística de Hosmer-Lemeshow y su precisión se evaluó calculando el área bajo la curva (AUC) de la curva de características operativas del receptor (ROC) con intervalos de confianza del 95%. La variabilidad explicada de los modelos basándose en los SNP se evaluó por medio del R^2 Nagelkerke. Para medir el impacto de los SNP y las variables incluidas en los modelos de los fenotipos analizados, se calculó la sensibilidad, la especificidad y el cociente de verosimilitud positiva ($LR+ = \text{sensibilidad}/(1-\text{especificidad})$) por medio de curvas de ROC.

Resultados

[0365] Para diferenciar entre HAQ>2 (modelo 1) se introdujeron 3 predictores (2 SNP + 1 variable clínica) en el modelo de LR directa. Para el modelo 2 (múltiples erosiones) se introdujeron 4 predictores SNP + 2 variables clínicas en el modelo. Para el modelo 3 (3 o más tratamientos) se introdujeron 3 predictores SNP + 4 variables clínicas en el modelo. Para el modelo 4 (dejar el trabajo) se introdujeron solamente 3 predictores SNP en el modelo. Para el modelo 5 (prótesis articular) se introdujeron 4 predictores SNP + 2 variables clínicas en el modelo. Para el modelo 6 (intolerancia a metotrexato) se introdujo 1 predictor SNP + 1 variable clínica en el modelo. Por último, para el modelo 7 (conseguir un estado de remisión) se introdujeron 2 predictores SNP + 1 variable clínica en el modelo. Las variables de SNP y variables clínicas incluidas en cada modelo se enumeran en la Tabla 11A (Figura 11A).

[0366] La información respecto a las variables (clínicas y SNP) restantes en cada función se muestra en las Tablas 4 a 10. Se construyeron funciones de probabilidad de regresión usando el Paquete Estadístico para las Ciencias Sociales (SPSS Inc. Headquarters, Chicago, IL, Estados Unidos) versión 14.0. SPSSv14. B es el coeficiente asociado a cada genotipo en la función de probabilidad. ET es el error en el cálculo de B. Wald es la prueba estadística. GL son los grados de libertad. Sig. P valor de B para la prueba de Wald. Exp (B) es el riesgo relativo.

[0367] La contribución de los factores genéticos y clínicos a los fenotipos de RA estudiados puede demostrarse adicionalmente mediante la proporción sustancial de varianza (R^2 Nagelkerke) explicada por las funciones (26,0% para el modelo 1; 43,6% para el modelo 2; 20,1% para el modelo 3; 18,3% para el modelo 4; 15,3% para el modelo

5; 18,3% para el modelo 6 y 26% para el modelo 7). Se obtuvieron las funciones de probabilidad y curvas de ROC para cada fenotipo analizado. Se proporcionan las curvas de ROC, la sensibilidad, la especificidad y los cocientes de verosimilitud positiva (LR+) de todos los modelos en las Figuras 4 a 10.

5 **[0368]** Nagelkerke R2 es una forma de medir la proporción de variantes explicadas por la función. El área bajo la curva de ROC (AUC de ROC) es una medida de rendimiento de ensayo o precisión diagnóstica. El cociente de verosimilitud positiva (LR+) se calcula como sensibilidad/1-especificidad.

10 Discusión

10 **[0369]** Los cocientes de verosimilitud son una forma útil y práctica de expresar la potencia de ensayos de diagnóstico. Cuatro de los siete modelos descritos en esta patente presentan un valor de LR relativamente elevado, poniendo por tanto de manifiesto la capacidad de las combinaciones de SNP estudiadas para predecir ese fenotipo particular. Las altas AUC de ROC obtenidas para estos modelos proporcionan pruebas adicionales de la elevada potencia discriminativa de las combinaciones de SNP usadas. La utilidad de la magnitud del AUC de ROC como herramienta para evaluar la fuerza de la relación entre genotipos y enfermedad se ha descrito anteriormente. El uso de estos SNP para obtener un perfil genético del paciente proporciona una herramienta extra para el médico para diferenciar entre pacientes con diferente curso de la enfermedad, posibilidad de remisión, presencia de erosiones principales en los rayos X y número de tratamientos necesarios. Para los modelos restantes se obtuvo un menor LR probablemente debido al hecho de que el fenotipo estudiado también depende de factores ambientales no calculados en las funciones de los presentes inventores. Sin embargo, con estos modelos basados principalmente en factores genéticos, los médicos pueden considerar de forma razonable la posibilidad de una intervención quirúrgica, de que un paciente pudiera dejar su trabajo y la posibilidad de una terapia diferente. Los síntomas clínicos junto con ensayos bioquímicos de rutina forman parte del diagnóstico de la RA y son necesarios para una orientación terapéutica y de pronóstico correctas. Los modelos descritos en la presente memoria basándose tanto en la genética como en datos bioquímicos de laboratorio son adecuados para su uso durante el seguimiento de pacientes con RA, y permiten la identificación de sus tipos de pacientes bien definidos, proporcionando indicios importantes sobre su futuro tratamiento.

30

REIVINDICACIONES

1. Un método de pronóstico de un fenotipo de remisión de la artritis reumatoide (RA) en un sujeto que tiene RA, comprendiendo dicho método determinar (i) el genotipo de un individuo al menos en la posición de polimorfismo de un solo nucleótido en rs2476601; y (ii) una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC BI) en el suero del sujeto.
2. Un método de acuerdo con la reivindicación 1, en el que el método comprende:
- (I) obtener resultados:
- (i) de la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en el alelo nulo del gen de la Glutación S-transferasa M1 (GSTM1) y en rs2476601; y
- (ii) de una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC BI) en el suero del sujeto; y
- (II) usar los resultados obtenidos en (I) para pronosticar el fenotipo de remisión de RA.
3. Un método de acuerdo con la reivindicación 1 o la reivindicación 2, que comprende además pronosticar un fenotipo de artritis reumatoide (RA) agresivo en el sujeto, comprendiendo dicho fenotipo de RA agresivo uno o más de: una puntuación de cuestionario de evaluación de la salud (HAQ) de > 2 (HAQ >2); la presencia de múltiples erosiones en las manos y en los pies (RX2); recibir 3 o más tratamientos diferentes (TTO_30M); verse forzado a dejar el empleo debido a la RA (dejar el trabajo); y recibir una prótesis articular (intervención quirúrgica), comprendiendo dicho método:
- (I) obtener resultados de acuerdo con el modelo 1, modelo 2, modelo 3, modelo 4 y/o modelo 5 para el sujeto; y
- (II) usar los resultados obtenidos en (I) para pronosticar el fenotipo de RA agresivo,
- en el que:
- (a) cuando el fenotipo de RA agresivo comprende una puntuación de cuestionario de evaluación de la salud (HAQ) de > 2 (HAQ >2), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 1, comprendiendo dichos resultados:
- (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs1800790 y rs2070874;
- (ii) el nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC) en el suero del sujeto; y
- (iii) la velocidad de sedimentación eritrocitaria (VSG) en la sangre del sujeto,
- (b) cuando el fenotipo de RA agresivo comprende la presencia de múltiples erosiones en las manos y en los pies (RX2), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 2, comprendiendo dichos resultados:
- (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs1544410, rs2071592, rs187238 y rs1801275;
- (ii) una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC BI) en el suero del sujeto; y
- (iii) el número de articulaciones dolorosas (NoARTDOL) mencionadas por el sujeto,
- (c) cuando el fenotipo de RA agresivo comprende recibir 3 o más tratamientos diferentes (TTO_30M), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 3, comprendiendo dichos resultados:
- (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs7531668, rs419598 y rs1024611;
- (ii) el nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCC) en el suero del sujeto;
- (iii) el nivel de anticuerpo Factor Reumatoide (RF) en el suero del sujeto;
- (iv) el nivel de proteína C reactiva (PCR) en la sangre del sujeto; y
- (v) la edad en años a la que se le diagnosticó al sujeto la RA,
- (d) cuando el fenotipo de RA agresivo comprende verse forzado a dejar el empleo debido a la RA (dejar el trabajo), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 4, comprendiendo dichos resultados:
- (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo

nucleótido en cada uno de rs1042522, el alelo nulo del gen de la Glutación S-transferasa T1 (GSTT1) y rs1801133,

5 (e) cuando el fenotipo de RA agresivo comprende recibir una prótesis articular (intervención quirúrgica), dicha etapa (I) comprende obtener los resultados de acuerdo con dicho modelo 5, comprendiendo dichos resultados:

- 10 (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en cada uno de rs2268277, rs1042522, rs10735810 y rs3025039;
- (ii) una representación binaria del nivel de anticuerpo anti-péptido cíclico citrulinado (ANTI-PCCBI) en el suero del sujeto; y
- (iii) la velocidad de sedimentación eritrocitaria (VSG) en la sangre del sujeto.

15 4. El método de acuerdo con la reivindicación 1 o la reivindicación 2, en el que el método comprende además pronosticar el fenotipo de RA de intolerancia a metotrexato, en el que:

- dicha etapa (I) comprende además obtener resultados de acuerdo con el modelo 6, comprendiendo dichos resultados:
- (i) la identidad del nucleótido en el ADN genómico del sujeto en la posición del polimorfismo de un solo nucleótido en rs1801133; y
- 20 (ii) la edad en años a la que se le diagnosticó al sujeto la RA.

5. El método de acuerdo con cualquiera de las reivindicaciones 2 a 4, en el que en la etapa (II), sólo se usan los resultados obtenidos en la etapa (I) para pronosticar el fenotipo de RA.

25 6. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la obtención de resultados para variables de SNP comprende la secuenciación o análisis de micromatrices.

30 7. El método de acuerdo con cualquiera de las reivindicaciones 1 a 5, en el que la obtención de resultados para variables de SNP comprende el uso de una o más parejas de sondas oligonucleotídicas indicadas en la Figura 2.

8. El método de acuerdo con cualquiera de las reivindicaciones 1 a 5, en el que la obtención de resultados para variables de SNP comprende la amplificación de ácido nucleico obtenido del sujeto, y el método comprende el uso de una o más parejas de cebadores oligonucleotídicos indicadas en la Figura 3.

35 9. El método de acuerdo con cualquiera de las reivindicaciones anteriores, en el que la etapa (II) comprende:

- (i) introducir los resultados determinados en la etapa (I) en una función de probabilidad, calculando de este modo un valor de función de probabilidad; y
- 40 (ii) comparar el valor de la función de probabilidad con los valores de la función de probabilidad calculados para individuos de fenotipo conocido.

10. El método de acuerdo con cualquiera de las reivindicaciones anteriores que comprende el uso de una función de probabilidad derivada de los datos de cualquiera de las Tablas 4 a 10.

45 11. Un método de selección de un tratamiento adecuado para tratar la RA en un sujeto, comprendiendo el método:

- (a) pronosticar un fenotipo de RA en el sujeto mediante un método de acuerdo con cualquiera de las reivindicaciones anteriores; y
- 50 (b) seleccionar un tratamiento que sea adecuado para el fenotipo de RA determinado.

Figura 1: Tabla 1A: Variaciones genéticas asociadas con RA

5	El polimorfismo C1672T de la familia de transportadores de soluto 22 A4 (SLC22A4).
	El polimorfismo C24658G del factor de transcripción relacionado con Runt 1 (RUNX1).
	El polimorfismo G117A de la peptidilarginina desiminasa tipo IV (PADI 4).
10	El polimorfismo G55S de la peptidilarginina desiminasa tipo IV (PADI 4).
	El polimorfismo V82A de la peptidilarginina desiminasa tipo IV (PADI 4).
15	El polimorfismo G112A de la peptidilarginina desiminasa tipo IV (PADI 4).
	El polimorfismo Hla25 del gen relacionado con la cadena A del complejo mayor de histocompatibilidad clase I (MICA).
20	El polimorfismo Hla28 del gen relacionado con la cadena A del complejo mayor de histocompatibilidad clase I (MICA).
	El polimorfismo Hla58 T/G de NOTCH4.
25	El polimorfismo G1154A del factor de crecimiento endotelial vascular (VEGF).
	El polimorfismo C634G del factor de crecimiento endotelial vascular (VEGF).
	El polimorfismo C936T del factor de crecimiento endotelial vascular (VEGF).
30	El polimorfismo A607C de interleucina 18 (IL18).
	El polimorfismo G-137C de interleucina 18 (IL18).
35	El polimorfismo Q551R del receptor de interleucina 4 (IL-4R).
	El polimorfismo T2018C del alelo 2 del antagonista del receptor de IL-1 (IL1RN*2).
	El polimorfismo C2073T del alelo 2 del antagonista del receptor de IL-1 (IL1RN*2).
40	El polimorfismo C872T de muerte programada 1 (PD-1, PDCD1).
	El polimorfismo Asp299Gly del receptor tipo peaje 4 (TLR4).
45	El polimorfismo Apos.VS7+283(b>B) del receptor de vitamina D (1,25- dihidroxivitamina D3)
	El polimorfismo C-547T del dominio de inmunoglobulina de célula T y dominio de mucina 3 (TIM3).
	El polimorfismo T4259G del dominio de inmunoglobulina de célula T y dominio de mucina 3 (TIM3).
50	El polimorfismo UTR 3' de uroquinasa (-A2849G).
	El antígeno de linfocito T citotóxico 4 (CTLA4A/G) en el exón 1.
55	El polimorfismo Ala566Thr del transportador asociado con el procesamiento de antígenos 2 (TAP2).
	El polimorfismo I kappa BL A-62T. Polimorfismo C1362T de poli (ADP-ribosa) polimerasa 1 (PARP-1).
	El polimorfismo T1858C de proteína tirosina fosfatasa.
60	El polimorfismo Fok I del receptor de vitamina D.
	El polimorfismo G38A de uteroglobina (UG).
65	El polimorfismo 6A/6A de metaloproteínasa de la matriz 3 (MMP3).

ES 2 375 577 T3

- Los genotipos Arg/Arg y Pro/Pro de proteína tumoral p53 en el codón 72.
El polimorfismo (-A2849G) de interleucina 10 (IL-10).
- 5 Los polimorfismos -308, +488, -857 y -238 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa).
El polimorfismo T587G del receptor del factor de necrosis tumoral 2 (TNFR2).
El polimorfismo +36 A/A del receptor del factor de necrosis tumoral 1 (TNFR1).
- 10 El polimorfismo G-455A del polipéptido beta de fibrinógeno B (FGB).
El polimorfismo de genotipo nulo de la variante 1 del transcrito de glutatión S-transferasa M1 GSTM1.
- 15 El polimorfismo de genotipo GSTT1-0 de la glutatión S-transferasa T1 (GSTT1).
El polimorfismo (-C3435T) de la proteína de multi-resistencia a fármacos (MDR-1).
El polimorfismo R241G de la molécula de adhesión 1 (ICAM1).
- 20 El polimorfismo C3954T de interleucina 1 beta (IL-1B).
El polimorfismo Val58Ile de condromodulina II (Chm II).
- 25 El polimorfismo -622 del gen de interleucina 6 (IL-6).
El 347 GG de 5-aminoimidazol-4-carboxamida ribonucleótido formiltransferasa/IMP ciclohidrolasa (ATIC).
Polimorfismo G80A del miembro 4 de la familia de transportadores de solutos 22 (SLC22A4).
- 30 El polimorfismo C677T y A1298C de metilentetrahidrofolato reductasa (MTHFR).
El polimorfismo 249, 365 y 720 del gen de linfotóxina alfa (LTA).
El alelo -1087 G/G de interleucina 10 (IL-10) y el alelo de factor de necrosis tumoral-alfa (TNF)-308 TNF1/TNF1 (Estas variaciones se asocian con RA en combinación, pero no en solitario).
- 35 El factor de crecimiento transformante beta TGFbeta1 (alelo C poco frecuente en el codón 25) y el antagonista del receptor de IL1 (alelo A2 en el intrón 2) (Estas variaciones se asocian con RA en combinación, pero no en solitario).
Polimorfismo T869C de codón 10.
- 40 El polimorfismo -1087 AA del gen de interleucina 10 (IL-10).
El polimorfismo 28 2C>T, 590G >A, 341T >C, 481C >T, 803A>G, 85 7G>A y 19 1G>A de N -acetiltransferasa 2 (arilamina N-acetiltransferasa).
- 45 Los polimorfismos G460A y A719G de tiopurina metiltransferasa (TPMT).
El polimorfismo -C196T de proteína de tipo receptor de Fc 3 (FCRL3).
- 50 El polimorfismo Arg260Thr del gen de tapasina (TPSN)
El haplotipo A de poli (ADP-ribosa) polimerasa 1 (PARP-1) (410T-[A](10)-[CA](10-12)-1362C, que incluye alelos CA de PARP-1 cortos) y haplotipo B (410C-[A](11)-[CA](13-20)-1362T).
- 55 Los polimorfismos Hla 76 y Hla 78 del antígeno de histocompatibilidad.
El polimorfismo del receptor de melatonina 1B (MTNR1B).
El polimorfismo (-168A→G) del transactivador del complejo mayor de histocompatibilidad clase II (MHC2TA).
- 60 El polimorfismo del receptor BETA-2-adrenérgico.
El polimorfismo A1298C de metilentetrahidrofolato reductasa (MTHFR).
- 65 El polimorfismo (-2518)G>A de la proteína quimiotáctica de monocitos 1 (MCP1).

ES 2 375 577 T3

- El polimorfismo menos 843C/T del gen del LIGANDO FAS.
- El polimorfismo IBD5 locusIGR2060a_1 del gen de la enfermedad inflamatoria del intestino.
- 5 El polimorfismo V158F (G/T) del gen cd16a.
- El polimorfismo menos 511a/c de interleucina 1 beta (IL1b).
- 10 El polimorfismo 34c/t de interleucina 4 (IL4).
- El polimorfismo C93T de caspasa 9.
- El polimorfismo K469E de la molécula de adhesión intracelular 1 (ICAM1).
- 15 El polimorfismo Gly54Asp del gen de unión a lectina, manosa (MBL).
- El polimorfismo -590 del gen de interleucina 4 (IL4).
- 20 El polimorfismo T145C (Leu49Ser) del gen de TIOPURINA S-METILTRANSFERASA 5 (TPMT5).
- El polimorfismo A539T del gen de TIOPURINA S-METILTRANSFERASA 6 (TPMT6).
- El polimorfismo T681G (His227Gln) del gen de TIOPURINA S-METILTRANSFERASA 7 (TPMT7).
- 25 El polimorfismo G644A (Arg215His) del gen de TIOPURINA S-METILTRANSFERASA 8 (TPMT8).
- El polimorfismo HLA b27 del antígeno de histocompatibilidad.
- 30 El G80A de la familia de transportadores de solutos 19 (folato reductasa) SLC19A1

Figura 1: Tabla 1B

n° SNP	Símbolo de Gen	Nombre de Gen	Polimorfismo
1	SLC22A4	Familia de transportadores de soluto 22 A4	C1672T
2	RUNX1	Factor de transcripción relacionado con Runt 1	C24658G
3	PADI4	Peptidilarginina desiminasa tipo IV	G117A
4	PADI4	Peptidilarginina desiminasa tipo IV	G55S
5	PADI4	Peptidilarginina desiminasa tipo IV	V82A
6	PADI4	Peptidilarginina desiminasa tipo IV	G112A
7	MICA	Gen relacionado con cadena A del complejo mayor de histocompatibilidad clase I	Hla25
8	MICA	Gen relacionado con cadena A del complejo mayor de histocompatibilidad clase I	Hla28
9	NOTCH4	Homólogo 4 de Notch	Hla58 T/G
10	VEGF	Factor de Crecimiento Endotelial Vascolar	G1154A
11	VEGF	Factor de Crecimiento Endotelial Vascolar	C634G
12	VEGF	Factor de Crecimiento Endotelial Vascolar	C836T
13	IL18	Interleucina 18	A607C
14	IL18	Interleucina 18	G-137C
15	IL-4R	Receptor de interleucina 4	Q551R
16	IL1RN*2	Alelo 2 del antagonista del receptor de IL-1	T2018C
17	IL1RN*2	Alelo 2 del antagonista del receptor de IL-1	C2073T
18	PD-1.PDCC1	Muerte programada 1	C872T
19	TLR4	Receptor de tipo peaje 4	Asp299Gly
20	VDR	Receptor de vitamina D (1,25-dihidroxitiamina D3)	AposV57+283b>B
21	TIM3	Dominio de inmunoglobulina de célula T y dominio de mucina 3	C-547T
22	TIM3	Dominio de inmunoglobulina de célula T y dominio de mucina 3	T4259G
23	PLAU	Activador de plasminogeno, uroquinasa	3' UTR -A2849G
24	CTLA4	Antígeno de linfocito T citotóxico 4	A/G en exón 1
25	TAP2	Transportador asociado con procesamiento de antígenos 2	Ala565Thr
26	NFKBIL1	Factor nuclear de potenciador del gen del polipéptido ligero kappa en tipo inhibidor de células B 1	A-62T
27	PARP-1	Poli ADP-ribosa polimerasa 1	C1362T
28	PTP	Proteína tirosina fosfatasa	T1858C
29	VDR	Receptor de vitamina D	FokI
30	UG	Uteroglobina	G38A
31	MMP3	Metaloproteinasa de la matriz 3	6A/6A
32	p53	Proteína tumoral	Genotipos Arg/Arg y Prol/Pro en el codón 72

33	IL-10	Interleucina 10		
34	TNF alfa	Promotor del Gen del Factor de Necrosis Tumoral Alfa	-A2849G	
35	TNF alfa	Promotor del Gen del Factor de Necrosis Tumoral Alfa	el -308	
36	TNF alfa	Promotor del Gen del Factor de Necrosis Tumoral Alfa	el +488	
37	TNF alfa	Promotor del Gen del Factor de Necrosis Tumoral Alfa	el -867	
38	TNFR2	Receptor del factor de necrosis tumoral 2	el -238	
39	TNFR1	Receptor del factor de necrosis tumoral 1	T587G	
40	FGB	Polipéptido beta de fibrinógeno B	+36 A/A	
41	GSTM1	Variante 1 del transcrito de glutatión S-transferasa M1	G-455A	
42	GSTT1	Glutatión S-transferasa T1	genotipo nulo	
43	MDR-1	proteína de multi-resistencia a fármacos	genotipo GSTT1-0	
44	ICAM1	Molécula de adhesión 1	-C3435T	
45	IL-1B	Interleucina 1 beta	R241G	
46	Chm II	Condromodulina II	C3954T	
47	IL-6	Gen de interleucina 6	Val58Ile	-622
48	ATIC	5-Aminoimidazol-4-carboxamida ribonucleótido formiltransferasa/IMP ciclohidrolasa	347 GG	
49	SLC22A4	Miembro 4 de la familia de transportadores de soluto 22	G80A	
50	MTHFR	Metilentetrahidfolato reductasa	C677T	
51	LTA	Gen de linfotoxina alfa	249	
52	LTA	Gen de linfotoxina alfa	365	
53	LTA	Gen de linfotoxina alfa	720	
54	IL-10	Interleucina 10	-1087G/G	
55	TNF	Factor de necrosis tumoral alfa	-308 TNFWNFI	
56	TGFbeta1 .	Factor de crecimiento transformante		
57	IL-10	Gen de interleucina 10	-1087 AA	
58	TPMT	Tiopurina metiltransferasa	G480A	
59	TPMT	Tiopurina metiltransferasa	A719G	
60	FCRL3	Proteína de tipo receptor de Fc 3	-C196T	
61	TPSN	Gen de tapasina	Arg260Thr	
62	NAT2	N-Acetiltransferasa 2	341T>C	
63	NAT2	N-Acetiltransferasa 2	481C>T	
64	NAT2	N-Acetiltransferasa 2	803A>G	
65	NAT2	N-Acetiltransferasa 2	867G>A	
66	NAT2	N-Acetiltransferasa 2	191G>A	

67 Hla	Antígeno de histocompatibilidad	78
68 Hla	Antígeno de histocompatibilidad	76
69 MTR1B.	Receptor de melatonina 1B	
70 MHC2TA.	Transactivador del complejo mayor de histocompatibilidad clase II	
71 ADRB2	Receptor BETA-2-Adrenérgico.	
72 MTHFR	Metilentetrahidrofolato reductasa	
73 MCP1.	Proteína quimiocítica de monocitos 1	A1288C
74 FASLG	Ligando Fas	menos 843C/T
75 BD5	Enfermedad inflamatoria del intestino 5	locusIGR2060a 1
76 CD16a	Receptor de fragmento Fc de IgG, baja afinidad IIIa	V158FG/T
77 IL1B	Interleucina 1 beta	menos 511a/c
78 IL4	Interleucina 4	34c/t
79 CASP9	Caspasa 9	C93T
80 ICAM1	Molécula de adhesión intercelular 1	K469E
81 MBL	Unión a lectina, manosa	Gly54Asp
82 IL4	Interleucina 4	-580
83 TPMT5	TIOPURINA S-METILTRANSFERASA 5	T145C Leu49Ser
84 TPMT6	TIOPURINA S-METILTRANSFERASA 6	A539T
85 TPMT7	TIOPURINA S-METILTRANSFERASA 7	T681G His227Gln
86 TPMT8	TIOPURINA S-METILTRANSFERASA 8	G644AArg215His
87 HLA	Antígeno de histocompatibilidad	polimorfismo b27
88 SLC19A1	Folato reductasa de la familia de transportadores de soluto 19	G80A
89 NAT2	N-Acetiltransferasa 2	282
90 NAT2	N-Acetiltransferasa 2	590

Figura 2: Tabla 2

1.- Polimorfismo C1672T de la familia de transportadores de soluto 22 A4 (SLC22A4) (sondas para detectar el polimorfismo C1672T del gen de la *familia de transportadores de soluto 22 A4*).

AAAAAAGGGTGAGGATTCCAATCAG

ÀAAAAAGGGTGAAGATTCCAATCAG

AAAAAAGGGTGAGGATTCCAATCA

AAAAAAGGGTGAAGATTCCAATCA

2.- Polimorfismo C24658G del factor de transcripción relacionado con Runt 1 (RUNX1)

TAAGAGCTGTCCCTGGGGCAGAT

TAAGAGCTGTCGCTGGGGCAGAT

ATCTGCCCCAGGGACAGCTCTTA

ATCTGCCCCAGCGACAGCTCTTA

3.- Polimorfismo G117A de peptidilarginina desiminasa tipo IV (PADI 4)

CAGAAATCTCCCTGTGCGCAGAC

CAGAAATCTCCTTGTGCGCAGAC

GTCTGCGCACAGGGAGATTTCTG

GTCTGCGCACAAGGAGATTTCTG

4.- Polimorfismo G55S de peptidilarginina desiminasa tipo IV (PADI 4)

ATATTGCCACGGCCCTCCAGCC

ATATTGCCACAGCCCTCCAGCC

GGCTGGAGGGCCGTGGGCAATAT

GGCTGGAGGGCTGTGGGCAATAT

5.- Polimorfismo V82A de peptidilarginina desiminasa tipo IV (PADI 4)

GACGATGAAAGTGGCCAGTGGTA

GACGATGAAAGCGGCCAGTGGTA

TACCACTGGCCACTTTCATCGTC

TACCACTGGCCGCTTTCATCGTC

6.- Polimorfismo G112A de peptidilarginina desiminasa tipo IV (PADI 4)

CTACCTCACCGCGGTGGGTAAGT

CTACCTCACCGGGGTGGGTAAGT

ACTTACCCACCGCGGTGAGGTAG

ACTTACCCACCCCGGTGAGGTAG

7.- Polimorfismo Hla25 del gen relacionado con cadena A del complejo mayor de histocompatibilidad clase I (MICA)

AGCGGGAGCAGAGGGACGTTTCC

AGCGGGAGCAGGGGGACGTTTCC

GGAAACGTCCCTCTGCTCCCGCT

GGAAACGTCCCCCTGCTCCCGCT

8.- Polimorfismo Hla28 del gen relacionado con cadena A del complejo mayor de histocompatibilidad clase I (MICA)

GCCCTGCCTGTGTAGCCCTTTGG

GCCCTGCCTGTTTAGCCCTTTGG

CCAAAGGGCTACACAGGCAGGGC

CCAAAGGGCTAAACAGGCAGGGC

9.- Polimorfismo Hla58 T/G de NOTCH4

CACACAGACAAATCAGTTCTTGA

CACACAGACAACTCAGTTCTTGA

TCAAGAACTGATTTGTCTGTGTG

TCAAGAACTGAGTTGTCTGTGTG

10.- Polimorfismo G1154A del factor de crecimiento endotelial vascular (VEGF)

CCGCGTGTGGAGGGGCTGAGGCT

AGCCTCAGCCCCTCCACACGCGG

CCGCGTGTGGAAGGGCTGAGGCT

AGCCTCAGCCCCTCCACACGCGG

11.- Polimorfismo C634G del factor de crecimiento endotelial vascular (VEGF).

AGCAGCGAAAGGGACAGGGGCAA

AGCAGCGAAAGCGACAGGGGCAA

TTGCCCTGTCCCTTTCGCTGCT

TTGCCCTGTGCTTTCGCTGCT

12.- Polimorfismo C936T del factor de crecimiento endotelial vascular (VEGF).

GTGACCCAGCACGGTCCCTCTTG

GTGACCCAGCATGGTCCCTCTTG

CAAGAGGGACCGTGCTGGGTCAC

CAAGAGGGACCATGCTGGGTCAC

13.- Polimorfismo A607C de interleucina 18 (IL18) .

AAAAATTATTACATAAAATTCTA

AAAAATTATTAATAAAATTCTA

TAGAATTTTATGTAATAATTTT

TAGAATTTTATTTAATAATTTT

14.- Polimorfismo G-137C de interleucina 18 (IL18) .

ACGGAAGAAAACATTTTCATGAAA

ACGGAAGAAAAGATTTTCATGAAA

TTTCATGAAATGTTTTCTCCGT

TTTCATGAAATCTTTTCTCCGT

15.- Polimorfismo Q551R del receptor de interleucina 4 (IL-4R).

CAGTGGCTATCGGGAGTTGTAC

CAGTGGCTATCAGGAGTTGTAC

GTACAAACTCCCGATAGCCACTG

GTACAAACTCCTGATAGCCACTG

16.- Polimorfismo T2018C del alelo 2 del antagonista del receptor de IL-1 (IL1RN*2).

CCAAGTAGTTGCTGGATACTTGCAA

CCAAGTAGTTGCCGGATACTTGCAA

CAACTAGTTGCTGGATACTTGCA

CAACTAGTTGCCGGATACTTGCA

17.- Polimorfismo C2073T del alelo 2 del antagonista del receptor de IL-1 (IL1RN*2)

GCCAGGAAAGCCAATGTATGT

GCCAGGAAAGTCAATGTATGT

GCCAGGAAAGCCAATGTATGT

GCCAGGAAAGTCAATGTATGT

18.- Polimorfismo C872T de muerte programada 1 (PD-1, PDCD1).

AGGGGCTCAGCCGACGGCCCTCG

AGGGGCTCAGCTGACGGCCCTCG

CGAGGGCCGTCGGCTGAGCCCT

CGAGGGCCGTCAGCTGAGCCCT

19.- Polimorfismo Asp299Gly del receptor de tipo peaje 4 (TLR4).

AAGTCAATAATATCATCGAGGTAGT

AAGTCAATAATACCATCGAGGTAGT

AGTCAATAATATCATCGAGGTAG

AGTCAATAATACCATCGAGGTAG

20. Polimorfismo Apos. VS7+283 (b>B) del receptor de vitamina D (1,25-dihidroxitamina D3).

GACAGGCCTGCGCATTCCAATA

GACAGGCCTGCACATTCCAATA

TATTGGGAATGCGCAGGCCTGTC

TATTGGGAATGTGCAGGCCTGTC

21.- Polimorfismo C-547T del dominio de inmunoglobulina de célula T y dominio de mucina 3 (TIM3).

CTGGGAGTTGCGATGGTCTGTAA

CTGGGAGTTGCTATGGTCTGTAA

TTACAGACCATCGCAACTCCCAG

TTACAGACCATAGCAACTCCCAG

22.-Polimorfismo T4259G del dominio de inmunoglobulina de célula T y dominio de mucina 3 (TIM3)

TGCACCGACTCTGCAGAGAGACT

TGCACCGACTCGGCAGAGAGACT

AGTCTCTCTGCAGAGTCGGTGCA

AGTCTCTCTGCCGAGTCGGTGCA

23.- Polimorfismo (-A2849G) de UTR 3' de uroquinasa

ACCCACCAGGGCGAACGACAATA

ACCCACCAGGGTGAACGACAATA

TATTGTCGTTCCGCCCTGGTGGGT

TATTGTCGTTCCACCCTGGTGGGT

24.-Antígeno de linfocito T citotóxico 4 (CTLA4A/G) en exón 1.

TGAACCTGGCTGCCAGGACCTGG

TGAACCTGGCTACCAGGACCTGG

CCAGGTCCTGGCAGCCAGGTTCA

CCAGGTCCTGGTAGCCAGGTTCA

25.- Polimorfismo Ala565Thr del transportador asociado con procesamiento de antígenos 2 (TAP2)

GGAACAACATTGCTTATGGGCTG

CAGCCCATAAGCAATGTTGTCC

GGAACAACATTACTTATGGGCTG

CAGCCCATAAGTAATGTTGTCC

26.-Polimorfismo I kappa BL A-62T

AGCAGAGACGCAGGTGGAGGACG

AGCAGAGACGCTGGTGGAGGACG

CGTCCTCCACCTGCGTCTCTGCT

CGTCCTCCACCAGCGTCTCTGCT

27- Polimorfismo C1362T de poli (ADP-ribosa) polimerasa 1 (PARP-1). Poli (ADP-ribosa) polimerasa 1 (PARP-1) haplotipo A (410T-[A](10)-[CA](10-12)-1362C, que incluye alelos CA de PARP-1 cortos) y haplotipo B (410C-[A](11)-[CA](13-20)-1362T.

AGGCGGGGGCCCGCGCTTCCCGG

AGGCGGGGGCCTGCGCTTCCCGG

CCGGGAAGCGCGGGCCCCCGCCT

CCGGGAAGCGCAGGCCCCCGCCT

28. Polimorfismo T1858C de proteína tirosina fosfatasa.

CACTTCTGTACGGACACCTGAA

CACTTCTGTATGGACACCTGAA

TTCAGGTGTCCGTACAGGAAGTG

TTCAGGTGTCCATACAGGAAGTG

29.- Polimorfismo Fok I de receptor de vitamina D.

TCTTACAGGGATGGAGGCAATGG

TCTTACAGGGACGGAGGCAATGG

CCATTGCCTCCATCCCTGTAAGA

CCATTGCCTCCGTCCCTGTAAGA

30.- Polimorfismo G38A de uteroglobina (UG).

GAACCAGAGACGGGCCAGAGCAT

GAACCAGAGACAGGCCAGAGCAT

ATGCTCTGGCCCGTCTCTGGTTC

ATGCTCTGGCCTGTCTCTGGTTC

31. Polimorfismo 6A/6A de metaloproteínasa de la matriz 3 (MMP3).

ACATGGTTTTTTCCCCCATCAA

ACATGGTTTTTTCCCCCATCAA

TTGATGGGGGGAAAAAACCATGT

TTGATGGGGGGAAAAAACCATGT

32.- Genotipos Arg/Arg y Pro/Pro en el codón 72 de la proteína tumoral p53

CAGGGGCCACGGGGGAGCAGCC

CAGGGGCCACGGGGGAGCAGCC

GGCTGCTCCCCGCGTGGCCCCTG

GGCTGCTCCCCCGTGGCCCCTG

33.- Polimorfismo (-A2849G) de interleucina 10 (IL-10).

ACCTTATGATCTGCCCGCCTTGG

ACCTTATGATCCGCCCGCCTTGG

CCAAGGCGGGCAGATCATAAGGT

CCAAGGCGGGCGGATCATAAGGT

34.- El -308 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa).

ACCCCGTCCCCATGCCCT

ACCCCGTCCTCATGCCCT

ACCCCGTCCCCATGCCCT

ACCCCGTCCTCATGCCCT

35.- El +488 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa).

GAGAAAAAACATGGAGAAAGAC

GAGAAAAAACGTGGAGAAAGAC

AGGTCCTGGCAGCCAGGTT

AGGTCCTGGTAGCCAGGTT

36.- El -857 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa).

GAACCAGAGACGGGCCAGAGCAT

GAACCAGAGACAGGCCAGAGCAT

ATGCTCTGGCCCGTCTCTGGTTC

ATGCTCTGGCCTGTCTCTGGTTC

37. El -238 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa).

CTCCCTGCTCCGATTCCGAGG

CTCCCTGCTCTGATTCCGAGG

TCCCTGCTCCGATTCCGAG

TCCCTGCTCTGATTCCGAG

38. Polimorfismo T587G del receptor del factor de necrosis tumoral 2 (TNFR2).

GAATGCAAGCATGGATGCAGTCT

GAATGCAAGCAGGGATGCAGTCT

GACTGCATCCATGCTTGCATT

GACTGCATCCCTGCTTGCATT

39. Polimorfismo +36 A/A del receptor del factor de necrosis tumoral 1 (TNFR1).

CTGCTGCTGCCGCTGGTGAGACC

CTGCTGCTGCCACTGGTGAGACC

GGTTCACCAGCGGCAGCAGCAG

GGTTCACCAGTGGCAGCAGCAG

40. Polimorfismo G-455A del polipéptido beta de fibrinógeno B (FGB).

TTGATTTTAATGGCCCCTTTGA

TTGATTTTAATAGCCCCTTTGA

TCAAAGGGGCCATTAAAATCAA

TCAAAGGGGCTATTAAAATCAA

41. Polimorfismo de genotipo nulo GSTM1 de la variante 1 del transcrito de glutatión S-transferasa M1.

ACACATTCTTGGCCTTCTGCAGA

ACACATTCTTGCCTTCTGCAGA

TCTGCAGAAGGCCAAGAATGTGT

TCTGCAGAAGGTCAAGAATGTGT

42. Polimorfismo del genotipo GSTT1-0 de glutatión S-transferasa T1 (GSTT1)

TGCCTAGTGGGTTACCTGCCCA

TGCCTAGTGGGGTACCTGCCCA

TGGGCAGGTGAACCCACTAGGCA

TGGGCAGGTGACCCCACTAGGCA

43. Polimorfismo (-C3435T) de proteína de multi-resistencia a fármacos (MDR-1).

AGGAAGAGATCGTGAGGGCAG

AGGAAGAGATTGTGAGGGCAG

AGGAAGAGATCGTGAGGGCAG

AGGAAGAGATTGTGAGGGCAG

44. Polimorfismo R241G de molécula de adhesión 1 (ICAM1).

TTCCTGGACGGGCTGTCCC

TTCCTGGACAGGCTGTCCC

GGGAACAGCCCGTCCAGGGAA

GGGAACAGCCTGTCCAGGGAA

45. Polimorfismo C3954T de interleucina 1 beta (IL-1B).

ACCTATCTTCTTCGACACATGGGAT

ACCTATCTCTTTGACACATGGGAT

ATCCCATGTGTCAAGAAGATAGGT

ATCCCATGTGTCAAAGAAGATAGGT

46. Polimorfismo Val58Ile de condromodulina II (Chm II).

AGGGTGTGGACGTCTTGTGCTCT

AGGGTGTGGACATCTTGTGCTCT

AGAGCACAAGACGTCCACACCCT

AGAGCACAAGATGTCCACACCCT

47. Polimorfismo - 622 del gen de interleucina 6 (IL-6).

TGGGATTTCCGATGCTAAAGGA

TGGGATTTCCCATGCTAAAGGA

TCCTTTAGCATCGGAAAATCCCA

TCCTTTAGCATGGGAAAATCCCA

48. 347 GG de 5-aminoimidazol-4-carboxamida ribonucleótido formiltransferasa/IMP ciclohidrolasa (ATIC).

TCCAGGTGTAAGTGTGAGGAGG

TCCAGGTGTAAGTGTGAGGAGG

CCTCCTCAACAGTTACACCTGGA

CCTCCTCAACACTTACACCTGGA

49. Polimorfismo G80A del miembro 4 de la familia de transportadores de soluto 22 (SLC22A4).

AAAAAAGGGTGAGGATTCCAATCAG

AAAAAAGGGTGAAGATTCCAATCAG

AAAAAAGGGTGAGGATTCCAATCA

AAAAAAGGGTGAAGATTCCAATCA

50. Polimorfismo C677T de metilentetrahidrofolato reductasa (MTHFR).

GATGAAATCGGCTCCCGCAGA

GATGAAATCGACTCCCGCAGA

GATGAAATCGGCTCCCGCAGA

GATGAAATCGACTCCCGCAGA

51. Polimorfismo 249 del gen de linfotoxina alfa (LTA).

ACAGAGAGGAATCATGGCAGAAA

ACAGAGAGGAACCATGGCAGAAA

TTCTGCCATGATTCCCTCTCTG

TTCTGCCATGGTTCCCTCTCTG

52. Polimorfismo 365 del gen de linfotoxina alfa (LTA)

CGGGGGTGCTCTCTCCAGGGC

CGGGGGTGCTGTCTCCAGGGC

GCCCTGGGAGAGACACCCCCG

GCCCTGGGAGACAGACACCCCCG

53. Polimorfismo 720 del gen de linfotoxina alfa (LTA).

CAGGTTTGAGGGTGCTGTGGGCA

CAGGTTTGAGGTTGCTGTGGGCA

GGTTTGAGGGTGCTGTGGG

GGTTTGAGGTTGCTGTGGG

54. Polimorfismo -1087 G/G de interleucina 10 (IL-10)

CTTCTTTGGGAAGGGGAAGTAGG

CTTCTTTGGGAGGGGAAGTAGG

CCTACTTCCCCTAGGGTTTCTTC

CCTACTTCCCCCAGGGTTTCTTC

55. Polimorfismo -308 TNF1/TNF1 del factor de necrosis tumoral alfa (TNF).

ACCCCGTCCCCATGCCCT

ACCCCGTCCTCATGCCCT

ACCCCGTCCCCATGCCCT

ACCCCGTCCTCATGCCCT

56. Polimorfismo de factor de crecimiento transformante TGFbeta1.

GACGCCTGGCCGGCCGGCCGCGG

GACGCCTGGCCCGCCGGCCGCGG

CCGCGGCCGGCCGGCCAGGCGTC

CCGCGGCCGGCCGGCCAGGCGTC

57. Polimorfismo -1087 AA del gen de interleucina 10 (IL-10).

CTTCTTTGGGAGGGGAAGTAGG

CTTCTTTGGGAAGGGGAAGTAGG

CCTACTTCCCCCAGGGTTTCTTC

CCTACTTCCCCTAGGGTTTCTTC

58. Polimorfismo G460A de tiopurina metiltransferasa (TPMT).

TGGGATAGAGGAGCATTAGTTGCCA

TGGGATAGAGGAACATTAGTTGCCA

GGGATAGAGGAGCATTAGTTGCC

GGGATAGAGGAACATTAGTTGCC

59. Polimorfismo A719G de tiopurina metiltransferasa (TPMT).

TCTGTAAGTAGATATAACTTTTCAA

TCTGTAAGTAGACATAACTTTTCAA

CTGTAAGTAGATATAACTTTTCA

CTGTAAGTAGACATAACTTTTCA

60. Polimorfismo -C196T de proteína de tipo receptor de Fc 3 (FCRL3)

ACGGGAAGTCCTTGATCTGTACA

ACGGGAAGTCCTTGATCTGTACA

TGTACAGATCAAGGACTTCCCGT

TGTACAGATCAGGGACTTCCCGT

61. Polimorfismo Arg260Thr del gen de tapasina (TPSN).

CTGGCTGCCTAGAGTTCAACCCT

CTGGCTGCCTACAGTTCAACCCT

AGGGTTGAACTCTAGGCAGCCAG

AGGGTTGAACTGTAGGCAGCCAG

62. Polimorfismo 341T>C de N-acetiltransferasa 2 (NAT2).

GCAGGTGACCATTGACGGCAGGA

TCCTGCCGTCAATGGTCACCTGC

GCAGGTGACCACTGACGGCAGGA

TCCTGCCGTCAATGGTCACCTGC

63. Polimorfismo 481C>T de N-acetiltransferasa 2 (NAT2).

GAATCTGGTACCTGGACCAAATC

GATTTGGTCCAGGTACCAGATTC

GAATCTGGTACTGGACCAAATC

GATTTGGTCCAAGTACCAGATTC

64. Polimorfismo 803A>G de N-acetiltransferasa 2 (NAT2).

AGAAGTGCTGAAAAATATATTTA

TAAATATATTTTCAGCACTTCT

AGAAGTGCTGAGAAATATATTTA

TAAATATATTTCTCAGCACTTCT

65. Polimorfismo 857G> de N-acetiltransferasa 2 (NAT2).

ACCTGGTGATGGATCCCTTACTA

TAGTAAGGGATCCATCACCAGGT

ACCTGGTGATGAATCCCTTACTA

TAGTAAGGGATTCATCACCAGGT

66. Polimorfismo 191G>A de N-acetiltransferasa 2 (NAT2).

AAGAAGAAACCGGGGTGGGTGGT

ACCACCCACCCCGGTTTCTTCTT

AAGAAGAAACCGGGGTGGGTGGT

ACCACCCACCCTGGTTTCTTCTT

67 Polimorfismo Hla 78 del antígeno de histocompatibilidad.

TTCCCCCAGCGTGGCAGGATCT

TTCCCCCAGCATGGCAGGATCT

AGATCCTGCCACGCTGGGGGGAA

AGATCCTGCCATGCTGGGGGGAA

68. Polimorfismo Hla 76 del antígeno de histocompatibilidad.

TCAAGAACTGATTTGTCTGTGTG

TCAAGAACTGAGTTGTCTGTGTG

CACACAGACAAATCAGTTCTTGA

CACACAGACAACTCAGTTCTTGA

69 Polimorfismo de receptor de melatonina 1B (MTNR1B).

TTCCCCCAGCATGGCAGGATCT

TTCCCCCAGCGTGGCAGGATCT

AGATCCTGCCA TGCTGGGGGGAA

AGATCCTGCCACGCTGGGGGGAA

70. Polimorfismo de transactivador del complejo mayor de histocompatibilidad clase II (MHC2TA).

TTCAGAGGTGTAGGGAGGGCTTA

TTCAGAGGTGTGGGGAGGGCTTA

TAAGCCCTCCCTACACCTCTGAA

TAAGCCCTCCCCACACCTCTGAA

71. Polimorfismo de receptor BETA-2-adrenérgico.

TGGCACCCAATGGAAGCCATGCG

TGGCACCCAATAGAAGCCATGCG

CGCATGGCTTCCATTGGGTGCCA

CGCATGGCTTCTATTGGGTGCCA

72. Polimorfismo A1298C de metilentetrahidrofolato reductasa (MTHFR).

GACCAGTGAAGAAAGTGCTTTG

GACCAGTGAAGCAAGTGCTTTG

CAAAGACACTTCTTCACTGGTC

CAAAGACACTTGCTTCACTGGTC

73. Proteína quimiotáctica de monocitos 1 (MCP1).

GGCAGACAGCTGTCACTTTCCAG

GGCAGACAGCTATCACTTTCCAG

CTGGAAAGTGACAGCTGTCTGCC

CTGGAAAGTGATAGCTGTCTGCC

74. Polimorfismo menos 843C/T de LIGANDO FAS.

TGAAAACATTGCGAAATACAAAG

TGAAAACATTGTGAAATACAAAG

TTTGTATTTGCAATGTTTTTTC

TTTGTATTTTACAATGTTTTTTC

75. Polimorfismo locus IGR2060a_1 de IBD5.

CCTTGCAACCCTGGCAAAGGTAATG

CCTTGCAACCCTCGCAAAGGTAATG

CATTACCTTTGCCAGGGTTGCAAGG

CATTACCTTTGCGAGGGTTGCAAGG

76. Polimorfismo V158F (G/T) de cd16a.

TGCAGGGGGCTTGTGGGAGTAAAA

TGCAGGGGGCTTTTTGGGAGTAAAA

TTACTCCCAACAAGCCCCCTG

TTACTCCCAAAAAGCCCCCTG

77. Polimorfismo menos 511a/c de interleucina 1 beta.

AAGAGAATCCCAGAGCAGCCTGT

AAGAGAATCCCCGAGCAGCCTGT

ACAGGCTGCTCTGGGATTCTCTT

ACAGGCTGCTCGGGGATTCTCTT

78. Polimorfismo 34c/t de interleucina 4 (IL4)

TAAACTAATTGCCTCACATTGTC

TAAACTAATTGTCTCACATTGTC

GACAATGTGAGGCAATTAGTTTA

GACAATGTGAGACAATTAGTTTA

79. Polimorfismo C93T de caspasa 9.

GTCCTGCTGAGCCGCGAGCTGTT

GTCCTGCTGAGTCGCGAGCTGTT

ACAGCTCGCGGCTCAGCAGGA

ACAGCTCGCGACTCAGCAGGA

80. Polimorfismo K469E de ICAM1.

AGGTCACCCGCAAGGTGACCGTG

CACGGTCACCTTGCGGGTGACCT

AGGTCACCCGCGAGGTGACCGTG

CACGGTCACCTCGCGGGTGACCT

81. Polimorfismo Gly54Asp de UNIÓN A LECTINA, MANOSA (MBL).

TGGGCGTGATGACACCAAGGGAG

CTCCCTTGGTGTTCATCACGCCCA

TGGGCGTGATGGCACCAAGGGAG

CTCCCTTGGTGCCATCACGCCCA

82. Polimorfismo -590 de interleucina 4.

GAGAACATTGTCCCCCAGTGCTG

CAGCACTGGGGGACAATGTTCTC

GAGAACATTGTTCCCCCAGTGCTG

CAGCACTGGGGAACAATGTTCTC

83. Polimorfismo T145C (Leu49Ser) de tiopurina S-metiltransferasa 5 TPMT5.

CTTCAGGCTATTAAAGAAGCATT

CTTCAGGCTATCAAAGAAGCATT

AATGCTTCTTTAATAGCCTGAAG

AATGCTTCTTTGATAGCCTGAAG

84. Polimorfismo A539T de tiopurina S-metiltransferasa 6 TPMT6.

GAAGTTTCAGTATCTCCTGTGTG

GAAGTTTCAGTTTCTCCTGTGTG

CACACAGGAGAT ACTGAAACTTC

CACACAGGAGAAACTGAAACTTC

85. Polimorfismo T681G (His227Gln) de tiopurina S-metiltransferasa 7 TPMT7.

GAAGAACGACATAAAAGTTGGGG

GAAGAACGACAGAAAAGTTGGGG

CCCCAACTTTT ATGTCGTTCTTC

CCCCAACTTTT CTGTCGTTCTTC

86. Polimorfismo G644A (Arg215His) de tiopurina S-metiltransferasa 8 TPMT 8.

ATGCAATATACGTTGTCTTGAGA

ATGCAATATACATTGTCTTGAGA

TCTCAAGACAACGTATATTGCAT

TCTCAAGACAATGTATATTGCAT

87. Polimorfismo HLA b27 del antígeno de histocompatibilidad.

Gacagcgacgccgagtcgag

Ctcggactcgcggtcgctgtc

Gacagcgacgccgagtcgag

Ctcggactcgcggtcgctgtc

88. G80A de familia de transportadores de soluto 19 (folato reductasa) SLC19A1.

GTCTGGCGGCGCCTCGTGTGCT

GTCTGGCGGCACCTCGTGTGCT

AGCACACGAGGCGCCGAGGAC

AGCACACGAGGTGCCGAGGAC

89. Polimorfismo 282 de N-acetiltransferasa 2 (NAT2).

GGTATTTTTACATCCCTCCAGT

GGTATTTTTATATCCCTCCAGT

GGTATTTTTACATCCCTCCAG

GGTATTTTTATATCCCTCCAG

90. Polimorfismo 590 de N-acetiltransferasa 2 (NAT2).

CTTCAATTGTTGAGGTTCAAGC

CTTCAATTGTTGAGGTTCAAGC

GCTTGAACCTCGAACAATTGAAG

GCTTGAACCTCAAACAATTGAAG

Figura 3: Tabla 3

Polimorfismo C1672T de la familia de transportadores de soluto 22 A4 (SLC22A4) (sondas para detectar el polimorfismo C1672T del gen de la familia de transportadores de soluto 22 A4)

GCTAGATGAAGAGCAAGCGCCAAGAGTGCCCAGAGAGTCC
TACAACCGACAGATGTATGTTTCTCCCTAAGGCATTTTGGT

Polimorfismo C24658G del factor de transcripción relacionado con Runt 1 (RUNX1)

GCTAGATGAAGAGCAAGCGCTCTCCACTGTGCTGATGAGG
TACAACCGACAGATGTATGTTTTTCTTGCGTGCTGACAC

Polimorfismo G117A de peptidilarginina desiminasa tipo IV (PADI 4)

GCTAGATGAAGAGCAAGCGCACCCCTCACCAACCTCTCCTC
TACAACCGACAGATGTATGTACACGGAATACGTGGGACAG

Polimorfismo G55S de peptidilarginina desiminasa tipo IV (PADI 4)

GCTAGATGAAGAGCAAGCGCCCCATGTGTCTTGTCCACAG
TACAACCGACAGATGTATGTTTCGTCAGGGTCACCTCTACC

Polimorfismo V82A de peptidilarginina desiminasa tipo IV (PADI 4)

GCTAGATGAAGAGCAAGCGC CCCATGTGTCTTGTCCACAG
TACAACCGACAGATGTATGTCCCCACAGCTATGACACTCA

Polimorfismo G112A de peptidilarginina desiminasa tipo IV (PADI 4)

GCTAGATGAAGAGCAAGCGCTCCAGTGGGTGTTTGTGA
TACAACCGACAGATGTATGTGGATGAGACGGCÀCTCTAGG

Polimorfismo Hla25 del gen relacionado con cadena A del complejo mayor de histocompatibilidad clase I (MICA)

GCTAGATGAAGAGCAAGCGCCATCTTCCCTTTTGCACCT
TACAACCGACAGATGTATGTGATCCCGAGGAGGACTGAA

Polimorfismo Hla28 del gen relacionado con cadena A del complejo mayor de histocompatibilidad clase I (MICA)

GCTAGATGAAGAGCAAGCGCATGCTAAGGGCCTGGATGAT

TACAACCGACAGATGTATGTCAGGACCCTCCCTGTTACAA

Polimorfismo H1a58 T/G de NOTCH4

GCTAGATGAAGAGCAAGCGCGGCCAAAACAACCATCTGAG

TACAACCGACAGATGTATGTTCCCTCATGGTTGGGTTAAA

Polimorfismo G1154A del factor de crecimiento endotelial vascular (VEGF)

GCTAGATGAAGAGCAAGCGCTTTTCAGGCTGTGAACCTTG

TACAACCGACAGATGTATGTGAGATCCTCCCCGCTACCAG

Polimorfismo C634G del factor de crecimiento endotelial vascular (VEGF).

GCTAGATGAAGAGCAAGCGCGGATTTTGAAACCAGCAGA

TACAACCGACAGATGTATGTCTGTCTGTCTGTCCGTCAGC

Polimorfismo C936T del factor de crecimiento endotelial vascular (VEGA).

GCTAGATGAAGAGCAAGCGCAGGGTTTCGGGAACCAGATC

TACAACCGACAGATGTATGTCTCGGTGATTTAGCAGCAAG

Polimorfismo A607C de interleucina 18 (IL18) .

GCTAGATGAAGAGCAAGCGCTCCCTCTCCCCAAGCTTACT

TACAACCGACAGATGTATGTTTCAGGACTTCCCCTTCCTC

Polimorfismo G-137C de interleucina 18 (IL18) .

GCTAGATGAAGAGCAAGCGCGTGCTGAAGTGTGACCAGGA

TACAACCGACAGATGTATGTGGGCAATGGAAGTCGAAATA

Polimorfismo Q551R del receptor de interleucina 4 (IL-4R).

GCTAGATGAAGAGCAAGCGCCAACCTGAGCCAGAAACCTG

TACAACCGACAGATGTATGTCCACATTTCTCTGGGGACAC

Polimorfismo T2018C del alelo 2 del antagonista del receptor de IL-1 (IL1RN*2).

GCTAGATGAAGAGCAAGCGCACAAAGTTCTGGGGGACACAG

TACAACCGACAGATGTATGTATTGCACCTAGGGTTTGTGC

Polimorfismo C872T de muerte programada 1 (PD-1, PDCD1)

GCTAGATGAAGAGCAAGCGCGTGCCTGTGTTCTCTGTGGA

TACAACCGACAGATGTATGTCCAAGAGCAGTGTCCATCCT

Polimorfismo C2073T del alelo 2 del antagonista del receptor de IL-1 (IL1RN*2)

GCTAGATGAAGAGCAAGCGCCTCTAGAGGGCCTGTGCAAT

TACAACCGACAGATGTATGTTCAATGTGGGAAACTGTCCA

Polimorfismo Apos. VS7+283 (b>B) del receptor de vitamina D (1,25-dihidroxitamina D3).

GCTAGATGAAGAGCAAGCGCCCTCACTGCCCTTAGCTCTG

TACAACCGACAGATGTATGTCCCGCAAGAAACCTCAAATA

Polimorfismo C-547T del dominio de inmunoglobulina de célula T y dominio de mucina 3 (TIM3).

GCTAGATGAAGAGCAAGCGCTGAGGCTTATGCTGGGAGTT

TACAACCGACAGATGTATGTAAATGGAGCATGTCCGAGTC

Polimorfismo T4259G del dominio de inmunoglobulina de célula T y dominio de mucina 3 (TIM3)

GCTAGATGAAGAGCAAGCGCCCCCTTTGATTCCCTGAAATA

TACAACCGACAGATGTATGTGCCCCATGCATAGTTACCTG

Polimorfismo (-A2849G) de UTR 3' de uroquinasa

GCTAGATGAAGAGCAAGCGCAGGCTCTGCACAGATGGATT

TACAACCGACAGATGTATGTATGCCCTGCCCTTTTTAACT

Antígeno de linfocito T citotóxico 4 (CTLA4A/G) en exón 1.

GCTAGATGAAGAGCAAGCGCCTGAACACCGCTCCCATAAA

TACAACCGACAGATGTATGTCTCCTCCATCTTCATGCTC

Polimorfismo Ala565Thr del transportador asociado con procesamiento de antígenos 2 (TAP2)

GCTAGATGAAGAGCAAGCGC GCACCAGGTGTTTCATTCTGA

TACAACCGACAGATGTATGT CCGCCATCACCTTATCATCT

Polimorfismo I kappa BL A-62T

GCTAGATGAAGAGCAAGCGCCCCAAGGCCTGTGTTTAAGA

TACAACCGACAGATGTATGTGCCTCCAGTCCAGTCTTCTG

Polimorfismo C1362T de poli (ADP-ribosa) polimerasa 1 (PARP-1). Poli (ADP-ribosa) polimerasa 1 (PARP-1) haplotipo A (410T-[A](10)-[CA](10-12)-1362C, que incluye alelos CA de PARP-1 cortos) y haplotipo B (410C-[A](11)-[CA](13-20)-1362T.

GCTAGATGAAGAGCAAGCGCAGGAGGGTGGACCTAGCACT

TACAACCGACAGATGTATGTCAAGTGATCTTCACGCCTCA

Polimorfismo T1858C de proteína tirosina fosfatasa.

GCTAGATGAAGAGCAAGCGCGGATAGCAACTGCTCCAAGG

TACAACCGACAGATGTATGTTTACACATCAGCTTCCCAA

Polimorfismo Fok I de receptor de vitamina D

GCTAGATGAAGAGCAAGCGCGGCTGCTTGCTGTTCTTAC

TACAACCGACAGATGTATGTTGCTTCTTCTCCCTCCCTTT

Polimorfismo G38A de uteroglobina (UG).

GCTAGATGAAGAGCAAGCGCCTTGCTGGGCATGTCTCATA

TACAACCGACAGATGTATGTGGGGTCTGAGAGTTCCTA

Polimorfismo 6A/6A de metaloproteínasa de la matriz 3 (MMP3).

GCTAGATGAAGAGCAAGCGCTCACTGCCACCACTCTGTTC

TACAACCGACAGATGTATGTGCCTCAACCTCTCAAAGTGC

Genotipos Arg/Arg y Pro/Pro en el codón 72 de la proteína tumoral p53

GCTAGATGAAGAGCAAGCGCCTGCCCTGGTAGGTTTTCTG

TACAACCGACAGATGTATGTGAAGACCCAGGTCCAGATG

Polimorfismo (-A2849G) de interleucina 10 (IL-10).

GCTAGATGAAGAGCAAGCGCTCTCAGCTCACTGCAAGCTC

TACAACCGACAGATGTATGTACAGTGGCTCATGCCTGTAA

Polimorfismo -308 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa)

GCTAGATGAAGAGCAAGCGCAAAGATGTGCGCTGATAGGG

TACAACCGACAGATGTATGTTTCCCCATCTCTTCTCA

Polimorfismo -238 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa)

GCTAGATGAAGAGCAAGCGCAAAGATGTGCGCTGATAGGG

TACAACCGACAGATGTATGTTTTCCCCATCTCTCTTCTCA

Polimorfismo +488 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa)

GCTAGATGAAGAGCAAGCGCACCTGGTCCCCAAAAGAAAT

TACAACCGACAGATGTATGTAAAGTTGGGGACACACAAGC

Polimorfismo -857 del promotor del gen del factor de necrosis tumoral alfa (TNF alfa)

GCTAGATGAAGAGCAAGCGCTGTCCAGGGCTATGGAAGTC

TACAACCGACAGATGTATGTTTTTCATTCTGACCCGGAGAC

Polimorfismo T587G del receptor del factor de necrosis tumoral 2 (TNFR2)

GCTAGATGAAGAGCAAGCGCCTCTCCTATCCTGCCTGCTG

TACAACCGACAGATGTATGTGGCTGGGGTAAGTGTACTGC

Polimorfismo +36 A/A del receptor del factor de necrosis tumoral 1 (TNFR1)

CTGCTGCTGCCGCTGGTGAGACC GGTCTCACCAGCGGCAGCAGCAG

CTGCTGCTGCCACTGGTGAGACC GGTCTCACCAGTGGCAGCAGCAG

Polimorfismo G-455A del polipéptido beta de fibrinógeno B (FGB).

GCTAGATGAAGAGCAAGCGCGGAGTGAGAGGCCATAGCTG

TACAACCGACAGATGTATGTGCAGTGGCTGAGGTTAGGAC

Polimorfismo de genotipo nulo GSTM1 de la variante 1 del transcrito de glutatión S-transferasa M1

GCTAGATGAAGAGCAAGCGCATGGTTTGCAGGAAACAAGG

TACAACCGACAGATGTATGTAAAGCGGGAGATGAAGTCCT

Polimorfismo del genotipo GSTT1-0 de glutatión S-transferasa T1 (GSTT1)

GCTAGATGAAGAGCAAGCGCGGCAGCATAAGCAGGACTTC

TACAACCGACAGATGTATGTGTTGCTCGAGGACAAGTTCC

Polimorfismo (-C3435T) de proteína de multi-resistencia a fármacos (MDR-1).

GCTAGATGAAGAGCAAGCGC TGCTCCCAGGCTGTTTATT

TACAACCGACAGATGTATGT TGTTTTTCAGCTGCTTGATGG

Polimorfismo R241G de molécula de adhesión 1 (ICAM1).

GCTAGATGAAGAGCAAGCGCGAATGAAATGCCCCAGAGAA
TACAACCGACAGATGTATGTACTGTGGGGTTCAACCTCTG

Polimorfismo C3954T de interleucina 1 beta (IL-1B).

GCTAGATGAAGAGCAAGCGCTGTTCTTAGCCACCCCACTC
TACAACCGACAGATGTATGTGTGATCGTACAGGTGCATCG

Polimorfismo Val58Ile de condromodulina II (Chm II).

GCTAGATGAAGAGCAAGCGC CTGCAAACAGTGGGCTAACA
TACAACCGACAGATGTATGT CCCACAATCATTCCAGTGAA

Polimorfismo -622 del gen de interleucina 6 (IL-6).

GCTAGATGAAGAGCAAGCGCTCCCCCTAGTTGTGTCTTGC
TACAACCGACAGATGTATGTTTCATGGGAAAATCCCACATT

347 GG de 5-aminoimidazol-4-carboxamida ribonucleótido formiltransferasa/IMP ciclohidrolasa (ATIC).

GCTAGATGAAGAGCAAGCGCGACAGTGGCTTCTCCAGGTG
TACAACCGACAGATGTATGTTCCCAAACACAATCCAGAA

Polimorfismo G80A del miembro 4 de la familia de transportadores de soluto 22 (SLC22A4).

GCTAGATGAAGAGCAAGCGCCAAGAGTGCCCAGAGAGTCC
TACAACCGACAGATGTATGTTTCTCCCTAAGGCATTTTGGT

Polimorfismo C677T de metilentetrahidrofolato reductasa (MTHFR).

GCTAGATGAAGAGCAAGCGCGCCTCTCCTGACTGTCATCC
TACAACCGACAGATGTATGTTTCAAAGCGGAAGAATGTG

Polimorfismo A1298C de metilentetrahidrofolato reductasa (MTHFR).

GCTAGATGAAGAGCAAGCGCGCCTCTCCTGACTGTCATCC
TACAACCGACAGATGTATGTTTCAAAGCGGAAGAATGTG

Polimorfismo 249 del gen de linfotoxina alfa (LTA).

GCTAGATGAAGAGCAAGCGCGGGTTTGGTTTTGGTTTCCT

TACAACCGACAGATGTATGTCCCGAGAGAGAGATCGACAG

Polimorfismo 365 del gen de linfotoxina alfa (LTA)

GCTAGATGAAGAGCAAGCGCTGTCTCCCTCTGCTCACCTT

TACAACCGACAGATGTATGTAGGAGGAGGTGTAGGGTGGT

Polimorfismo 720 del gen de linfotoxina alfa (LTA).

GCTAGATGAAGAGCAAGCGCAGGTGAGGCAGCAGGAGAAT

TACAACCGACAGATGTATGTATGAGGTGAGCAGCAGGTTT

Polimorfismo -1087 G/G de interleucina 10 (IL-10)

GCTAGATGAAGAGCAAGCGCTTCCCCAGGTAGAGCAACAC

TACAACCGACAGATGTATGTGATGGGGTGGAAGAAGTTGA

Polimorfismo -308 TNF1/TNF1 del factor de necrosis tumoral alfa (TNF).

GCTAGATGAAGAGCAAGCGCACCTGGTCCCCAAAAGAAAT

TACAACCGACAGATGTATGTAAAGTTGGGGACACACAAGC

Factor de crecimiento transformante beta TGF beta1 (alelo C poco frecuente en el codón 25).

GCTAGATGAAGAGCAAGCGCCGAGGCCCTCCTACCTTTT

TACAACCGACAGATGTATGTTTCGATAGTCTTGCAGGTGGA

Polimorfismo -1087 AA del gen de interleucina 10 (IL-10).

GCTAGATGAAGAGCAAGCGCTTCCCCAGGTAGAGCAACAC

TACAACCGACAGATGTATGTGATGGGGTGGAAGAAGTTGA

Polimorfismo 282C>T de N-acetiltransferasa 2 (arilamina N-acetiltransferasa) NAT2

GCTAGATGAAGAGCAAGCGCCCTGCCAAAGAAGAAACACC

TACAACCGACAGATGTATGTGATGAAGCCCACCAAACAGT

Polimorfismo 590G>A de N-acetiltransferasa 2 (arilamina N-acetiltransferasa) NAT2

GCTAGATGAAGAGCAAGCGCACTGGGCTCTGACCACAATC

TACAACCGACAGATGTATGTGGCTGATCCTTCCCAGAAAT

Polimorfismo G460A de tiopurina metiltransferasa (TPMT).

GCTAGATGAAGAGCAAGCGCCCAGGTCCACACATTCCTCT

TACAACCGACAGATGTATGTTTACCATTTGCGATCACCTG

Polimorfismo A719G de tiopurina metiltransferasa (TPMT).

GCTAGATGAAGAGCAAGCGCCCAGGTCCACACATTCCTCT

TACAACCGACAGATGTATGTTTACCATTTGCGATCACCTG

Polimorfismo -C196T de proteína de tipo receptor de Fc 3 (FCRL3)

GCTAGATGAAGAGCAAGCGCGGGGTGGAACCTCTTTGATT

TACAACCGACAGATGTATGTGCTTAATGAGCGTGGTGTGA

Polimorfismo Arg260Thr del gen de tapasina (TPSN).

GCTAGATGAAGAGCAAGCGCTGGGTAAGGGACATCTGCTC

TACAACCGACAGATGTATGTCAGATAGGTGCCCTCCTGAA

Polimorfismo Asp299Gly de receptor de tipo peaje 4 (TLR4)

GCTAGATGAAGAGCAAGCGCCTCTAGAGGGCCTGTGCAAT

TACAACCGACAGATGTATGTTCAATGTGGGAAACTGTCCA

G80A de la familia de transportadores de soluto 19 (folato reductasa) SLC19A1

GCTAGATGAAGAGCAAGCGCAGGCACAGTGTACCTTCGT

TACAACCGACAGATGTATGTGCCGTAGAAGCAAAGGTAGC

Polimorfismo Hla 76 del antígeno de histocompatibilidad

GCTAGATGAAGAGCAAGCGCGCACAATTCAACACCTCTGC

TACAACCGACAGATGTATGTCTGCAACCTTCCACTGTCCT

Polimorfismos Hla 78 del antígeno de histocompatibilidad

GCTAGATGAAGAGCAAGCGGATCTGCAGAGCCATCTTCC

TACAACCGACAGATGTATGTTGAGGTCTTCAGCTCCAGT

Polimorfismo de receptor de melatonina 1B (MTNR1B).

GCTAGATGAAGAGCAAGCGCGCAGAAGAGCCCAACTCCTT

TACAACCGACAGATGTATGTTGCATGTGTGGTTGTGATTG

Polimorfismo de transactivador del complejo mayor de histocompatibilidad clase II (MHC2TA).

GCTAGATGAAGAGCAAGCGCAAGGTTCCCCAACAGACTT
TACAACCGACAGATGTATGTCAAGCTAAGCCAACATGCAA

Polimorfismo de receptor BETA-2-adrenérgico.

GCTAGATGAAGAGCAAGCGCGCTCACCTGCCAGACTGC
GCTAGATGAAGAGCAAGCGCGCCAGGACGATGAGAGACAT

Polimorfismo (-2518) G>A de la proteína quimiotáctica de monocitos 1 (MCP1).

GCTAGATGAAGAGCAAGCGCCCAGCCAAATGCATTCTCTT
TACAACCGACAGATGTATGTCACAGGGAAGGTGAAGGGTA

Polimorfismo menos 843C/T del gen de LIGANDO FAS.

GCTAGATGAAGAGCAAGCGCCTTGAGCCCAGGAGTTTGAG
TACAACCGACAGATGTATGTATCAGAGGCTGCAAACCAGT

El polimorfismo IBD5 locus IGR2060a_1 del gen de la enfermedad inflamatoria del intestino.

GCTAGATGAAGAGCAAGCGCCATACAGCACCTTCGGGTCT
TACAACCGACAGATGTATGTGGGCAGACTTTGGAAGTCTCAG

El polimorfismo V158F (G/T) del gen cd16a

GCTAGATGAAGAGCAAGCGC CCAAAGCCACACTCAAAGAC
TACAACCGACAGATGTATGT CTTGAGTGATGGTGATGTTC

El polimorfismo menos 511a/c de interleucina 1 beta (IL1b)

GCTAGATGAAGAGCAAGCG CAGGCAGAGAGGGAAGGAGAG
TACAACCGACAGATGTATGT AACAGCGAGGGAGAACTG

El polimorfismo 34c/t de interleucina 4 (IL4)

GCTAGATGAAGAGCAAGCGCTCATTTCCCTCGGTTTCAG
TACAACCGACAGATGTATGTAGAACAGAGGGGAAGCAGT

El polimorfismo C93T de la caspasa 9

GCTAGATGAAGAGCAAGCGCGCGCCTGGAGTCTTAGTT

TACAACCGACAGATGTATGTGATCATATGGGGCCTGAACA

El polimorfismo R241G de la molécula de adhesión 1 (ICAM1)

GCTAGATGAAGAGCAAGCGCGAATGAAATGCCCCAGAGAA

TACAACCGACAGATGTATGTACTGTGGGGTTCAACCTCTG

El polimorfismo -590 del gen de interleucina 4 (IL4)

GCTAGATGAAGAGCAAGCGACCCAACTAGGCCTCACCT

TACAACCGACAGATGTATGTACAGGTGGCATCTTGAAAC

El polimorfismo T145C (Leu49Ser) del gen de la TIOPURINA S-METILTRANSFERASA 5 (TPMT5)

GCTAGATGAAGAGCAAGCGCCACTGACTGGGTGTGTGTC

TACAACCGACAGATGTATGTCTCAGTCCACTCTTGCCTTT

El polimorfismo A539T del gen de la TIOPURINA S-METILTRANSFERASA 6 (TPMT6)

GCTAGATGAAGAGCAAGCGCTTCCTTCCCTGCCTTTTGT

TACAACCGACAGATGTATGTCCCAACAACCTTACCTGGATG

El polimorfismo T681G (His227Gln) del gen de la TIOPURINA S-METILTRANSFERASA 7 (TPMT7)

GCTAGATGAAGAGCAAGCGCCATCCATTACATTTTCAGGCTTT

TACAACCGACAGATGTATGTGGTTGATGCTTTTGAAGAACG

El polimorfismo G644A (Arg215His) del gen de la TIOPURINA S-METILTRANSFERASA 8 (TPMT8)

GCTAGATGAAGAGCAAGCGTCCTGACCTCAAGTGATCCA

TACAACCGACAGATGTATGTTCGTTCTTCAAAGCATCAACC

El polimorfismo 341T>C de la N-acetiltransferasa 2 (arilamina N-acetiltransferasa)

GCTAGATGAAGAGCAAGCGTGTTGCTCTCCAGGTCAATCA

TACAACCGACAGATGTATGTGGCTGATCCTTCCCAGAAAT

El polimorfismo 481C>T de la N-acetiltransferasa 2 (arilamina N-acetiltransferasa)

GCTAGATGAAGAGCAAGCGTGACGGCAGGAATTACATTG

TACAACCGACAGATGTATGTTGTTTCTTCTTGGCAGGAGA

El polimorfismo 803A>G de la N-acetiltransferasa 2 (arilamina N-acetiltransferasa)

GCTAGATGAAGAGCAAGCGCCTGCCAAAGAAGAAACACC

TACAACCGACAGATGTATGTTCGTGAGGGTAGAGAGGATATCTG

El polimorfismo 857G>A de la N-acetiltransferasa 2 (arilamina N-acetiltransferasa)

GCTAGATGAAGAGCAAGCGCCTGCCAAAGAAGAAACACC

TACAACCGACAGATGTATGTTCGTGAGGGTAGAGAGGATATCTG

El polimorfismo 191G>A de la de N-Acetiltransferasa 2 (arilamina N-acetiltransferasa)

GCTAGATGAAGAGCAAGCGCCATGGAGTTGGGCTTAGAG

TACAACCGACAGATGTATGTCCATGCCAGTGCTGTATTTG

El polimorfismo Gly54Asp del gen de unión a lectina, manosa (MBL)

GCTAGATGAAGAGCAAGCGTGGCAGCGTCTTACTCAGAA

TACAACCGACAGATG⁸TATGTAGAACAGCCCAACACGTACC

Polimorfismo Hla B27 del antígeno de histocompatibilidad

GCTAGATGAAGAGCAAGCGCGCTACGTGGACGACACGCT

TACAACCGACAGATGTATGTCAGTCTGTGCCTTGGCGTTGC

Figura 4A: Tabla 4

Variables de ecuación

Etapa	B	E.T.	Wald	gl	Sig.	Exp(B)	I.C. 95,0% para EXP(B)	
							Inferior	Superior
Etapa 1ª	ANTI_PCC	0,001	4,890	1	0,027	1,001	1,000	1,002
	VSG	0,021	4,661	1	0,031	1,021	1,002	1,041
	@8REC(1)	0,797	2,247	1	0,134	2,220	0,782	6,297
	@40REC(1)	1,027	2,572	1	0,109	2,792	0,796	9,795
	@78REC(1)	1,651	4,365	1	0,037	5,211	1,108	24,513
Constante	-5,839	1,067	29,940	1	0,000	0,003		
Etapa 2ª	ANTI_PCC	0,001	5,919	1	0,015	1,001	1,000	1,002
	VSG	0,023	5,803	1	0,016	1,023	1,004	1,043
	@40REC(1)	1,130	3,286	1	0,070	3,097	0,912	10,511
	@78REC(1)	1,680	4,572	1	0,032	5,366	1,150	25,029
	Constante	-5,748	1,073	28,705	1	0,000	0,003	

a. Variable(s) introducidas en la etapa 1: ANTI_PCC, VSG, @8REC, @40REC, @78REC.

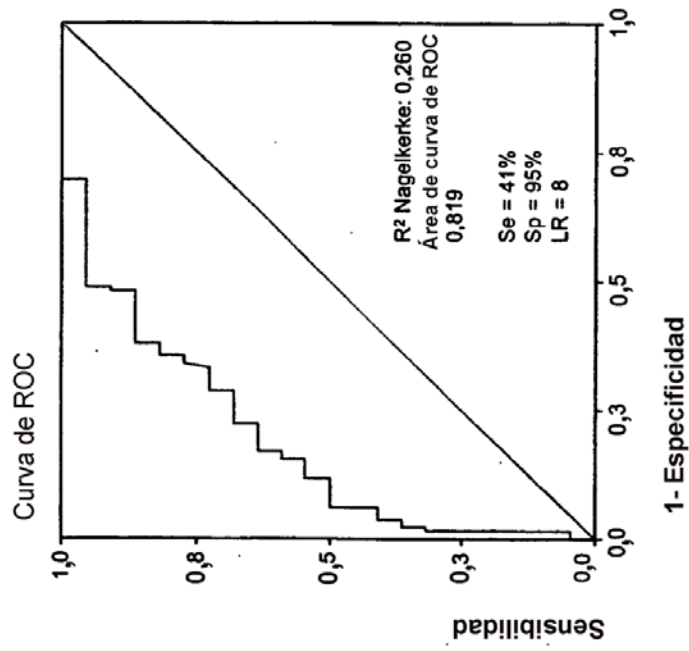


Figura 4B

Figura 5A: Tabla 5

Variables de ecuación

		B	E.T.	Wald	gl	Sig.	Exp(B)
Etapa 1 ^a	A PCC BI(1)	2,784	0,585	22,614	1	0,000	16,187
	Constante	-3,725	0,506	54,228	1	0,000	0,024
Etapa 2 ^b	A_PCC_BI(1)	2,842	0,593	22,968	1	0,000	17,156
	@41(1)	2,034	1,065	3,649	1	0,056	7,648
	Constante	-5,520	1,140	23,462	1	0,000	0,004
Etapa 3 ^c	A PCC BI(1)	2,892	0,602	23,114	1	0,000	18,030
	@41(1)	2,012	1,074	3,509	1	0,061	7,476
	@28_REC(1)	1,994	1,075	3,437	1	0,064	7,344
	Constante	-7,256	1,547	21,999	1	0,000	0,001

a. Variable(s) introducidas en la etapa 1: A_PCC_BI.
b. Variable(s) introducidas en la etapa 2: @41.
c. Variable(s) introducidas en la etapa 3: @28_REC

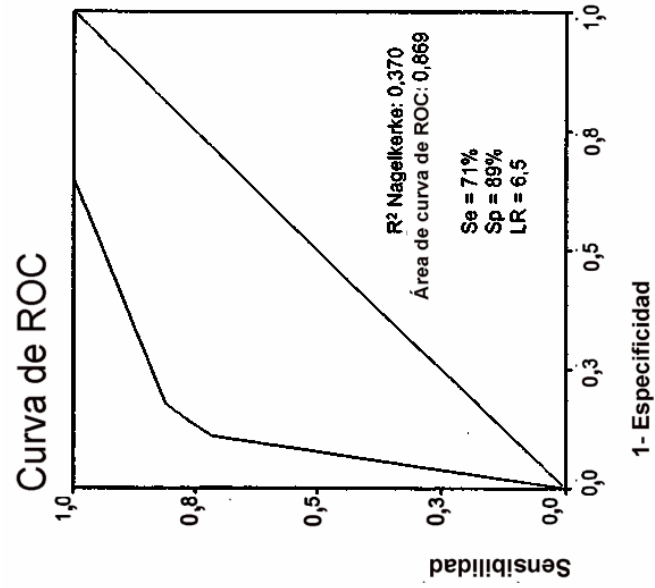


Figura 5B

Figura 6A: Tabla 6

Variables de ecuación

Etapa	B	E.T.	Wald	gl	Sig.	Exp(B)
@20			10,165	2	0,006	
@20(1)	-0,987	0,472	4,377	1	0,036	0,373
@20(2)	-2,175	0,706	9,489	1	0,002	0,114
@26			0,778	2	0,678	
@26(1)	-0,384	0,459	0,698	1	0,403	0,681
@26(2)	-9,082	31,635	0,082	1	0,774	0,000
@14			4,844	2	0,089	
@14(1)	0,976	0,447	4,769	1	0,029	2,654
@14(2)	-8,422	31,897	0,070	1	0,792	0,000
@15			5,454	2	0,065	
@15(1)	-1,114	0,478	5,425	1	0,020	0,328
@15(2)	-7,793	44,305	0,031	1	0,860	0,000
A_PCC_BI(1)	2,059	0,770	7,160	1	0,007	7,840
N°ARTDOL	0,108	0,032	11,416	1	0,001	1,114
Constante	-2,414	0,948	6,481	1	0,011	0,089

a. Variable(s) introducidas en la etapa 1: @20, @26, @14, @15, A_PCC_BI, N°ARTDOL.

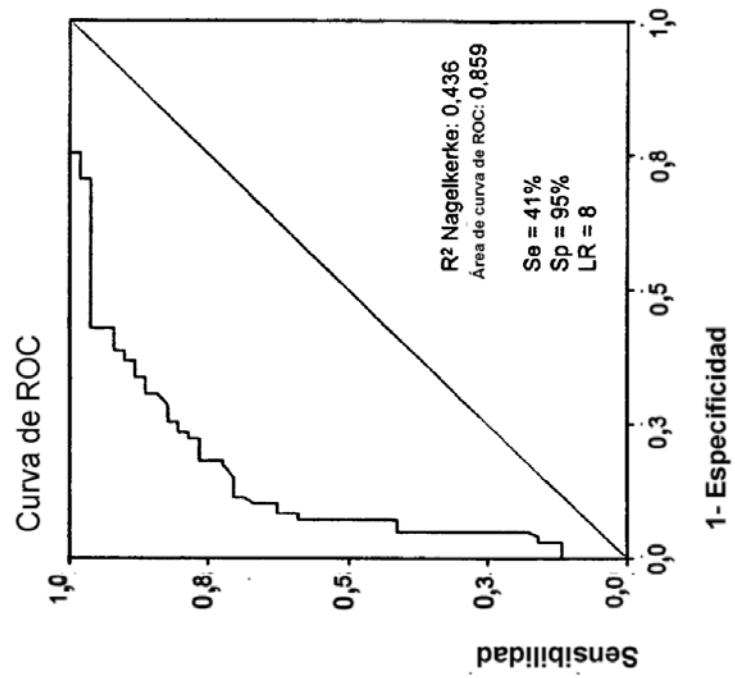


Figura 6B

Figura 7A: Tabla 7

Variables de ecuación

Etapa	B	E.T.	Wald	gl	Sig.	Exp(B)
1 ^a						
@27			6,158	2	0,046	
@27(1)	0,288	0,281	1,052	1	0,305	1,334
@27(2)	1,314	0,534	6,061	1	0,014	3,720
@16			5,916	2	0,052	
@16(1)	0,665	0,278	5,700	1	0,017	1,944
@16(2)	0,006	0,613	0,000	1	0,993	1,006
@73			7,101	2	0,029	
@73(1)	0,523	0,275	3,632	1	0,057	1,688
@73(2)	-1,026	0,693	2,192	1	0,139	0,358
EDADDEBU	-0,024	0,011	5,117	1	0,024	0,976
A_PCC_BI(1)	0,419	0,333	1,585	1	0,208	1,520
FR	0,001	0,001	4,409	1	0,036	1,001
PCR	0,091	0,043	4,537	1	0,033	1,095
Constante	-0,434	0,617	0,495	1	0,482	0,648

a. Variable(s) introducidas en la etapa 1: @27, @16, @73, EDADDEBU, A_PCC_BI, FR, PCR.

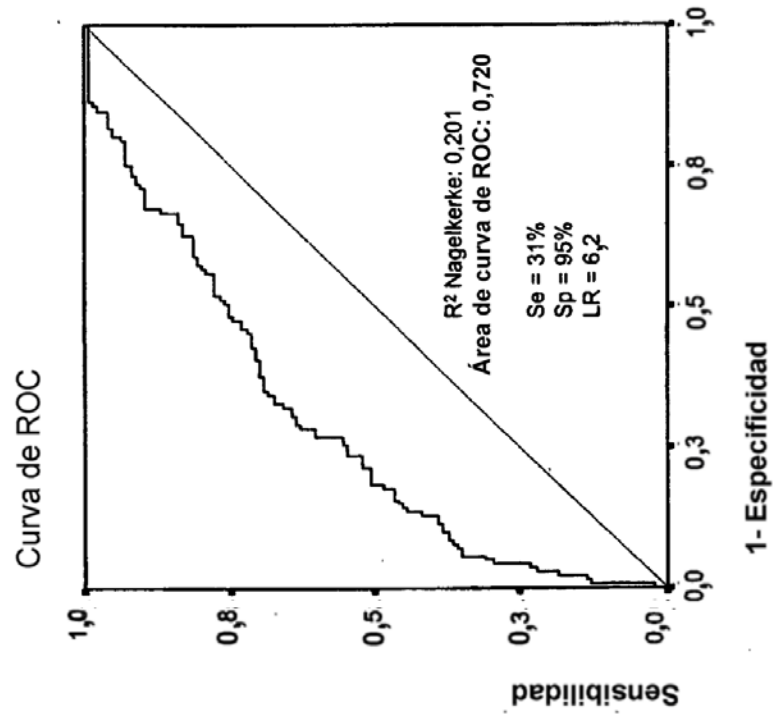


Figura 7B

Figura 8A: Tabla 8

Variables de ecuación

Etapa	B	E.T.	Wald	gl	Sig.	Exp(B)
1ª						
@2	-0,046	0,342	5,694	2	0,058	0,955
@2(1)	0,940	0,431	0,018	1	0,892	2,559
@2(2)			4,761	1	0,029	
@32			4,818	2	0,090	
@32(1)	0,678	0,316	4,613	1	0,032	1,970
@32(2)	-6,640	15,264	0,189	1	0,664	0,001
@29			4,048	2	0,132	
@29(1)	0,662	0,335	3,910	1	0,048	1,939
@29(2)	0,232	0,535	0,188	1	0,664	1,262
@12			5,972	2	0,050	
@12(1)	0,004	0,353	0,000	1	0,991	1,004
@12(2)	2,915	1,196	5,947	1	0,015	18,457
A_PCC_BI(1)	0,473	0,396	1,425	1	0,233	1,605
VSG	0,007	0,006	1,467	1	0,226	1,007
Constante	-2,886	0,532	29,477	1	0,000	0,056

a. Variable(s) introducidas en la etapa 1: @2, @32, @29, @12, A_PCC_BI, VSG.

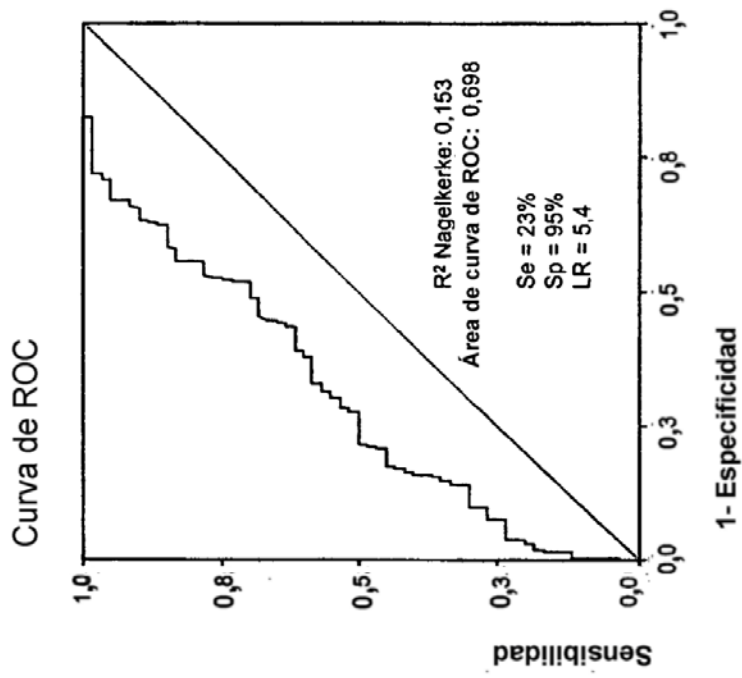


Figura 8B

Figura 9A: Tabla 9

Variables de ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)
Etapa 1 ^a						
@32REC(1)	-0,934	0,375	6,196	1	0,013	0,393
Constante	-0,234	0,217	1,158	1	0,282	0,792
Etapa 2 ^b						
@32REC(1)	-0,947	0,395	5,729	1	0,017	0,388
@50			7,129	2	0,028	
@50(1)	-0,134	0,408	0,108	1	0,743	0,875
@50(2)	1,213	0,529	5,262	1	0,022	3,364
Constante	-0,381	0,295	1,669	1	0,196	0,683
Etapa 3 ^c						
@32REC(1)	-0,914	0,406	5,065	1	0,024	0,401
@42(1)	1,588	0,626	6,430	1	0,011	4,892
@50			8,797	2	0,012	
@50(1)	-0,359	0,431	0,693	1	0,405	0,698
@50(2)	1,228	0,530	5,365	1	0,021	3,413
Constante	-0,475	0,301	2,496	1	0,114	0,622

a. Variable(s) introducidas en la etapa 1: @32REC.

b. Variable(s) introducidas en la etapa 2: @50.

c. Variable(s) introducidas en la etapa 3: @42.

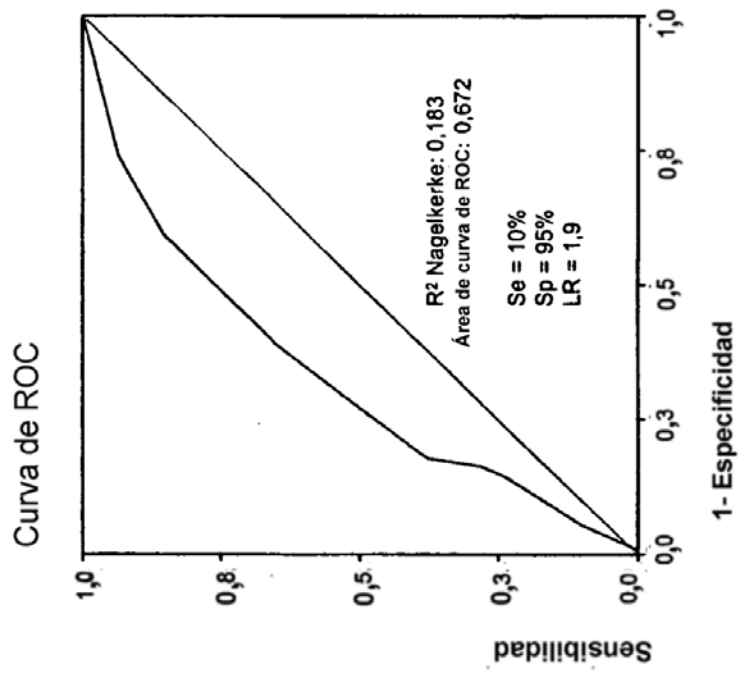


Figura 9B

Figura 10A: Tabla 10

Variables de ecuación

	B	E.T.	Wald	gl	Sig.	Exp(B)
Etapa 1 ^a	EDADDEBU	0,032	6,674	1	0,010	0,969
	Constante	0,041	0,572	1	0,943	1,042
Etapa 2 ^b	@:50	0,443	9,025	2	0,011	1,557
	@:50(1)	1,440	1,365	1	0,243	4,219
	@:50(2)	1,440	8,910	1	0,003	0,968
	Constante	-0,381	0,633	1	0,548	0,683

a. Variable(s) introducidas en la etapa 1: EDADDEBU

b. Variable(s) introducidas en la etapa 2: @:50

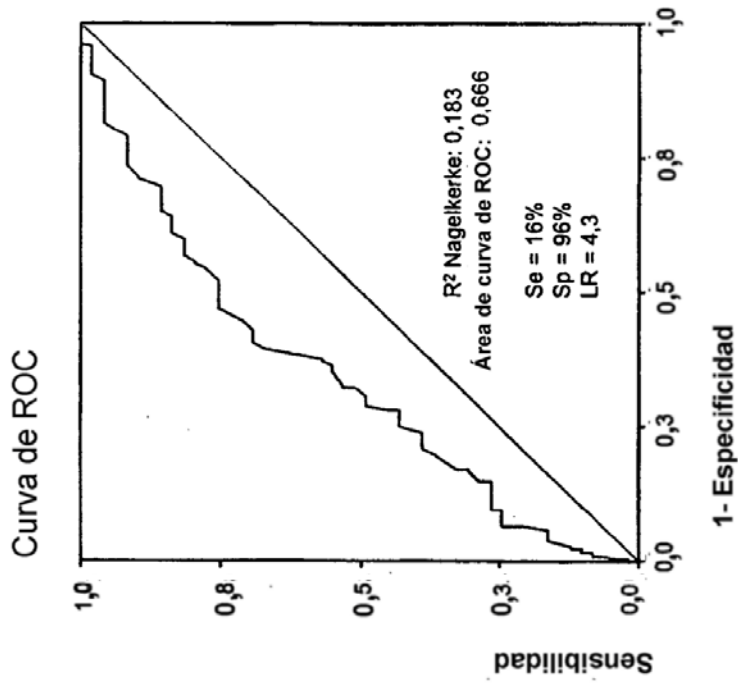


Figura 10B

Figura 11A: Tabla 11A

		MODELO POR FENOTIPO						
	1	2	3	4	5	6	7	
Variable	HAQ>2	Múltiples erosiones (RX2)	>3 Tratamientos diferentes (TTO_3OM)	Dejar el trabajo	Intervención quirúrgica	Intolerancia a metotrexato (Respuesta a la terapia)	Remisión	
1 SNP (identificador rs)								
2 GSTM1 (no disp.)								
3 FGB (rs1800790)	A							
4 IL4 (rs2070874)	T							
5 PTPN22 (rs2476601)							T	
6 VDR (rs1544410)	G							
7 NFKBIL1 (rs2071592)	T							
8 IL18 (rs187238)	G							
9 IL4R (rs1801275)	A							
10 PRR1 (rs7531668)	A							
11 IL1RN (rs419598)	T							
12 MCP1 (rs1024611)	C				T			
13 RUNX1 (rs2268277)				G	G			
14 TP53 (rs1042522)					A			
15 VDR (rs10735810)					T			
16 VEGF (rs3025039)				alelo nulo				
17 GSTT1 (no disp.)				T		T		
MTHFR(rs1801133)								
Clinica								
ANTI-PCC	altos niveles - alto riesgo		altos niveles - alto riesgo					
ANTI-PCC BI		Presencia - alto riesgo					Presencia - baja probabilidad	
VSG	altos niveles - alto riesgo				Presencia - alto riesgo	altos niveles - alto riesgo		
NoARTDOL		alto número - alto riesgo						
RF			altos niveles - alto riesgo					
PCR			altos niveles - alto riesgo					
Edad aparición			Aparición más temprana se asocia con 3 o más tratamientos			Aparición más temprana se asocia con intolerancia		

Figura 11B: Tabla 11B

Nº SNP	Símbolo de Gen	Nombre de Gen	Polimorfismo	ID (rs)
41	GSTM1	Glutión S-transferasa M1	alelo nulo	no disponible
40	FGB	Polipéptido beta de fibrinógeno B	G455A	rs1800790
78	IL4	Interleucina 4	C34T	rs2070874
28	PTPN22	Prot-tyr fosfatasa, no receptora 2	1858C/T	rs2476601
20	VDR	Receptor de vitamina D	AposVS7+283(b>B)	rs1544410
26	NFKBIL1	NF de pot de pp ligero κ tipo inh células B 1	T-62A	rs2071592
14	IL18	Interleucina 18	G-137C	rs187238
15	IL4R	Receptor de interleucina 4	Q551R	rs1801275
27	PARP1	Poli ADP-ribosa polimerasa 1	C1362T	rs7531668
16	IL1RN	El antagonista de receptor de IL-1	T2018C(IL-1RN*1)	rs419598
73	MCP1	Proteína quimiotáctica de monocitos 1	(-2518)G>A	rs1024611
2	RUNX1	Factor de transcripción relacionado con Runt	G2677T	rs2268277
32	TP53	Proteína tumoral p53	codón 72	rs1042522
29	VDR	Receptor de vitamina D	Fok1	rs10735810
12	VEGF	Factor de crecimiento endotelial vascular	C936T	rs3025039
42	GSTT1	Glutión S-transferasa T1	alelo nulo	no disponible
50	MTHFR	Metilтетrahidrofolato reductasa	C677T	rs1801133

Figura 12: Tabla 12

Fenotipo y SNP	Valor p
HAQ>2	
1 SNP 78 (rs2070874)	0,01723806
2 SNP 89 (rs1041983)	0,01859894
3 SNP 40 (rs1800790)	0,01905114
4 SNP 87 (rs no disponible)	0,0480329
RX2	
5 SNP 42 (rs no disponible)	0,00351956
6 SNP 26 (rs2071592)	0,02199691
7 SNP 15 (rs1801275)	0,03075416
8 SNP 34 (rs1800629)	0,03181104
TTO 3	
9 SNP 20 (rs1544410)	0,00625435
10 SNP 87 (rs no disponible)	0,04457328
Intervención quirúrgica	
11 SNP 41 (rs no disponible)	0,01281179
Remisión	
12 SNP 87 (rs no disponible)	0,02296336
13 SNP 28 (rs2476601)	0,0237744
14 SNP 27 (rs7531668)	0,02581015
Dejar el trabajo	
15 ^r SNP 32 (rs1042522)	0,0350
Intolerancia Metotrexato	
16 SNP 50 (rs1801133)f	0,0100
17 SNP 58 (rs2842934)	0,0360
18 SNP 78 (rs2070874)	0,0270

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante es únicamente para la comodidad del lector. No forma parte del documento de la patente europea. A pesar del cuidado tenido en la recopilación de las referencias, no se pueden excluir errores u omisiones y la EPO niega toda responsabilidad en este sentido.

Documentos de patentes citados en la descripción

- 10 • GB 0613844 A [0001] [0037] • WO 200600796 W [0157]

Literatura diferente de patentes citada en la descripción

- 15 • Arnett et al. *Arthritis and Rheumatism*, March 1988, vol. 31 (3), 315-24 [0005]
- 20 • Cutler DJ ; Zwick ME ; Carrasquillo MN ; Yohn CT ; Tobi KP ; Kashuk C ; Mathews DJ ; Shah N ; Eichler EE ; Warrington JA. *Genome Research*, 2001, vol. 11, 1913-1925 [0025]
- 25 • Johansson M. et al. *Arthritis Research & Therapy*, 2006, vol. 8, R19 [0025]
- Sokka T ; Krishnan E ; Häkkinen A ; Hannonen P. *Arthritis Rheum.*, January 2003, vol. 48 (1), 59-63 [0051]
- 30 • Rossi F ; Di Dia F ; Galipò O ; Pistorio A ; Valle M ; Magni-Manzoni S ; Ruperto N ; Tomà P ; Martini A ; Ravelli A. *Arthritis Rheum.*, 15 October 2006, vol. 55 (5), 717-23 [0052]
- 35 • Verstappen SM ; Hoes JN ; Ter Borg EJ ; Bijlsma JW ; Blaauw AA ; van Albada-Kuipers GA ; van Booma-Frankfort C ; Jacobs JW. *Ann Rheum Dis.*, November 2006, vol. 65 (11), 1506-11 [0055]
- McKendry RJ ; Dale P. *J Rheumatol.*, November 1993, vol. 20 (11), 1850-6 [0057]
- 40 • Arnette et al. *Arthritis and Rheumatism*, 1988 [0060]
- Vives-Corrons JL ; Jou JM. *Sangre (Barc)*, 1982, vol. 27 (4A), 573-8 [0069]
- Greiner A ; Plischke H ; Kellner H ; Gruber R. *Ann N Y Acad Sci.*, June 2005, vol. 1050, 295-303 [0071]
- Miller VM ; Redfield MM ; McConnell JP. *Curr Vasc Pharmacol.*, January 2007, vol. 5 (1), 15-25 [0072]
- PCR protocols; A Guide to Methods and Applications. Academic Press, 1990 [0136]
- Mullis et al. Cold Spring Harbor Symp. Quant. Biol. Stockton Press, 1989, vol. 51, 263 [0136]
- Ehrlich et al. *Science*, 1991, vol. 252, 1643-1650 [0136]
- Molecular Cloning: a Laboratory Manual. Cold Spring Harbor Laboratory Press, 1989 [0146]
- Current Protocols in Molecular Biology. John Wiley & Sons, 1992 [0146]
- Tejedor et al. *Clinical Chemistry*, 2005, vol. 51, 1137-1144 [0157]
- Eggerding FA ; Iovannisci DM ; Brinson E. ; Grossman P. ; Winn-Deen E.S. *Human Mutation*, 1995, vol. 5, 153-65 [0255] [0261]
- Pastinen T ; Raitio M ; Lindroos K ; Tainola P ; Peltonen L ; Syvanen AC. *Genome Research*, 2000, vol. 10, 1031-42 [0258]
- Sambrook et al. *Molecular Cloning: a laboratory manual*. 1989 [0329]