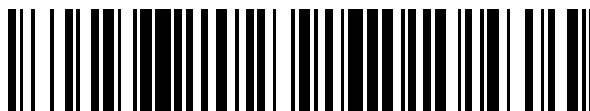


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 661 249**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **21.05.2008 E 12192837 (8)**

97 Fecha y número de publicación de la concesión europea: **20.12.2017 EP 2612924**

54 Título: **Métodos y composiciones para identificar y tratar el lupus**

30 Prioridad:

21.05.2007 US 939156 P

12.12.2007 US 13283 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
28.03.2018

73 Titular/es:

GENENTECH, INC. (100.0%)

1 DNA Way

South San Francisco, CA 94080-4990, US

72 Inventor/es:

BEHRENS, TIMOTHY W.;

HOM, GEOFFREY;

ORTMANN, WARD A. y

GRAHAM, ROBERT ROYAL

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 661 249 T3

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

DESCRIPCIÓN

Métodos y composiciones para identificar y tratar el lupus

5 Remisión a solicitudes relacionadas

Campo de la invención

10 La presente invención se refiere en líneas generales a un conjunto único de polimorfismos genéticos asociados con lupus y a composiciones y métodos para evaluar el riesgo de desarrollar lupus, así como para diagnosticar y tratar el lupus.

Antecedentes

15 El lupus es una enfermedad autoinmunitaria que implica anticuerpos que atacan al tejido conjuntivo. Se estima que la enfermedad afecta a casi 1 millón de americanos, principalmente mujeres entre las edades de 20-40. La forma principal de lupus es una diseminada (lupus eritematoso diseminado; SLE). El lupus eritematoso diseminado (SLE) es una enfermedad autoinmunitaria crónica con fuertes componentes genéticos, así como ambientales (véase, por ejemplo, Hochberg MC, Dubois' Lupus Erythematosus. 5.^a ed., Wallace DJ, Hahn BH, eds. Baltimore: Williams y
20 Wilkins (1997); Wakeland EK, *et al.*, Immunity 2001;15(3):397-408; Nath SK, *et al.*, Curr. Opin. Immunol. 2004;16(6):794-800). Los autoanticuerpos desempeñan una función importante en la patogénesis de SLE, y las diversas manifestaciones clínicas de la enfermedad se deben al depósito de complejos inmunitarios que contienen anticuerpo en los vasos sanguíneos, lo que da lugar a inflamación en el riñón, el cerebro y la piel, junto con efectos patogénicos directos de los autoanticuerpos que contribuyen a anemia hemolítica y trombocitopenia. El SLE se
25 caracteriza generalmente como un trastorno autoinmunitario del tejido conjuntivo con una amplia gama de características clínicas, que afecta predominantemente a mujeres, especialmente de ciertos grupos étnicos. D'Cruz *et al.*, Lancet (2007), 369:587-596. El SLE está asociado con la producción de anticuerpos antinucleares, complejos inmunitarios en circulación y activación del sistema del complemento. El SLE tiene una incidencia de aproximadamente 1 de cada 700 mujeres entre las edades de 20 y 60. El SLE puede afectar a cualquier sistema orgánico y puede causar daño tisular grave. Numerosos autoanticuerpos de diferente especificidad están presentes en SLE. Los pacientes con SLE a menudo producen autoanticuerpos que tienen especificidad anti-ADN, anti-Ro y antiplaqueta y que son capaces de iniciar características clínicas de la enfermedad, tales como glomerulonefritis, artritis, serositis, bloqueo cardíaco completo en recién nacidos y anomalías hematológicas. Estos autoanticuerpos también están posiblemente relacionados con alteraciones del sistema nervioso central. Arbuckle *et al.* describe el
35 desarrollo de autoanticuerpos antes de la aparición clínica de SLE (Arbuckle *et al.* N. Engl. J. Med. 349(16): 1526-1533 (2003)). El diagnóstico definitivo de lupus, incluyendo SLE, no es fácil, provocando que los médicos tengan que recurrir a una estrategia de clasificación de signos multifactoriales y basada en los síntomas. Gill *et al.*, American Family Physician (2003), 68(11): 2179-2186.

40 El lupus no tratado puede ser mortal según progresa de ataque de la piel y las articulaciones a los órganos internos, incluyendo pulmón, corazón y riñones (siendo la enfermedad renal la preocupación principal), haciendo de este modo que el diagnóstico temprano y preciso y/o la evaluación de riesgos de desarrollar lupus sea particularmente crítico. El lupus aparece principalmente como una serie de erupciones, con periodos intermedios de poca o ninguna manifestación de la enfermedad. El daño renal, medido por la cantidad de proteinuria en la orina, es una de las
45 áreas más agudas de daño asociadas con la patogenicidad en SLE, y representa al menos un 50 % de la mortalidad y la morbilidad de la enfermedad.

Uno de los retos más difíciles en el tratamiento clínico de las enfermedades autoinmunitarias complejas tales como lupus es la identificación precisa y temprana de la enfermedad en un paciente. Durante los años, se han realizado
50 muchos estudios de ligamiento y de genes candidatos para identificar factores genéticos que contribuyen a la susceptibilidad a SLE. Los haplotipos que portan los alelos de HLA de clase II *DRB1*0301* y *DRB1*1501* están claramente asociados con la enfermedad, así como la presencia de anticuerpos contra autoantígenos nucleares. Véase, por ejemplo, Goldberg MA, *et al.*, Arthritis Rheum 1976; 19(2):129-32; Graham RR, *et al.*, Am J Hum Genet 2002; 71(3):543-53; y Graham RR, *et al.*, Eur J Hum Genet 2007; 15(8):823-30). Más recientemente, se
55 descubrieron variantes del factor 5 regulador de interferón (*IRF5*) y el transductor de señales y activador de la transcripción 4 (*STAT4*) como factores de riesgo significativos para SLE. Véase, por ejemplo, Sigurdsson S, *et al.*, Am J Hum Genet 2005; 76(3):528-37; Graham RR, *et al.*, Nat Genet 2006; 38(5):550-55; Graham RR, *et al.*, Proc Natl Acad Sci U S A 2007; 104(16):6758-63; y Remmers EF, *et al.*, N Engl J Med 2007; 357(10):977-86. La identificación de *IRF5* y *STAT4* como genes de riesgo de SLE proporciona un apoyo al concepto de que la ruta de interferón de tipo I es central para la patogénesis de la enfermedad. Véase, por ejemplo, Ronnblom L, *et al.*, J Exp Med 2001; 194(12):F59-63; Baechler EC, *et al.*, Curr Opin Immunol 2004; 16(6):801-07; Banchereau J, *et al.*, Immunity 2006; 25(3):383-92; Miyagi T, *et al.*, J Exp Med 2007; publicación electrónica 10 de sept.

65 Para este fin, sería muy ventajoso tener métodos de diagnóstico de base molecular que puedan usarse para identificar de forma objetiva la presencia de y/o para clasificar la enfermedad en un paciente. Las variaciones genéticas o polimorfismos son variaciones genéticas que están presentes en el genoma de un organismo. Los

polimorfismos incluyen polimorfismos de un único nucleótido (SNP). Véase, por ejemplo, Carlson *et al.*, Nature 2004; 429:446-452; Bell, Nature 2004; 429:453-463; Evans y Relling, Nature 2004; 429:464-468. Los SNP se han correlacionado fuertemente con el riesgo y/o la presencia de enfermedades graves tales como diabetes (Sladek *et al.*, Nature 2007; 445: 881-828; Zeggini *et al.*, Science 2007; 26 de abril; Scott *et al.*, Science 2007; 26 de abril; y Saxena *et al.*, Science 2007; 26 de abril); enfermedad de Crohn (por ejemplo, Hampe *et al.*, Nat. Genet. 2007; febrero ;39(2):207-11); artritis reumatoide (por ejemplo, publicación de patente de Estados Unidos n.º 2007/0031848); y otras enfermedades autoinmunitarias inflamatorias (por ejemplo, patente de Estados Unidos 6.900.016; patente de Estados Unidos 7.205.106).

Hasta hace poco, no ha sido posible examinar de forma exhaustiva el genoma para variantes que modifiquen el riesgo a enfermedades complejas tales como lupus. Sin embargo, la generación de un catálogo extenso de la variación humana común (véase, por ejemplo, Nature 2005; 437(7063):1299-320) acoplado con los avances tecnológicos que permiten el genotipado rentable y preciso de cientos de miles de variantes, ha impulsado una revolución en la genética humana. Por primera vez, es posible realizar exploraciones de asociación de todo el genoma bien potenciadas para ensayar más completamente la hipótesis de que variantes comunes influyen en el riesgo. En los dos últimos años, esta tecnología se ha validado intensamente. Véase, por ejemplo, Dewan A, *et al.*, Science 2006; 314(5801):989-92; Nature 2007; 447(7145):661-78, Matarin M, *et al.*, Lancet neurology 2007; 6(5):414-20; Moffatt MF, *et al.*, Nature 2007; 448(7152):470-73; Plenge RM, *et al.*, N Engl J Med 2007; Saxena R, *et al.*, Science 2007; 316(5829):1331-36; Scott LJ, *et al.*, Science 2007; 316(5829):1341-45; Scuteri A, *et al.*, PLoS Genet 2007;3(7):e115. Los loci de riesgo identificados están proporcionando nuevas perspectivas sobre las rutas moleculares reguladas incorrectamente en enfermedades humanas.

Sin embargo, sigue habiendo una ausencia significativa de información creíble sobre las asociaciones de SNP con enfermedades complejas tales como lupus, por tanto, está claro que sigue existiendo la necesidad de identificar polimorfismos asociados con dichas enfermedades. Dichas asociaciones beneficiarían enormemente la identificación de la presencia de lupus en pacientes o la determinación de la susceptibilidad a desarrollar la enfermedad. Además, la información estadística y biológicamente significativa y reproducible respecto a la asociación de un SNP con una enfermedad compleja tal como lupus podría utilizarse como componente integrado en los esfuerzos para identificar subconjuntos específicos de pacientes que se esperaría que se beneficiaran significativamente del tratamiento con un agente terapéutico particular, por ejemplo, cuando el agente terapéutico es o ha demostrado ser en estudios clínicos de beneficio terapéutico en dicha subpoblación específica de pacientes con lupus.

La invención descrita en este documento cumple las necesidades descritas anteriormente y proporciona otros beneficios.

Breve resumen de la invención

La invención proporciona métodos precisos, simples y rápidos y composiciones para identificar el lupus, y para evaluar el riesgo de desarrollar lupus, basados al menos en parte en la identificación de una o más variaciones genéticas, por ejemplo, SNP, que está correlacionados con alta significación estadística y biológica con la presencia, subtipos y/o subpoblaciones de pacientes de lupus, como se define en las reivindicaciones. Más específicamente, la invención, se refiere a la identificación de un conjunto único de SNP, combinaciones únicas de dichos SNP y regiones de desequilibrio del ligamiento que están asociadas con lupus y sus subtipos, y subpoblaciones de pacientes que las padecen.

En particular, el conjunto único y/o combinaciones de SNP pueden usarse como un perfil o característica genética indicativo de un sujeto en riesgo de desarrollar lupus, o indicativo de la enfermedad o síntoma o afección de la misma. Los polimorfismos divulgados en este documento son útiles como biomarcadores para evaluar el riesgo de desarrollar lupus, así como dianas para el diseño de reactivos de diagnóstico. En algunos aspectos el SNP no está asociado con un gen. En otros aspectos, el SNP está asociado con un gen y puede estar localizado en una región intergénica o intragénica, y más particularmente, puede estar localizado en una región codificante o no codificante. Los genes asociados con un SNP de la presente invención pueden estar asociados con un gen desconocido o pueden estar asociados con un gen conocido, por ejemplo, ITGAM o BLK.

Los SNP identificados en este documento proporcionan dianas para el desarrollo de agentes terapéuticos para su uso en el diagnóstico y tratamiento de pacientes con lupus identificados genéticamente, incluyendo el diagnóstico y el tratamiento dirigido de subpoblaciones de pacientes con lupus que muestran una característica genética distintiva que comprende uno o más de los SNP de la invención. Por ejemplo, en un aspecto, los genes que contienen las variaciones genéticas identificadas en este documento y el ácido nucleico (por ejemplo, ADN o ARN) asociado con estos genes y proteínas codificadas por estos genes pueden usarse como dianas para el desarrollo de agentes terapéuticos (por ejemplo, compuestos de molécula pequeña, anticuerpos, agentes antisentido/iARN, etc.) o usarse directamente como agentes terapéuticos (por ejemplo, proteínas terapéuticas, etc.) para el tratamiento del lupus.

Por consiguiente, en un aspecto, la divulgación proporciona un conjunto de uno o más SNP que forman una característica genética única para evaluar el riesgo de desarrollar lupus. En un aspecto, la característica genética única comprende aproximadamente 1-10, 10-20, 20-30, 30-40 o 40-50 SNP seleccionados de cualquiera los SNP

expuestos en las figuras 1-17 y las tablas 1-10.

En un aspecto, la característica genética única comprende 1 o más SNP, 2 o más SNP, 3 o más SNP, 4 o más SNP, 5 o más SNP, 6 o más SNP, 7 o más SNP, 8 o más SNP, 9 o más SNP, 10 o más SNP, 11 o más SNP, 12 o más SNP, 13 o más SNP, 14 o más SNP, 15 o más SNP, 16 o más SNP, 17 o más SNP, 18 o más SNP, 19 o más SNP o 20 o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, los SNP de la característica genética se seleccionan de la tabla 6. Otro aspecto, los SNP se seleccionan del grupo que consiste en rs9888739, rs13277113, rs7574865, rs2269368, rs6889239, rs2391592 y rs21177770. En otro aspecto, los SNP se seleccionan del grupo que consiste en rs2187668, rs10488631, rs7574865, rs9888739, rs13277113, rs2431697, rs6568431, rs10489265, rs2476601, rs2269368, rs1801274, rs4963128, rs5754217, rs6445975, rs3129860, rs10516487, rs6889239, rs2391592 y rs2177770.

En otro aspecto, la divulgación proporciona métodos de evaluación de si un sujeto está en riesgo de desarrollar lupus detectando en una muestra biológica obtenida de dicho sujeto la presencia de una característica genética indicativa de riesgo de desarrollar lupus, donde dicha característica genética comprende un conjunto de uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, el conjunto de SNP comprende aproximadamente 1-10, 10-20, 20-30, 30-40 o 40-50 SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el conjunto de SNP comprende 2 o más SNP, 3 o más SNP, 4 o más SNP, 5 o más SNP, 6 o más SNP, 7 o más SNP, 8 o más SNP, 9 o más SNP, 10 o más SNP, 11 o más SNP, 12 o más SNP, 13 o más SNP, 14 o más SNP, 15 o más SNP, 16 o más SNP, 17 o más SNP, 18 o más SNP, 19 o más SNP o 20 o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el conjunto de SNP comprende 1-19 SNP seleccionados de la tabla 6. En otro aspecto, el conjunto de SNP comprende un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10. En otro aspecto, el conjunto de SNP comprende un SNP de ITGAM seleccionado de cualquiera de los SNP de ITGAM expuestos en las tablas 7-10. En otro aspecto, el conjunto de SNP comprende además un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10. En otro aspecto, el conjunto de SNP comprende uno o más SNP seleccionados de siguiente grupo de SNP: rs2187668, rs10488631, rs7574865, rs9888739, rs13277113, rs2431697, rs6568431, rs10489265, rs2476601, rs2269368, rs1801274, rs4963128, rs5754217, rs6445975, rs3129860, rs10516487, rs6889239, rs2391592 y rs2177770.

En otro aspecto, la divulgación proporciona métodos de diagnóstico del lupus en un sujeto detectando en una muestra biológica obtenida de dicho sujeto la presencia de una característica genética indicativa de lupus, donde dicha característica genética comprende un conjunto de uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

En otro aspecto, la divulgación proporciona un polinucleótido aislado o fragmento del mismo que es de al menos aproximadamente 10 nucleótidos de longitud, donde el polinucleótido o fragmento del mismo comprende: (a) una variación genética en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, o (b) el complemento de (a). En un aspecto, el polinucleótido aislado es un ADN genómico que comprende un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el polinucleótido aislado es un ARN que comprende un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

En un aspecto, la divulgación proporciona un polinucleótido asociado a PRO aislado o fragmento del mismo que es de al menos aproximadamente 10 nucleótidos de longitud, donde el polinucleótido asociado a PRO o fragmento del mismo comprende: (a) una variación genética en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, o (b) el complemento de (a). En un aspecto, el polinucleótido asociado a PRO es un ADN genómico que codifica un gen (y/o región reguladora del gen) que comprende un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el SNP está en una región de un cromosoma que no codifica un gen. En otro aspecto, el SNP está en una región intergénica de un cromosoma. En otro aspecto, el polinucleótido asociado a PRO es un cebador. En otro aspecto, el polinucleótido asociado a PRO es un oligonucleótido.

En otro aspecto, la divulgación proporciona un oligonucleótido que es (a) un oligonucleótido específico de alelo que hibrida con una región de un polinucleótido que comprende una variación genética en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en las figuras 1-17 y las tablas 1-10, o (b) el complemento de (a). En un aspecto, el SNP es un polinucleótido asociado a PRO que codifica un gen (o su región reguladora) que comprende un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el SNP está en un ADN genómico que codifica un gen (o su región reguladora) que comprende un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el SNP es una región no codificante del gen. En otro aspecto, el SNP está en una región codificante del gen. En otro aspecto, el oligonucleótido específico de alelo es un cebador específico de alelo.

En otro aspecto, la divulgación proporciona un kit que comprende uno cualquiera de los oligonucleótidos anteriores y, opcionalmente al menos una enzima. En un aspecto, la al menos una enzima es una polimerasa. En otro aspecto, la al menos una enzima es una ligasa.

- 5 En otro aspecto, la invención proporciona una micromatriz que comprende cualquiera de los oligonucleótidos anteriores.

10 En otro aspecto, la divulgación proporciona un método de detección de la ausencia o presencia de una variación en un polinucleótido en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en las figuras 1-17 y las tablas 1-10, comprendiendo el método (a) poner en contacto el ácido nucleico sospechoso de comprender la variación con un oligonucleótido específico de alelo que es específico para la variación en condiciones adecuadas para la hibridación del oligonucleótido específico de alelo con el ácido nucleico; y de detectar la ausencia o presencia de hibridación específica de alelo. En un aspecto, la variación comprende un SNP expuesto en figuras 1-17 y las tablas 1-10. En un aspecto, el polinucleótido es un polinucleótido asociado a PRO.

15 En otro aspecto, la divulgación proporciona un método de amplificación de un ácido nucleico que comprende una variación en un polinucleótido en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, comprendiendo el método (a) poner en contacto el ácido nucleico con un cebador que hibrida con el ácido nucleico en una secuencia 3' de la variación y (b) prolongar el cebador para generar un producto de amplificación que comprende la variación. En un aspecto, el polinucleótido es un polinucleótido asociado a PRO.

20 En otro aspecto la divulgación proporciona un método de determinación del genotipo de una muestra biológica de un mamífero, comprendiendo el método detectar, en un material de ácido nucleico derivado de la muestra biológica, la ausencia o presencia de una variación en un polinucleótido en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, el polinucleótido es un polinucleótido asociado a PRO.

25 En otro aspecto, la muestra biológica es conocida por o sospechosa de comprender un polinucleótido de la presente invención, donde el polinucleótido comprende una variación en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, la muestra biológica es un tejido enfermo. En otro aspecto, la detección comprende realizar un proceso seleccionado de un ensayo de prolongación de cebador; un ensayo de prolongación de cebador específico de alelo; un ensayo de incorporación de nucleótidos específica de alelo; un ensayo de hibridación de oligonucleótido específico de alelo; un ensayo de 5' nucleasa; un ensayo que emplea balizas moleculares; y un ensayo de ligamiento de oligonucleótidos.

30 En otro aspecto, la divulgación proporciona un método de subclasificación del lupus en un mamífero, comprendiendo el método detectar la presencia de uno o más de los SNP expuestos en las figuras 1-17 y las tablas 1-10, en una muestra biológica obtenida del mamífero, donde la muestra biológica es conocida por o sospechosa de comprender al menos un polinucleótido que comprende un SNP seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, el polinucleótido es un polinucleótido asociado a PRO.

35 En otro aspecto, la detección comprende realizar un proceso seleccionado de un ensayo de prolongación de cebador; un ensayo de prolongación de cebador específico de alelo; un ensayo de incorporación de nucleótidos específica de alelo; un ensayo de hibridación de oligonucleótido específico de alelo; un ensayo de 5' nucleasa, un ensayo que emplea balizas moleculares; y un ensayo de ligamiento de oligonucleótidos.

40 En otro aspecto, la divulgación proporciona un método para predecir si un sujeto con lupus responderá a un agente terapéutico contra el lupus, comprendiendo el método determinar si el sujeto comprende una variación de un polinucleótido en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en figuras 1-17 y las tablas 1-10, donde la presencia de una variación indica que el sujeto responderá al agente terapéutico. En otro aspecto, el polinucleótido es un polinucleótido asociado a PRO.

45 En otro aspecto, la divulgación proporciona un método de diagnóstico o pronóstico del lupus en un sujeto, comprendiendo el método detectar la presencia de una variación en un polinucleótido derivado de una muestra biológica obtenida del sujeto, donde: (a) la muestra biológica es conocida por comprender o es sospechosa de comprender un polinucleótido que comprende la variación; (b) la variación comprende o está localizada en una posición nucleotídica correspondiente a un SNP seleccionado de cualquiera de los SNP expuestos en figuras 1-17 y las tablas 1-10; y (c) la presencia de la variación es un diagnóstico o pronóstico de lupus en el sujeto.

50 En otro aspecto, la divulgación proporciona un método de diagnóstico o pronóstico del lupus en un sujeto, comprendiendo el método detectar la presencia de una variación en PRO o polinucleótido asociado a PRO derivado de una muestra biológica obtenida del sujeto, donde: (a) la muestra biológica es conocida por comprender o es

sospechosa de comprender PRO o polinucleótido asociado a PRO que comprende la variación; (b) la variación comprende o está localizada en una posición nucleotídica correspondiente a un SNP expuesto en las figuras 1-17 y las tablas 1-10; y (c) la presencia de la variación es un diagnóstico o pronóstico de lupus en el sujeto.

5 En otro aspecto, la divulgación proporciona un método de ayuda al diagnóstico o pronóstico del lupus en un sujeto, comprendiendo el método detectar la presencia de una variación en un polinucleótido derivado de una muestra biológica obtenida del sujeto, donde: (a) la muestra biológica es conocida por comprender o sospechosa de comprender un polinucleótido que comprende la variación; (b) la variación comprende o está localizada en una posición nucleotídica correspondiente a un SNP seleccionado de cualquiera de los SNP expuestos en las
10 figuras 1-17 y las tablas 1-10; y (c) la presencia de la variación es un diagnóstico o pronóstico de una afección o síntoma de lupus en el sujeto.

En otro aspecto, la divulgación proporciona un método de ayuda al diagnóstico o pronóstico del lupus en un sujeto, comprendiendo el método detectar la presencia de una variación en PRO o polinucleótido asociado a PRO derivado
15 de una muestra biológica obtenida del sujeto, donde: (a) la muestra biológica es conocida por comprender o sospechosa de comprender PRO o polinucleótido asociado a PRO que comprende la variación; (b) la variación comprende o está localizada en una posición nucleotídica correspondiente a un SNP seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10; y (c) la presencia de la variación es un diagnóstico o pronóstico de una afección o síntoma de lupus en el sujeto.

20 En otro aspecto, el polinucleótido comprende una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como expone en las figuras 1-17 y las tablas 1-10. En un aspecto, la variación está en ADN genómico que comprende un SNP seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, el SNP es una región cromosómica que no codifica un gen. En otro aspecto, el SNP está en una región intergénica.
25

En otro aspecto, el polinucleótido asociado a PRO codifica PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10. En un aspecto, la variación está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las
30 tablas 1-10. En un aspecto, el SNP está en una región no codificante del gen. En otro aspecto, el SNP está en una región codificante del gen.

En otro aspecto, la divulgación proporciona un método de identificación de un agente terapéutico eficaz para tratar el lupus en una subpoblación de pacientes, comprendiendo el método correlacionar la eficacia del agente con la presencia en el paciente de uno o más de los SNP seleccionados de cualquiera de los SNP expuestos en las
35 figuras 1-17 y las tablas 1-10, identificando de ese modo el agente como eficaz para tratar el lupus en dicha subpoblación de pacientes.

40 En otro aspecto, la divulgación proporciona un método de identificación de un agente terapéutico eficaz para tratar el lupus en una subpoblación de paciente, comprendiendo el método correlacionar la eficacia del agente con la presencia de una combinación de los SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, identificando de ese modo el agente como eficaz para tratar el lupus en dicha subpoblación de
45 pacientes.

En otro aspecto, la divulgación proporciona un método de tratamiento de una afección de lupus en un sujeto en que se sabe que está presente una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para tratar la afección.
50

En otro aspecto, la divulgación proporciona un método de tratamiento de un sujeto que tiene una afección de lupus, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para tratar la afección en un sujeto que tiene una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.
55

En otro aspecto, la divulgación proporciona un método de tratamiento de un sujeto que tiene una afección de lupus, comprendiendo el método administrar al sujeto un agente terapéutico que ha demostrado ser eficaz para tratar dicha afección en al menos un estudio clínico, donde el agente se administró a al menos cinco sujetos humanos que tenían cada uno una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único
60 nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, los al menos cinco sujetos tenían dos o más SNP diferentes en total para el grupo de al menos cinco sujetos. En otro aspecto, los al menos cinco sujetos tenían el mismo SNP para el grupo completo de al menos cinco sujetos.

65 En otro aspecto, la divulgación proporciona un método de tratamiento de un sujeto con lupus de una subpoblación de pacientes específicos con lupus, donde la subpoblación se caracteriza al menos en parte por la asociación con la

variación genética en una posición nucleotídica correspondiente a un SNP seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, y donde el método comprende administrar al sujeto una cantidad eficaz de un agente terapéutico que está aprobado como agente terapéutico para dicha subpoblación. En un aspecto, la subpoblación tiene nefritis lúpica. En otro aspecto, la subpoblación es femenina. En otro aspecto, la subpoblación es de ascendencia europea.

En otro aspecto, la divulgación proporciona un método que comprende fabricar un agente terapéutico contra el lupus, y envasar el agente con instrucciones para administrar el agente a un sujeto que tiene o se cree que tiene lupus y que tiene una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

En otro aspecto, la divulgación proporciona un método de especificación de un agente terapéutico para su uso en una subpoblación de pacientes con lupus, comprendiendo el método proporcionar instrucciones para administrar el agente terapéutico a una subpoblación de pacientes caracterizada por una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10

En otro aspecto, la divulgación proporciona un método para preparar un agente terapéutico para su uso en una subpoblación de pacientes con lupus, comprendiendo el método informar a una audiencia determinada acerca del uso del agente terapéutico para tratar la subpoblación de pacientes caracterizada por la presencia, en pacientes de dicha subpoblación, de una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

En otro aspecto, la divulgación proporciona un método para modular la señalización a través del receptor de linfocitos B en un sujeto en que se sabe que está presente una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para modular la señalización a través del receptor de linfocitos B.

En otro aspecto, la divulgación proporciona un método para modular la diferenciación de células Th17 en un sujeto en que se sabe que está presente una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para modular la diferenciación de células Th17.

En otro aspecto, la divulgación proporciona un conjunto de SNP que comprenden una característica genética indicativa del riesgo de desarrollar lupus, donde dicho conjunto de SNP comprende uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, el conjunto de SNP comprende aproximadamente 1-10, 10-20, 20-30, 30-40 o 40-50 SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el conjunto de SNP comprende uno o más SNP seleccionados del grupo que consiste en rs9888739, rs13277113, rs7574865, rs2269368, rs6889239, rs2391592 y rs21177770. En otro aspecto, el conjunto de SNP comprende 2 o más SNP, 3 o más SNP, 4 o más SNP, 5 o más SNP, 6 o más SNP, 7 o más SNP, 8 o más SNP, 9 o más SNP, 10 o más SNP, 11 o más SNP, 12 o más SNP, 13 o más SNP, 14 o más SNP, 15 o más SNP, 16 o más SNP, 17 o más SNP, 18 o más SNP, 19 o más SNP o 20 o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otro aspecto, el conjunto de SNP comprende 1-19 SNP seleccionados de la tabla 6. En otro aspecto, el conjunto de SNP comprende un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10. En otro aspecto, el conjunto de SNP comprende un SNP de ITGAM seleccionado de cualquiera de los SNP de ITGAM expuestos en las tablas 7-10. En otro aspecto, el conjunto de SNP comprende además un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10. En otro aspecto, el conjunto de SNP comprende uno o más SNP seleccionados del siguiente grupo de SNP: rs2187668, rs10488631, rs7574865, rs9888739, rs13277113, rs2431697, rs6568431, rs10489265, rs2476601, rs2269368, rs1801274, rs4963128, rs5754217, rs6445975, rs3129860, rs10516487, rs6889239, rs2391592 y rs2177770.

En otro aspecto, la divulgación proporciona un conjunto de SNP que comprende una característica genética indicativa de lupus donde dicho conjunto de SNP comprende uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

Breve descripción de los dibujos

La figura 1 representa los resultados de una exploración de asociación en todo el genoma en SLE que identifica 5 genes principales. Los datos representan 502 033 variantes de SNP tipadas en 3 series de muestra, para un total de 1311 casos de SLE y 3340 controles. El panel A muestra un diagrama cuantil-cuantil de la distribución del valor P observado frente a la distribución esperada del valor P nulo. Los rombos representan todos los valores P y los círculos representan los valores P después de la exclusión de las variantes de la región *HLA*, *IRF5* y *STAT4*. El panel B es una representación gráfica de los valores $-\log_{10} P$ del análisis combinado organizado por cromosoma. No

se muestran variantes de la región HLA adicionales ($N = 34$) con $P < 1 \times 10^{-13}$ en el panel B.

La figura 2 muestra que variantes asociadas de la región *BLK/C8orf13* se correlacionan con los niveles de expresión en linfocitos B transformados. A) los valores $-\log_{10} P$ de la región *BLK/C8orf13* se presentan. El color de los rombos representa las correlaciones r^2 con rs13277113. Todos los genes RefSeq en la región se presentan por encima de un diagrama que muestra el LD en la región determinada por análisis de cromosomas de control. Cabe destacar que esta región asociada en el cromosoma 8 está dentro de una inversión intracromosómica polimórfica común de 4,2 Mb (véase, por ejemplo, Giglio *et al.* Am J Hum Genet 2001;68(4):874-83 y Sugawara *et al.* Genomics 2003;82(2):238-44) y está asociada con niveles inusualmente bajos de LD prolongado en toda la región, como se muestra. Sin embargo, la asociación de *BLK/C8orf13* a SLE es independiente de la inversión. La expresión de *BLK* (B) y *C8orf13* (C) líneas de linfocitos B transformados de 210 creadores HapMap CEU sanos no relacionados se muestra estratificada por el genotipo en rs13277113. La significación de la expresión diferencial se determinó usando ensayos T de Student para datos independientes.

La figura 3 muestra que las variantes dentro del locus *ITGAM/ITGAX* están asociadas con SLE. El panel A muestra los valores $-\log_{10} P$ de la región *ITGAM/ITGAX*. El color de los rombos representa las correlaciones r^2 con rs11574637. Todos los genes RefSeq en la región se presentan por encima de un diagrama que muestra el LD en la región determinado por los cromosomas de control estudiados. El panel B representa la estructura genómica de *ITGAM*, los dominios proteicos principales conservados y la relación entre rs11574637 y dos alelos no sinónimos de *ITGAM*.

La figura 4 representa la frecuencia de características clínicas en las series de SLE 1-3 y los casos suecos.

La figura 5 representa los 50 locis principales asociados en SLE en una exploración de genoma completo en 1311 casos y 3340 controles.

La figura 6 representa los niveles de expresión de *BLK*, *C8orf1* y genes de control en 210 líneas de linfocitos B transformados de individuos HapMap.

La figura 7 representa la expresión de *BLK* en linfocitos B transformados de las poblaciones HapMap.

La figura 8 representa la asociación de variantes de la región *C8orf13/BLK* y *ITGAM/ITGAX* con SLE por serie de casos/control.

La figura 9 representa la asociación de las variantes *C8orf13/BLK* y *ITGAM/ITGAX* con los 11 criterios clínicos ACR para las series de SLE 1-3.

La figura 10 representa la asociación de las variantes *C8orf13/BLK* y *ITGAM/ITGAX* con los 11 criterios clínicos ACR para 521 casos suecos de SLE. En las muestras suecas, se examinaron 521 casos para una asociación de los criterios ACR. La significación estadística se evaluó por tablas de contingencia 2×2 y un ensayo de ji cuadrado. Los valores P calculados no se ajustaron para ensayo múltiple, ya que se sabe que los criterios ACR están correlacionados y una simple corrección de Bonferroni de $X = 0,05 / 11 = 0,0045$ probablemente sería demasiado conservativa.

La figura 11 representa la fórmula usada para combinar los valores Z corregidos ponderados para el tamaño de la serie y ajustados para el factor de inflación de control genómico residual (λ_{gc}). La varianza (σ^2) de cada serie se calculó donde p = la frecuencia del alelo en los casos y los controles. El valor Z combinado para las 3 series de SLE (Z^*) se calculó donde, Z_1 , Z_2 y Z_3 es igual al valor Z basado en el EIGENSTRAT corregido por ji cuadrado para la asociación de una variante a SLE de cada serie, y donde λ_1 , λ_2 y λ_3 es el factor de inflación de control genómico residual (λ_{gc}) después de la corrección de EIGENSTRAT para cada serie. Se aplica la siguiente clave a los encabezados de las figuras 12-17:

SNP n.º	Numeración arbitraria de SNP en un subconjunto/grupo de estudio de pacientes específico
Región n.º	Numeración arbitraria de regiones de desequilibrio del ligamiento en los subconjuntos específicos de pacientes
SNP ID	Número rsID de SNP
EIG P	Valor P de la estadística ji cuadrado de EIGENSTRAT.
P en principal	Valor P para este SNP en el grupo principal
P en mujeres	Valor P para este SNP en el subconjunto de femenino
Coordenada	Pares de bases del SNP en su cromosoma
cM	centiMorgans desde el inicio del cromosoma
MAF CEU	Frecuencia del alelo minoritario del SNP de muestras HapMap CEU
SNP antes de la región	El SNP inmediatamente antes de la región de desequilibrio del ligamiento (LD) que contiene el SNP indicado en SNP_ID.
SNP después de la región	El SNP inmediatamente después de la región LD
Por qué los genes elegidos (región)	Racional fundamento para que se elija esta región como región relevante
1.º gen en la región	El primero de los genes en la región indicada, enumerado en orden por coordenadas (pares de bases)
Descripción (Descr.)	Descripción del gen del Comité de Nomenclatura de Genes HUGO
IRIS	Relación de la media inmunitaria más alta/media no inmunitaria más alta; del estudio IRIS
SNP de la región LD	Cualquier SNP localizado en la región de desequilibrio del ligamiento definida por (i) la coordenada A y la coordenada B, o (ii) SNP A y SNP B, inclusive.

La figura 12 (A) y (B), juntas, representan el análisis de un subconjunto de nefritis lúpica, que muestra 11 regiones que contenían 20 SNP candidatos considerados probables de contener al menos un alelo de riesgo para nefritis lúpica. (C) y (D), juntos, proporcionan caracterización adicional de regiones de desequilibrio del ligamiento, identidad de ciertos genes dentro de estas regiones y criterios para identificar dichos genes.

La figura 13 (A) representa el análisis de un subconjunto femenino, que muestra 6 regiones adicionales que contienen 9 SNP candidatos considerados probables de contener al menos un alelo de riesgo. (B) proporciona caracterización adicional de regiones de desequilibrio del ligamiento, identidad de ciertos genes dentro de estas regiones y criterios para identificar dichos genes.

La figura 14 (A) representa el análisis del grupo principal que muestra 6 regiones adicionales que contienen 8 SNP candidatos considerados probables de contener al menos un alelo de riesgo. La figura 14 (B) proporciona caracterización adicional de regiones de desequilibrio del ligamiento, ciertos genes dentro de estas regiones y criterios para identificar dichos genes.

La figura 15 representa la definición de regiones de desequilibrio del ligamiento y SNP contenidos en las mismas, basándose en ciertos datos de la figura 12.

La figura 16 representa la definición de regiones de desequilibrio del ligamiento y SNP contenidos en las mismas, basándose en ciertos datos de la figura 13.

La figura 17 representa la definición de regiones de desequilibrio del ligamiento y SNP contenidos en las mismas, basándose en ciertos datos de la figura 14.

Descripción detallada de la invención

La invención proporciona métodos precisos, simple y rápidos y composiciones para identificar el lupus, y para evaluar el riesgo de desarrollar lupus, basados al menos en parte en la identificación de una o más variaciones genéticas, por ejemplo, SNP, que están correlacionadas con alta significación estadística y biológica con la presencia, subtipos y/o subpoblaciones de pacientes de lupus. Más específicamente, la invención se refiere a la identificación de un conjunto único de SNP, combinaciones únicas de dichos SNP y regiones de desequilibrio del ligamiento que están asociados con lupus y sus subtipos, y subpoblaciones de pacientes que padecen los mismos.

En particular, el conjunto único y/o combinaciones de SNP pueden usarse como perfil genético o característica indicativa de un sujeto en riesgo de desarrollar lupus, o indicativa de la enfermedad o síntomas o afección de la misma. Los polimorfismos divulgados en este documento son útiles como biomarcadores para evaluar el riesgo de desarrollar lupus, así como para dianas para el diseño de reactivos de diagnóstico. En algunas realizaciones, el SNP no está asociado con un gen. En otras realizaciones, el SNP está asociado con un gen y puede estar localizado en una región intergénica o intragénica, y más particularmente, puede estar localizado en una región codificante o no codificante. Los genes asociados con un SNP de la presente invención pueden estar asociados con un gen desconocido o pueden estar asociados con un gen conocido, por ejemplo, ITGAM o BLK.

Los SNP identificados en este documento proporcionan dianas para el desarrollo de agentes terapéuticos para su uso en el diagnóstico y tratamiento de pacientes con lupus identificados genéticamente, incluyendo el diagnóstico y el tratamiento dirigido de subpoblaciones de pacientes con lupus que muestran una característica genética distinta que comprende uno o más de los SNP de la presente invención. Por ejemplo, en una realización, los genes que contienen las variaciones genéticas identificadas en este documento y el ácido nucleico (por ejemplo, ADN o ARN) asociado con estos genes, y las proteínas codificadas por estos genes, pueden usarse como dianas para el desarrollo de agentes terapéuticos (por ejemplo, compuestos de molécula pequeña, anticuerpos, agentes antisentido/iARN, etc.) o usarse directamente como agentes terapéuticos (por ejemplo, proteínas terapéuticas, etc.) para el tratamiento del lupus.

Técnicas generales

La práctica de la presente invención empleará, salvo que se indique de otro modo, técnicas convencionales de biología molecular (incluyendo técnicas recombinantes), microbiología, biología celular, bioquímica e inmunología, que pertenecen a las habilidades de la técnica. Dichas técnicas se explican completamente en la bibliografía, tal como, "Molecular Cloning: A Laboratory Manual", segunda edición (Sambrook *et al.*, 1989); "Oligonucleotide Synthesis" (M. J. Gait, ed., 1984); "Animal Cell Culture" (R. I. Freshney, ed., 1987); "Methods in Enzymology" (Academic Press, Inc.); "Current Protocols in Molecular Biology" (F. M. Ausubel *et al.*, eds., 1987, y actualizaciones periódicas); "PCR: The Polymerase Chain Reaction", (Mullis *et al.*, eds., 1994).

Los cebadores, oligonucleótidos y polinucleótidos empleados en la presente invención pueden generarse usando técnicas convencionales conocidas en la técnica.

Salvo que se definan de otro modo, los términos técnicos y científicos usados en este documento tienen el mismo significado que el comprendido habitualmente por los expertos en la materia a la que pertenece esta invención. Singleton *et al.*, Dictionary of Microbiology and Molecular Biology 2.^a ed., J. Wiley & Sons (New York, N.Y. 1994), y March, Advanced Organic Chemistry Reactions, Mechanisms and Structure 4.^a ed., John Wiley & Sons (New York, N.Y. 1992), proporcionan a los expertos en la materia unas directrices generales para muchos de los términos usados en la presente solicitud.

I. Definiciones

Con fines de interpretación de esta memoria descriptiva, las siguientes definiciones tendrán aplicación y, siempre que sea apropiado, los términos usados en la forma singular también incluirán la forma plural y viceversa. En el caso de que cualquier definición expuesta a continuación entre en conflicto con cualquier documento incorporado en este documento por referencia, prevalecerá la definición expuesta a continuación.

"Lupus" o "afección de lupus", como se usa en este documento es una enfermedad o trastorno autoinmunitario que en general implica anticuerpos que atacan al tejido conjuntivo. La forma principal de lupus es una diseminada, el lupus eritematoso diseminado (SLE), incluyendo SLE cutáneo y SLE cutáneo subagudo, así como otros tipos de lupus (incluyendo nefritis, extrarrenal, cerebritis, pediátrico, no renal, discoide y alopecia). Véase, en líneas generales, D'Cruz *et al.*, *supra*.

El término "polinucleótido" o "ácido nucleico", que se usan indistintamente en este documento, se refieren a polímeros de nucleótidos de cualquier longitud e incluyen ADN y ARN. Los nucleótidos pueden ser desoxirribonucleótidos, ribonucleótidos, nucleótidos o bases modificadas, y/o sus análogos, o cualquier sustrato que pueda incorporarse en un polímero de ADN o ARN polimerasa. Un polinucleótido puede comprender nucleótidos modificados, tales como nucleótidos metilados y sus análogos. Si está presente, la modificación de la estructura nucleotídica puede conferirse antes o después del ensamblaje del polímero. La secuencia de nucleótidos puede interrumpirse por componentes no nucleotídicos. Un polinucleótido puede modificarse adicionalmente después de la polimerización, tal como por conjugación con un componente de marcaje. Otros tipos de modificaciones incluyen, por ejemplo, "capuchones", sustitución de uno o más nucleótidos de origen natural con un análogo, modificaciones internucleotídicas tales como, por ejemplo, aquellas con enlaces no cargados (por ejemplo, metilfosfonatos, fosfotriésteres, fosfoamidatos, carbamatos, etc.) y con enlaces cargados (por ejemplo, fosforotioatos, fosforoditioatos, etc.), aquellos que contienen restos colgantes, tales como, por ejemplo, proteínas (por ejemplo, nucleasas, toxinas, anticuerpos, péptido señal, poli-L-lisina, etc.), aquellos con intercaladores (por ejemplo, acridina, psoraleno, etc.), aquellos que contienen quelantes (por ejemplo, metales, metales radiactivos, boro, metales oxidantes, etc.), aquellos que contienen alquilantes, aquellos con enlaces modificados (por ejemplo, ácidos nucleicos alfa anoméricos, etc.), así como formas no modificadas del polinucleótido o polinucleótidos. Además, cualquiera de los grupos hidroxilo habitualmente presentes en los azúcares puede remplazarse, por ejemplo, por grupos fosfonato, grupos fosfato, protegerse por grupos protectores convencionales o activarse para preparar enlaces adicionales a nucleótidos adicionales, o puede conjugarse a soportes sólidos. El OH 5' o 3' terminal puede fosforilarse o sustituirse con aminas o restos de grupos de capuchón orgánicos de 1 a 20 átomos de carbono. Otros hidroxilos también pueden derivatizarse en grupos protectores convencionales. Los polinucleótidos también pueden contener formas análogas de azúcares de ribosa o desoxirribosa que son generalmente conocidas en la técnica incluyendo, por ejemplo, 2'-O-metil-2'-O-alil, 2'-fluoro o 2'-ácido-ribosa, análogos carbocíclicos de azúcar, azúcares α -anoméricos, azúcares epiméricos tales como arabinosa, xilasas o lixosas, azúcares de piranosa, azúcares de furanosa, sedoheptulosas, análogos acíclicos y análogos nucleosídicos abásicos tales como ribósido de metilo. Uno o más enlaces fosfodiéster pueden remplazarse por grupos de enlace alternativos. Estos grupos de enlace alternativos incluyen, aunque sin limitación, realizaciones donde el fosfato se reemplaza por P(O)S("tioato"), P(S)S("ditiato"), "(O)NR₂("amidato"), P(O)R, P(O)OR', CO o CH₂("formacetal"), en que cada R o R' es independientemente H o alquilo (1-20 C) sustituido o sin sustituir que contiene opcionalmente un enlace éter (--O--), arilo, alquenilo, cicloalquilo, cicloalquenilo o araldilo. No todos los enlaces en un polinucleótido tienen que ser idénticos. La descripción precedente se aplica a todos los polinucleótidos mencionados en este documento, incluyendo ARN y ADN.

"Oligonucleótido", como se usa en este documento se refiere a polinucleótidos monocatenarios cortos que son de al menos aproximadamente siete nucleótidos de longitud y de menos de aproximadamente 250 nucleótidos de longitud. Los oligonucleótidos pueden ser sintéticos. Los términos "oligonucleótido" y "polinucleótido" no son mutuamente excluyentes. La descripción anterior para polinucleótidos es igual y completamente aplicable a oligonucleótidos.

El término "cebador" se refiere a un polinucleótido monocatenario que es capaz de hibridar con un ácido nucleico y permitir la polimerización de un ácido nucleico complementario, generalmente proporcionando un grupo 3'-OH libre.

El término "PRO" se refiere a un polipéptido codificado por cualquier gen codificado por una secuencia de ácido nucleico localizada dentro de una región de desequilibrio del ligamiento (región LD), donde la región LD se determina de acuerdo con el conjunto de información expuesta en las figuras 1-17 y las tablas 1-10. En una realización, un PRO de la invención no incluye un polipéptido conocido en la técnica por causar lupus. En una realización, un PRO de la invención no incluye un polipéptido conocido en la técnica por estar asociado con lupus, por ejemplo, IRF5, o cualquier polipéptido codificado por un indicado en las tablas 5-9 del documento WO 2007/019219. La expresión "polinucleótido asociado a PRO" o "ácido nucleico asociado con PRO" se refiere a una molécula de ácido nucleico que comprende una secuencia contigua, donde la secuencia contigua comprende una posición identificada en este documento que muestra variación genética. En una realización, la posición que muestra variación genética está localizada en el extremo 5' o 3' de la secuencia contigua. En una realización, la posición que muestra variación genética en la secuencia contigua está flanqueada, en cualquiera o ambas de sus

regiones 5' y/o 3', por uno o más nucleótidos que constituyen la secuencia flanqueante de origen natural de la posición. En una realización, una posición que muestra variación genética es una posición correspondiente a un SNP indicado en cualquiera de las figuras 1-17 y las tablas 1-10.

5 La expresión "variación genética" o "variación nucleotídica" se refiere a un cambio en una secuencia de nucleótidos (por ejemplo, una inserción, eliminación, inversión o sustitución de uno o más nucleótidos, tal como un polimorfismo de un único nucleótido (SNP)) respecto a una secuencia de referencia (por ejemplo, una secuencia encontrada habitualmente y/o de tipo silvestre y/o la secuencia de un alelo principal). La expresión también abarca el cambio correspondiente en el complemento de la secuencia de nucleótidos, salvo que se indique de otro modo. En una
10 realización, una variación genética es un polimorfismo somático. En una realización, una variación genética es un polimorfismo de la línea germinal.

Un "polimorfismo de un único nucleótido" o "SNP" se refiere a una única posición de base en una molécula de ARN o ADN (por ejemplo, un polinucleótido), en que existen diferentes alelos o nucleótidos alternativos en una población.
15 La posición del SNP (mencionado indistintamente en este documento como SNP, sitio SNP, locus SNP) está habitualmente precedida y seguida por secuencias altamente conservadas del alelo (por ejemplo, secuencias que varían en menos de 1/100 o 1/1000 miembros de las poblaciones). Un individuo puede ser homocigótico o heterocigótico para un alelo en cada posición SNP.

20 La expresión "variación de aminoácido" se refiere a un cambio en una secuencia de aminoácidos (por ejemplo, una inserción, sustitución o eliminación de uno o más aminoácidos, tal como una eliminación interna o un truncamiento N o C-terminal) respecto a una secuencia de referencia.

El término "variación" se refiere a una variación nucleotídica o una variación de aminoácido.

25 La expresión "una variación genética en una posición nucleotídica correspondiente a un SNP", "una variación nucleotídica en una posición nucleotídica correspondiente a un SNP" y variantes gramaticales de las mismas se refieren a una variación nucleotídica en una secuencia polinucleotídica en el relativo correspondiente a la posición nucleotídica ocupada por el SNP en el genoma. La expresión también abarca la variación correspondiente en el
30 complemento de la secuencia de nucleótidos, salvo que se indique de otro modo. En algunas realizaciones la variación nucleotídica es una secuencia polinucleotídica asociada a PRO en el relativo correspondiente a la posición nucleotídica ocupada por el SNP en el genoma.

La expresión "SNP de la región de desequilibrio del ligamiento" o "SNP de la región LD" se refiere a un SNP
35 presente en una región específica de ADN, estando definida dicha región por marcadores apropiados de ácido nucleico/genómicos, por ejemplo, coordenadas o SNP. En una realización, una región LD está definida por una primera coordenada (por ejemplo, coordenada A) y una segunda coordenada (por ejemplo, coordenada B), refiriéndose ambas coordenadas al mismo cromosoma. En una realización, una región LD está definida por un primer SNP (por ejemplo, SNP A) y un segundo SNP (por ejemplo, SNP B). Por tanto, en una realización, un SNP de
40 la región LD se refiere a un SNP localizado en una región de ácido nucleico (por ejemplo, región genómica) que varía de una primera coordenada hasta una segunda coordenada, o de un primer SNP hasta un segundo SNP. Los ejemplos de dichas regiones LD y SNP de región LD se muestran en las figuras 1-17 y las tablas 1-10.

El término "matriz" o "micromatriz" se refiere a una disposición ordenada de elementos en serie hibridables, preferiblemente sondas polinucleotídicas (por ejemplo, oligonucleótidos) en su sustrato. El sustrato puede ser un sustrato sólido, tal como un portaobjetos de vidrio o un sustrato semisólido, tal como una membrana de nitrocelulosa.

El término "amplificación" se refiere al proceso de producir una o más copias de una secuencia de ácido nucleico de referencia o su complemento. La amplificación puede ser lineal o exponencial (por ejemplo, PCR). Una "copia" no significa necesariamente complementariedad o identidad de secuencia perfecta respecto a la secuencia molde. Por ejemplo, las copias pueden incluir análogos nucleotídicos tales como desoxiinosina, alteraciones intencionadas de la secuencia (tales como alteraciones de la secuencia introducidas a través de un cebador que comprende una secuencia que es hibridable, pero no completamente complementaria al molde) y/o errores de secuencia que se
55 producen durante la amplificación.

La expresión "oligonucleótido específico de alelo" se refiere a un oligonucleótido que hibrida con una región de un ácido nucleico diana que comprende una variación nucleotídica (generalmente una sustitución). "Hibridación específica de alelo" significa que, cuando se hibrida un oligonucleótido específico de alelo con su ácido nucleico diana, un nucleótido en el oligonucleótido específico de alelo forma pares de base específicas con la variación nucleotídica. Un oligonucleótido específico de alelo con capacidad de hibridación específica de alelo con respecto a una variación nucleotídica particular se dice que es "específico para" esa variación.

La expresión "cebador específico de alelo" se refiere un oligonucleótido específico de alelo que es un cebador.

65 La expresión "ensayo de prolongación de cebador" se refiere a un ensayo en que se añaden nucleótidos a un ácido

nucleico, produciendo un ácido nucleico más largo, o "producto de extensión", que se detecta directa o indirectamente. Los nucleótidos pueden añadirse para prolongar el extremo 5' o 3' del ácido nucleico.

5 La expresión "ensayo de incorporación de nucleótidos específica de alelo" se refiere a un ensayo de prolongación de cebador en que un cebador (a) hibrida con un ácido nucleico diana en una región que está 3' o 5' de una variación nucleotídica y (b) se prolonga por una polimerasa, incorporando de ese modo en el producto de extensión un nucleótido que es complementario a la variación nucleotídica.

10 La expresión "ensayo de prolongación de cebador específico de alelo" se refiere a un ensayo de prolongación de cebador en que se hibrida un cebador específico de alelo a un ácido nucleico diana y se prolonga.

15 La expresión "ensayo de hibridación de oligonucleótido específico de alelo" se refiere a un ensayo en que (a) un oligonucleótido específico de alelo hibrida con un ácido nucleico diana y (b) se detecta la hibridación directa o indirectamente.

La expresión "ensayo de 5' nucleasa" se refiere a un ensayo en que la hibridación de un oligonucleótido específico de alelo con un ácido nucleico diana permite la escisión nucleolítica de la sonda hibridada, produciendo una señal detectable.

20 La expresión "ensayo que emplea balizas moleculares" se refiere a un ensayo en que la hibridación de un oligonucleótido específico de alelo con un ácido nucleico diana produce un nivel de señal detectable que es mayor que el nivel de señal detectable emitida por el oligonucleótido libre.

25 La expresión "ensayo de ligamiento de oligonucleótido" se refiere a un ensayo en que un oligonucleótido específico de alelo y un segundo oligonucleótido hibridan adyacentes entre sí en un ácido nucleico diana y se ligan juntos (directa o indirectamente a través de nucleótidos intermedios) y el producto de ligamiento se detecta directa o indirectamente.

30 La expresión "secuencia diana", "ácido nucleico diana" o "secuencia de ácido nucleico diana" se refiere en líneas generales a una secuencia polinucleotídica de interés en que se sospecha o se sabe que reside una variación nucleotídica, incluyendo copias de dicho ácido nucleico diana generado por amplificación.

35 Como se usa en este documento, un sujeto "en riesgo" de desarrollar lupus puede tener o no enfermedad o síntomas de enfermedad detectables, y puede haber presentado o no enfermedad o síntomas de la enfermedad detectables antes de los métodos de tratamiento descritos en este documento. "En riesgo" indica que un sujeto tiene uno o más factores de riesgo, que son parámetros medibles que se correlacionan con el desarrollo de lupus, como se describe en este documento y sabe en la técnica. Un sujeto que tiene uno o más de estos factores de riesgo tiene una mayor probabilidad de desarrollar lupus que un sujeto sin uno o más de estos factores de riesgo. Por ejemplo, en algunas realizaciones, un sujeto "en riesgo" de desarrollar lupus tiene una característica genética que comprende uno o más de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En otra realización, un sujeto "en riesgo" de desarrollar lupus tiene una característica genética que comprende uno o más de los SNP expuestos en la tabla 6.

El término "detección" incluye cualquier medio de detección, incluyendo detección directa e indirecta.

45 El término "diagnóstico" se usa en este documento para hacer referencia a la identificación o clasificación de un estado molecular o patológico, enfermedad o afección. Por ejemplo, "diagnóstico" puede hacer referencia a la identificación de un tipo particular de afección de lupus, por ejemplo, SLE. "Diagnóstico" también puede hacer referencia a la clasificación de un subtipo particular de lupus, por ejemplo, por afectación tisular/orgánica (por ejemplo, nefritis lúpica) por características moleculares (por ejemplo, una subpoblación de pacientes caracterizada por una o más variaciones genéticas en un gen o región de ácido nucleico particular).

50 La expresión "ayudar al diagnóstico" se usa en este documento para hacer referencia a métodos que ayudan a hacer una determinación clínica respecto a la presencia, grado u otra característica, de un tipo particular de síntoma o afección de lupus. Por ejemplo, un método de ayuda al diagnóstico de lupus puede comprender medir la cantidad o detectar la presencia o ausencia de uno o más SNP en una muestra biológica de un individuo. En otro ejemplo, un método de ayuda al diagnóstico de lupus puede comprender medir la cantidad o detectar la presencia de uno o más SNP en una muestra biológica de un individuo.

60 El término "pronóstico" se usa en este documento para hacer referencia a la predicción de la probabilidad de síntomas de enfermedad atribuibles a afección autoinmunitaria incluyendo, por ejemplo, recidiva, brote y resistencia a fármacos, de una enfermedad autoinmunitaria tal como lupus. El término "predicción" se usa en este documento para hacer referencia a la probabilidad de que un paciente responda de forma favorable o desfavorable a un fármaco o conjunto de fármacos. En una realización, la predicción se refiere al grado de esas respuestas. En una realización, la predicción se refiere a si un paciente sobrevivirá o mejorará y/o la probabilidad de que un paciente sobreviva o mejores después del tratamiento, por ejemplo, tratamiento con un agente terapéutico particular o durante un cierto periodo de tiempo sin recidiva de la enfermedad. Los métodos predictivos de la invención pueden usarse

clínicamente para tomar decisiones de tratamiento eligiendo las modalidades de tratamiento más apropiadas para cualquier paciente particular. Los métodos predictivos de la presente invención son herramientas valiosas en la predicción de si un paciente probablemente responderá de forma favorable a un régimen de tratamiento, tal como un régimen terapéutico dado, incluyendo, por ejemplo, la administración de un agente terapéutico dado o combinación, intervención quirúrgica, tratamiento con esteroides, etc., o si es probable una supervivencia a largo plazo del paciente después de un régimen terapéutico. El diagnóstico de SLE puede ser de acuerdo con los actuales criterios del American College of Rheumatology (ACR). La enfermedad activa puede definirse por un criterio del British Isles Lupus Activity Group (BILAG) "A" o dos criterios BILAG "B". Algunos signos, síntomas u otros indicadores usados para diagnosticar SLE están adaptados de: Tan *et al.* "The Revised Criteria for the Classification of SLE" *Arth Rheum* 25 (1982) puede ser eritema vespertilio, como eritema en las mejillas, eritema discoide o parches rojos en relieve, fotosensibilidad tal como reacción a la luz del sol, que provoca el desarrollo o el aumento de la erupción cutánea, úlceras bucales tales como úlceras en la nariz o la boca, habitualmente indoloras, artritis, tales como artritis no erosiva que implica dos o más articulaciones periféricas (artritis en que los huesos alrededor de las articulaciones no quedan destruidos) serositis, pleuritis o pericarditis, trastorno renal tal como un exceso de proteínas en la orina (mayor de 0,5 g/día o 3+ en tiras reactivas) y/o costras celulares (elementos anómalos derivados de la orina y/o glóbulos blancos y/o células tubulares del riñón), signos, síntomas u otros indicadores neurológicos, agarrotamiento (convulsiones) y/o psicosis en ausencia de fármacos o alteraciones metabólicas que se sabe que causan dichos efectos, y signos, síntomas u otros indicadores hematológicos tales como anemia hemolítica o leucopenia (recuento de glóbulos blancos por debajo de 4000 células por milímetro cúbico) o linfopenia (menos de 1500 linfocitos por milímetro cúbico) o trombocitopenia (menos de 100 000 plaquetas por milímetro cúbico). La leucopenia y la linfopenia generalmente deben detectarse en dos o más ocasiones. La trombocitopenia generalmente debe detectarse en ausencia de fármacos que se sabe que la induce. La invención no está limitada a estos signos, síntomas u otros indicadores de lupus.

Como se usa en este documento, "tratamiento" se refiere a intervención clínica en un intento por alterar el curso natural del individuo o la de la célula que se está tratando, y puede realizarse antes o durante el curso de la patología clínica. Los efectos deseados del tratamiento incluyen prevención de la aparición o recidiva de una enfermedad o una afección o síntoma de la misma, alivio de una afección o síntoma de la enfermedad, disminución de cualquier consecuencia patológica directa o indirecta de la enfermedad, disminución de la tasa de progresión de la enfermedad, mejora o paliación del estado patológico y consecución de la remisión o del pronóstico mejorado. En algunas realizaciones, los métodos y composiciones de la invención son útiles en intentos por retardar el desarrollo de una enfermedad o trastorno.

Una "cantidad eficaz" se refiere a una cantidad eficaz, a dosificaciones y durante periodos de tiempo necesarios, para conseguir el resultado terapéutico o profiláctico deseado. Una "cantidad terapéuticamente eficaz" de un agente terapéutico puede variar de acuerdo con factores tales como el estado patológico, la edad, el género y el peso del individuo, y la capacidad del anticuerpo de provocar una respuesta deseada en el individuo. Una cantidad terapéuticamente eficaz también es una en que cualquier efecto tóxico o perjudicial del agente terapéutico se compensa por los efectos terapéuticamente beneficiosos. Una "cantidad profilácticamente eficaz" se refiere a una cantidad eficaz, a dosificaciones y durante periodos de tiempo necesarios, para conseguir el resultado profiláctico deseado. Normalmente, aunque no necesariamente, ya que una dosis profiláctica se usa en sujetos antes de o en una fase prematura de la enfermedad, la cantidad profilácticamente eficaz será menor que la cantidad terapéuticamente eficaz.

Un "individuo", "sujeto" o "paciente" es un vertebrado. En ciertas realizaciones, el vertebrado es un mamífero. Los mamíferos incluyen, aunque sin limitación, primates (incluyendo seres humanos y primates no humanos) y roedores (por ejemplo, ratones y ratas). En ciertas realizaciones, un mamífero es un ser humano.

Una "subpoblación de pacientes" y variaciones gramaticales de la misma, como se usan en este documento, se refiere a un subconjunto de pacientes caracterizado por tener una o más características medibles y/o identificables distintivas que distinguen el subconjunto de pacientes de otros en la categoría de enfermedad más amplia a la que pertenece. Dichas características incluyen subcategorías de enfermedad (por ejemplo, nefritis lúpica), género, estilo de vida, historial clínico, órganos/tejidos afectados, historial de tratamientos, etc. En una realización, una subpoblación de pacientes se caracteriza por características genéticas, incluyendo variaciones genéticas en particular las posiciones y/o regiones de nucleótidos (tales como SNP).

El término "muestra", como se usa en este documento, se refiere a una composición que se obtiene o deriva de un sujeto de interés que contiene una entidad celular y/o molecular diferente que se tiene que caracterizar y/o identificar, por ejemplo, basándose en características físicas, bioquímicas, químicas y/o fisiológicas. Por ejemplo, la expresión "muestra de enfermedad" y variaciones de la misma se refiere a cualquier muestra obtenida de un sujeto de interés que se esperaría o que se sabe que contiene la entidad celular y/o molecular que se tiene que caracterizar.

Por "muestra tisular o celular" se entiende un conjunto de células similares obtenidas de un tejido de un sujeto o paciente. La fuente de la muestra tisular o celular puede ser tejido sólido como de una muestra o biopsia o aspirado orgánico o tisular fresco, congelado y/o conservado, sangre o cualquier constituyente de la sangre; líquidos

corporales tales como líquido cefalorraquídeo, líquido amniótico, líquido peritoneal o líquido intersticial; células de cualquier momento en la gestación o desarrollo del sujeto. La muestra tisular también puede ser células o líneas celulares primarias o cultivadas. Opcionalmente, la muestra tisular o celular se obtiene de un tejido/órgano enfermo. La muestra tisular puede contener compuestos que no están entremezclados de forma natural con el tejido en la naturaleza tales como conservantes, anticoagulantes, tamponantes, fijadores, nutrientes, antibióticos o similares. Una "muestra de referencia", "célula de referencia", "tejido de referencia", "muestra de control", "célula de control" o "tejido de control", como se usan en este documento, se refiere a una muestra, célula o tejido obtenido de una fuente que se sabe o se cree que no está afectada con la enfermedad o afección para la que se está usando un método o composición de la invención para identificarla. En una realización, una muestra de referencia, una célula de referencia, un tejido de referencia, una muestra de control, una célula de control o un tejido de control se obtiene de una parte sana del organismo del mismo sujeto o paciente en que se está identificando una enfermedad o afección usando una composición o método de la invención. En una realización, una muestra de referencia, una célula de referencia, un tejido de referencia, una muestra de control, una célula de control o un tejido de control se obtiene de una parte sana del organismo de un individuo que no es el sujeto o paciente en que se está identificando una enfermedad o afección usando una composición o método de la invención.

Para los fines de este documento una "sección" de una muestra tisular significa una parte individual o trozo de una muestra tisular, por ejemplo, una porción fina de tejido o células cortadas de una muestra tisular. Se entiende que pueden tomarse múltiples secciones de muestras tisulares y someterse a análisis de acuerdo con la presente invención, con la condición de que se entienda que la presente invención comprende un método por el que se analiza la misma sección de muestra tisular a nivel tanto morfológico como molecular, o se analiza con respecto tanto a la proteína como al ácido nucleico.

Por "correlacionar" o "que se correlaciona" se entiende comparar, de cualquier manera, el rendimiento y/o los resultados de un primer análisis o protocolo con el rendimiento y/o los resultados de un segundo análisis o protocolo. Por ejemplo, se pueden usar los resultados de un primer análisis o protocolo en la realización de un segundo protocolo y/o se pueden usar los resultados de un primer análisis o protocolo para determinar si debe realizarse un segundo análisis o protocolo. Con respecto a la realización del análisis o protocolo de expresión génica, se pueden usar los resultados del análisis o protocolo de expresión génica para determinar si debe realizarse un régimen terapéutico específico.

La palabra "marcador" cuando se usa en este documento, se refiere a un compuesto o composición que se conjuga o fusiona directa o indirectamente con un reactivo tal como una sonda de ácido nucleico o un anticuerpo y facilita la detección del reactivo con el que está conjugado o fusionado. El marcador puede ser detectable por sí mismo (por ejemplo, marcadores radioisotópicos o marcadores fluorescentes) o, en el caso de un marcador enzimático, puede catalizar la alteración química de un compuesto o composición sustrato que es detectable.

Un "medicamento" es un fármaco activo para tratar una enfermedad, trastorno y/o afección. En una realización, la enfermedad, trastorno y/o afección es lupus o sus síntomas o efectos secundarios.

La expresión "resistencia aumentada" a un agente terapéutico u opción de tratamiento particular, cuando se usa de acuerdo con la invención, significa respuesta disminuida a una dosis convencional del fármaco o a un protocolo de tratamiento convencional.

La expresión "sensibilidad disminuida" a un agente terapéutico u opción de tratamiento particular, cuando se usa de acuerdo con la invención, significa respuesta disminuida a una dosis convencional del agente o a un protocolo de tratamiento convencional, donde la respuesta disminuida puede compensarse (al menos parcialmente) aumentando la dosis del agente o la intensidad del tratamiento.

La "respuesta del paciente" puede evaluarse usando cualquier criterio de valoración que indique un beneficio para el paciente incluyendo, sin limitación, (1) inhibición, en alguna medida, de la progresión de la enfermedad, incluyendo ralentización y detención completa; (2) reducción en la cantidad de episodios patológicos y/o síntomas; (3) reducción en el tamaño de la lesión; (4) inhibición (es decir, reducción, ralentización o detención completa) de infiltración de células enfermas en órganos periféricos y/o tejidos adyacentes; (5) inhibición (es decir, reducción, ralentización o detención completa) de la propagación de la enfermedad; (6) disminución de la respuesta autoinmunitaria que puede provocar, aunque no necesariamente, la regresión o la ablación de la lesión de la enfermedad; (7) alivio, en alguna medida, de uno o más síntomas asociados con el trastorno; (8) aumento en la duración de presentación sin enfermedad después del tratamiento; y/o (9) mortalidad disminuida en un punto dado de tiempo después del tratamiento.

El término "antagonista" se usa en el sentido más amplio e incluye cualquier molécula que inhiba parcial o completamente o neutralice una actividad biológica de un polipéptido, tal como PRO, o que inhiba parcial o completamente la transcripción o la traducción de un ácido nucleico que codifica el polipéptido. Las moléculas antagonistas a modo de ejemplo incluyen, aunque sin limitación, anticuerpos antagonistas, fragmentos polipeptídicos, oligopéptidos, moléculas orgánicas (incluyendo moléculas pequeñas) y ácidos nucleicos antisentido.

El término "agonista" se usa en el sentido más amplio e incluye cualquier molécula que imite parcial o completamente una actividad biológica de un polipéptido, tal como PRO, o que aumente la transcripción o la traducción de un ácido nucleico que codifica el polipéptido. Las moléculas agonistas a modo de ejemplo, aunque sin limitación, anticuerpos agonistas, fragmentos polipeptídicos, oligopéptidos, moléculas orgánicas (incluyendo moléculas pequeñas), polinucleótidos asociados a PRO, polipéptidos PRO o fusiones PRO-Fc.

Un "agente terapéutico dirigido a PRO o un polinucleótido asociado a PRO" significa cualquier agente que afecte a la expresión y/o la actividad de PRO o un polinucleótido asociado a PRO incluyendo, aunque sin limitación, cualquiera de los agonistas o antagonistas de PRO descritos en este documento, incluyendo dichos agentes terapéuticos que ya son conocidos en la técnica, así como aquellos que se desarrollarán posteriormente.

Un "agente terapéutico contra el lupus", un "agente terapéutico eficaz para tratar el lupus" y variaciones gramaticales de los mismos, como se usan en este documento, se refieren a un agente que cuando se proporciona en una cantidad eficaz los médicos saben que, se demuestra clínicamente que o se espera que proporcionen un beneficio terapéutico en un sujeto que tiene lupus. En una realización, la expresión incluye cualquier agente que esté comercializado por un fabricante o se use de otra manera por médicos con licencia, como un agente clínicamente aceptado que cuando se proporciona en una cantidad eficaz se esperaría que proporcione un efecto terapéutico en un sujeto que tiene lupus. En una realización, un agente terapéutico contra el lupus comprende un fármaco antiinflamatorio no esteroideo (AINE), que incluye ácido acetilsalicílico (por ejemplo, aspirina), ibuprofeno ((Motrin), naproxeno (Naprosyn), indometacina (Indocin), nabumetona (Relafen), tolmetina (Tolectin) y cualquier otra realización que comprenda uno o más ingredientes activos terapéuticamente equivalentes y formulación de los mismos. En una realización, un agente terapéutico contra el lupus comprende acetaminofeno (por ejemplo, Tylenol), corticosteroides o antipalúdicos (por ejemplo, cloroquina, hidroxiclороquina). En una realización, un agente terapéutico contra el lupus comprende un fármaco inmunomodulador (por ejemplo, azatioprina, ciclofosfamida, metotrexato, ciclosporina). En una realización, un agente terapéutico contra el lupus es un agente anti-linfocitos B (por ejemplo, anti-CD20 (por ejemplo, rituximab), anti-CD22), un agente anti-citocinas (por ejemplo, anti-factor de necrosis tumoral α , anti-receptor de interleucina-1 (por ejemplo, anakinra), anti-interleucina 10, anti-receptor de interleucina 6, anti-interferón alfa, anti-estimulador de linfocitos B), un inhibidor de la co-estimulación (por ejemplo, anti-CD154, CTLA4-Ig (por ejemplo, abatacept)), un modulador de la energía de linfocitos B (por ejemplo, LJP 394 (por ejemplo, abetimus)). En una realización, un agente terapéutico contra el lupus comprende tratamiento hormonal (por ejemplo, DHEA) y tratamiento anti-hormonal (por ejemplo, el agente anti-prolactina bromocriptina). En una realización, un agente terapéutico contra el lupus es un agente que proporciona inmunoadsorción, es un factor anti-complemento (por ejemplo, anti-C5a), vacunación de linfocitos T, transfección celular con cadena zeta del receptor de linfocitos T o tratamientos con péptidos (por ejemplo, edratida dirigida a idiotipos anti-ADN).

Un agente terapéutico que tiene "aprobación para su comercialización" o que se ha "aprobado como agente terapéutico" o variaciones gramaticales de estas expresiones, como se usan en este documento, se refieren a un agente (por ejemplo, en forma de una formulación de fármaco, medicamento) que está aprobado, tiene licencia, está registrado o autorizado por una entidad gubernamental relevante (por ejemplo, agencia, departamento, oficina reguladora federal, estatal o local) para venderse por y/o a través y/o en beneficio de una entidad comercial (por ejemplo, una entidad con ánimo de lucro) para el tratamiento de un trastorno particular (por ejemplo, lupus) o una subpoblación de pacientes (por ejemplo, pacientes con nefritis lúpica, pacientes de una etnia, género, estilo de vida, perfil de riesgo de enfermedad particular, etc.). Una entidad gubernamental relevante incluye, por ejemplo, la Food and Drug Administration (FDA), European Medicines Evaluation Agency (EMA) y equivalentes de las mismas.

"Anticuerpos" (Ab) e "inmunoglobulinas" (Ig) se refieren a glucoproteínas que tienen características estructurales similares. Aunque los anticuerpos muestran especificidad de unión a un antígeno específico, las inmunoglobulinas incluyen tanto anticuerpos como otras moléculas de tipo anticuerpo que generalmente carecen de especificidad antigénica. Los polipéptidos del último tipo se producen, por ejemplo, a bajos niveles por el sistema linfático y a niveles aumentados por mielomas.

Los términos "anticuerpo" y "inmunoglobulina" se usan indistintamente en el sentido más amplio e incluyen anticuerpos monoclonales (por ejemplo, anticuerpos monoclonales de longitud completa o intactos), anticuerpos policlonales, anticuerpos monovalentes, anticuerpos multivalentes, anticuerpos multiespecíficos (por ejemplo, anticuerpos biespecíficos siempre que muestren la actividad biológica deseada) y también pueden incluir ciertos fragmentos de anticuerpo (como se describe en mayor detalle en este documento). Un anticuerpo puede ser quimérico, humano, humanizado y/o de afinidad madurada.

La expresión "anticuerpo anti-PRO" o "un anticuerpo que se une a PRO" se refiere a un anticuerpo que es capaz de unirse a PRO con suficiente afinidad, de modo que el anticuerpo es útil como agente de diagnóstico y/o terapéutico para abordar PRO. Preferiblemente, el grado de unión de un anticuerpo anti-PRO a una proteína que no es PRO no relacionada es de menos de aproximadamente un 10 % de la unión del anticuerpo a PRO medida, por ejemplo, por un radioinmunoensayo (RIA). En ciertas realizaciones, un anticuerpo que se une a PRO tiene una constante de disociación (Kd) de $\leq 1 \mu\text{M}$, $\leq 100 \text{ nM}$, $\leq 10 \text{ nM}$, $\leq 1 \text{ nM}$ o $\leq 0,1 \text{ nM}$. En ciertas realizaciones, un anticuerpo anti-PRO se une a un epítipo de PRO que está conservado entre PRO de diferentes especies.

Las expresiones "anticuerpo de longitud completa", "anticuerpo intacto" y "anticuerpo completo" se usan en este documento indistintamente para hacer referencia a un anticuerpo en su forma sustancialmente intacta, no a fragmentos de anticuerpo como se define a continuación. Las expresiones se refieren particularmente a un anticuerpo con cadenas pesadas que contienen la región Fc.

Los "fragmentos de anticuerpo" comprende una parte de un anticuerpo intacto, preferiblemente comprendiendo la región de unión a antígeno del mismo. Los ejemplos de fragmentos de anticuerpo incluyen fragmentos Fab, Fab', F(ab')₂ y Fv; diacuerpos, anticuerpos lineales; moléculas de anticuerpo de cadena sencilla; y anticuerpos multiespecíficos formados por fragmentos de anticuerpo.

La digestión con papaína de los anticuerpos produce dos fragmentos de unión a antígeno idénticos, llamados fragmentos "Fab", cada uno con un único sitio de unión a antígeno, y un fragmento "Fc" residual, cuyo nombre refleja su capacidad de cristalizar fácilmente. El tratamiento con pepsina produce un fragmento F(ab')₂ que tiene dos sitios de combinación de antígeno y aún es capaz de entrecruzarse con el antígeno.

"Fv" es un fragmento mínimo de anticuerpo que contiene un sitio de unión a antígeno completo. En una realización, una especie Fv de dos cadenas consiste en un dímero de un dominio variable de cadena pesada y uno de cadena ligera en estrecha asociación no covalente. De forma colectiva, la seis CDR de un Fv confieren especificidad de unión a antígeno al anticuerpo. Sin embargo, incluso un único dominio variable (o la mitad de un Fv que comprende únicamente tres CDR específicas para un antígeno) tiene la capacidad de reconocer y unirse al antígeno, aunque a una afinidad inferior que el sitio de unión completo.

El fragmento Fab contiene los dominios variables de cadena pesada y cadena ligera y también contiene el dominio constante de la cadena ligera y el primer dominio constante (CH1) de la cadena pesada. Los fragmentos Fab' difieren de los fragmentos Fab por la adición de unos pocos restos en el extremo carboxi del dominio CH1 de cadena pesada que incluye una o más cisteínas de la región de bisagra del anticuerpo. Fab'-SH es la denominación en este documento para Fab' en que los uno o más restos de cisteína de los dominios constantes albergan un grupo tio libre. Los fragmentos de anticuerpo F(ab')₂ originalmente se produjeron como parejas de fragmentos Fab' que tienen cisteínas de la bisagra entre ellos. También se conocen otros acoplamientos químicos de fragmentos de anticuerpo.

Los fragmentos de anticuerpo "Fv de cadena sencilla" o "scFv" comprenden los dominios VH y VL de anticuerpo, donde estos dominios están presentes en una única cadena polipeptídica. Generalmente, el polipéptido scFv comprende además un conector polipeptídico entre los dominios VH y VL que posibilita que el scFv forme la estructura deseada para la unión al antígeno. Para una revisión de scFv véase Pluckthun, en *The Pharmacology of Monoclonal Antibodies*, vol. 113, Rosenberg y Moore eds., Springer-Verlag, Nueva York, pág. 269-315 (1994).

El término "diacuerpos" se refiere a fragmentos de anticuerpo pequeños con dos sitios de unión a antígeno, que son fragmentos que comprenden un dominio variable de cadena pesada (VH) conectado a un dominio variable de cadena ligera (VL) en la misma cadena polipeptídica (VH-VL). Mediante el uso de un conector que es demasiado corto para permitir el emparejamiento entre los dos dominios en la misma cadena, se fuerza a los dominios a emparejarse con los dominios complementarios de otra cadena y crear dos sitios de unión a antígeno. Los diacuerpos pueden ser bivalentes o biespecíficos. Los diacuerpos se describen más completamente en, por ejemplo, el documento EP 404,097; el documento WO93/1161; Hudson et al., (2003) *Nat. Med.* 9:129-134; y Hollinger et al., *Proc. Natl. Acad. Sci. USA* 90: 6444-6448 (1993). Los triacuerpos y los tetrauerpos también se describen en Hudson et al., (2003) *Nat. Med.* 9:129-134.

La expresión "anticuerpo monoclonal" como se usa en este documento se refiere a un anticuerpo obtenido de una población de anticuerpos sustancialmente homogéneos, es decir, los anticuerpos individuales que componen la población son idénticos excepto por posibles mutaciones, por ejemplo, mutaciones de origen natural que pueden estar presentes en cantidades mínimas. Por tanto, el modificador "monoclonal" indica el carácter del anticuerpo que no es una mezcla de anticuerpos concretos. En ciertas realizaciones, dicho anticuerpo monoclonal normalmente incluye un anticuerpo que comprende una secuencia polipeptídica que se une a una diana, donde la secuencia polipeptídica de unión a la diana se obtuvo por un proceso que incluye la selección de una única secuencia polipeptídica de unión a la diana de una pluralidad de secuencias polipeptídicas. Por ejemplo, el proceso de selección puede ser la selección de clon único de una pluralidad de clones, tal como una combinación de clones de hibridoma, clones de fago o clones de ADN recombinante. Debe entenderse que una secuencia de unión a la diana seleccionada puede alterarse adicionalmente, por ejemplo, para mejorar la afinidad por la diana, para humanizar la secuencia de unión a la diana, para mejorar su producción en cultivo celular, para reducir su inmunogenicidad *in vivo*, para crear un anticuerpo multiespecífico, etc., y que un anticuerpo que comprende la secuencia de unión a la diana alterada también es un anticuerpo monoclonal de esta invención. En contraste con las preparaciones de anticuerpos policlonales que normalmente incluyen diferentes anticuerpos dirigidos contra diferentes determinantes (epítomos), cada anticuerpo monoclonal de una preparación de anticuerpos monoclonales está dirigido contra un único determinante en un antígeno. Además de su especificidad, las preparaciones de anticuerpos monoclonales son ventajosas porque normalmente no están contaminadas por otras inmunoglobulinas.

El modificador "monoclonal" indica el carácter del anticuerpo que se obtiene de una población sustancialmente homogénea de anticuerpos, y no debe interpretarse como que requiere producción del anticuerpo por ningún método particular. Por ejemplo, los anticuerpos monoclonales a usarse de acuerdo con la presente invención pueden prepararse por diversas técnicas, incluyendo, por ejemplo, el método de hibridoma (por ejemplo, Kohler *et al.*, Nature, 256: 495 (1975); Harlow *et al.*, *Antibodies: A Laboratory Manual*, (Cold Spring Harbor Laboratory Press, 2ª ed. 1988); Hammerling *et al.*, en: *Monoclonal Antibodies and T-Cell Hybridomas* 563-681 (Elsevier, N.Y., 1981)), métodos de ADN recombinante (véase, por ejemplo, la patente de Estados Unidos n.º 4.816.567), tecnologías de presentación en fagos (véase, por ejemplo, Clackson *et al.*, Nature, 352: 624-628 (1991); Marks *et al.*, J. Mol. Biol. 222: 581-597 (1992); Sidhu *et al.*, J. Mol. Biol. 338(2): 299-310 (2004); Lee *et al.*, J. Mol. Biol. 340(5): 1073-1093 (2004); Fellouse, Proc. Natl. Acad. Sci. USA 101(34): 12467-12472 (2004); y Lee *et al.*, J. Immunol. Methods 284(1-2): 119-132(2004), y tecnologías para producir anticuerpos humanos o de tipo humano en animales que tienen partes de o todos los loci o genes de inmunoglobulina humana que codifican las secuencias de inmunoglobulina humana (véanse, por ejemplo, los documentos WO98/24893; WO96/34096; WO96/33735; WO91/10741; Jakobovits *et al.*, Proc. Natl. Acad. Sci. USA 90: 2551 (1993); Jakobovits *et al.*, Nature 362: 255-258 (1993); Bruggemann *et al.*, Year in Immunol. 7:33 (1993); las patentes de Estados Unidos n.º 5.545.807; 5.545.806; 5.569.825; 5.625.126; 5.633.425; 5.661.016; Marks *et al.*, Bio.Technology 10: 779-783 (1992); Lonberg *et al.*, Nature 368: 856-859 (1994); Morrison, Nature 368: 812-813 (1994); Fishwild *et al.*, Nature Biotechnol. 14: 845-851 (1996); Neuberger, Nature Biotechnol. 14: 826 (1996) y Lonberg y Huszar, Intern. Rev. Immunol. 13: 65-93 (1995).

Los anticuerpos monoclonales de este documento incluyen específicamente anticuerpos "quiméricos" en que una parte de la cadena pesada y/o ligera es idéntica u homóloga a secuencias correspondientes en anticuerpos derivados de una especie particular y que pertenecen a una clase o subclase particular de anticuerpo, mientras que el resto de la cadena o cadenas es idéntico u homólogo a secuencias correspondientes en anticuerpos derivados de otra especie o que pertenecen a otra clase o subclase de anticuerpo, así como fragmentos de dichos anticuerpos, siempre que muestren la actividad biológica deseada (patente de Estados Unidos n.º 4.816.567; y Morrison *et al.*, Proc. Natl. Acad. Sci. USA 81:6855-9855 (1984)).

Las formas "humanizadas" de anticuerpos no humanos (por ejemplo. murinos) son anticuerpos quiméricos que contienen la secuencia mínima derivada de inmunoglobulina no humana. En una realización, un anticuerpo humanizado es una inmunoglobulina humana (anticuerpo destinatario) en que los restos de una región hipervariable del destinatario se remplazan por restos de una región hipervariable de una especie no humana (anticuerpo donante) tal como ratón, rata, conejo o primate no humano que tiene la especificidad, afinidad y/o capacidad deseadas. En algunos casos, los restos de la región flanqueante (FR) de la inmunoglobulina humana se remplazan por los correspondientes restos no humanos. Además, los anticuerpos humanizados pueden comprender restos que no se encuentran en el anticuerpo destinatario o en el anticuerpo donante. Estas modificaciones pueden hacerse para refinar adicionalmente el rendimiento del anticuerpo. En general, un anticuerpo humanizado comprenderá sustancialmente todo de al menos uno, y normalmente dos, dominios variables, en que todos o sustancialmente todos los bucles hipervariables corresponden a los de una inmunoglobulina no humana, y todas o sustancialmente todas las FR son las de una secuencia de inmunoglobulina humana. El anticuerpo humanizado opcionalmente también comprenderá al menos una parte de una región constante de inmunoglobulina (Fc), normalmente la de una inmunoglobulina humana. Para detalles adicionales véase Jones *et al.*, Nature 321:522-525 (1986); Riechmann *et al.*, Nature 332:323-329 (1988); y Presta, Curr. Op. Struct. Biol. 2:593-596 (1992). Véanse también los siguientes artículos de revisión y las referencias citadas en los mismos: Vaswani y Hamilton, Ann. Allergy, Asthma & Immunol. 1:105-115 (1998); Harris, Biochem. Soc. Transactions 23:1035-1038 (1995); Hurlé y Gross, Curr. Op. Biotech. 5:428-433 (1994).

Un "anticuerpo humano" es uno que comprende una secuencia de aminoácidos correspondiente a la de un anticuerpo producido por un ser humano y que se ha preparado usando cualquiera de las técnicas para preparar anticuerpos humanos divulgadas en este documento. Dichas técnicas incluyen el cribado de bibliotecas combinatorias derivadas de seres humanos, tales como bibliotecas de presentación en fagos (véase, por ejemplo, Marks *et al.*, J. Mol. Biol., 222: 581-597 (1991) y Hoogenboom *et al.*, Nucl. Acids Res., 19: 4133-4137 (1991)); el uso de líneas celulares de mieloma humano y de heteromieloma de ratón-humano para la producción de anticuerpos monoclonales humanos (véase, por ejemplo, Kozbor J. Immunol., 133: 3001 (1984); Brodeur *et al.*, *Monoclonal Antibody Production Techniques and Applications*, pág. 55-93 (Marcel Dekker, Inc., Nueva York, 1987); y Boerner *et al.*, J. Immunol., 147: 86 (1991)); y la generación de anticuerpos monoclonales en animales transgénicos (por ejemplo, ratones) que son capaces de producir un repertorio completo de anticuerpos humanos en ausencia de producción endógena de inmunoglobulina (véase, por ejemplo, Jakobovits *et al.*, Proc. Natl. Acad. Sci USA, 90: 2551 (1993); Jakobovits *et al.*, Nature, 362: 255 (1993); Bruggemann *et al.*, Year in Immunol., 7: 33 (1993)). Esta definición de un anticuerpo humano excluye específicamente un anticuerpo humanizado que comprende restos de unión a antígeno de un animal no humano.

Un anticuerpo de "afinidad madurada" es uno con una o más alteraciones en una o más CDR del mismo que provoca una mejora en la afinidad del anticuerpo por el antígeno, en comparación con un anticuerpo precursor que no posee esas alteraciones. En una realización, un anticuerpo de afinidad madurada tiene afinidades nanomolares o incluso picomolares por el antígeno diana. Los anticuerpos de afinidad madurada se producen por procedimientos conocidos en la técnica. Marks *et al.* Bio/Technology 10:779-783 (1992) describe la maduración de afinidad por

reordenamiento del dominio VH y VL. La mutagénesis aleatoria de restos HVR y/o flanqueantes se describe por: Barbas *et al.* Proc Nat. Acad. Sci. USA 91:3809-3813 (1994); Schier *et al.* Gene 169:147-155 (1995); Yelton *et al.* J. Immunol. 155:1994-2004 (1995); Jackson *et al.*, J. Immunol. 154(7):3310-9 (1995); y Hawkins *et al.*, J. Mol. Biol. 226:889-896 (1992).

Un "anticuerpo de bloqueo" o un "anticuerpo antagonista" es uno que inhibe o reduce una actividad biológica del antígeno al que se une. Ciertos anticuerpos de bloqueo o anticuerpos antagonistas inhiben parcial o completamente la actividad biológica del antígeno.

Una "molécula pequeña" o "molécula orgánica pequeña" se define en este documento como una molécula orgánica que tiene un peso molecular por debajo de aproximadamente 500 Da.

Un "oligopéptido de unión a PRO" o un "oligopéptido que se une a PRO" es un oligopéptido que es capaz de unirse a PRO con suficiente afinidad de modo que el oligopéptido sea útil como agente de diagnóstico y/o terapéutico para abordar PRO. En ciertas realizaciones, el grado de unión de un oligopéptido de unión a PRO a una proteína que no es PRO no relacionada es de menos de aproximadamente un 10 % de la unión a PRO del oligopéptido de unión a PRO medida, por ejemplo, por un ensayo de resonancia de plasmón superficial. En ciertas realizaciones el oligopéptido de unión a PRO tiene una constante de disociación (K_d) de $\leq 1 \mu\text{M}$, $\leq 100 \text{ nM}$, $\leq 10 \text{ nM}$, $\leq 1 \text{ nM}$ o $\leq 0,1 \text{ nM}$.

Una "molécula orgánica de unión a PRO" o "una molécula orgánica que se une a PRO" es una molécula orgánica diferente a un oligopéptido o anticuerpo como se define en este documento, que es capaz de unirse a PRO con suficiente afinidad de modo que la molécula orgánica sea útil como agente de diagnóstico y/o terapéutico para abordar PRO. En ciertas realizaciones, el grado de unión de una molécula orgánica de unión a PRO a una proteína que no es PRO no relacionada es de menos de aproximadamente un 10 % de la unión a PRO de la molécula orgánica de unión a PRO medida, por ejemplo, por un ensayo de resonancia de plasmón superficial. En ciertas realizaciones, una molécula orgánica de unión a PRO tiene una constante de disociación (K_d) de $\leq 1 \mu\text{M}$, $\leq 100 \text{ nM}$, $\leq 10 \text{ nM}$, $\leq 1 \text{ nM}$ o $\leq 0,1 \text{ nM}$.

La constante de disociación (K_d) de cualquier molécula que se une a un polipéptido diana puede medirse convenientemente usando un ensayo de resonancia de plasmón superficial. Dichos ensayos pueden emplear un BIAcore™-2000 o un BIAcore™-3000 (BIAcore, Inc., Piscataway, NJ) a 25 °C con chips CM5 de polipéptido diana inmovilizado a ~ 10 unidades de respuesta (UR). En resumen, se activan chips biodetectores de dextrano carboximetilado (CM5, BIAcore Inc.) con clorhidrato de *N*-etil-*N'*-(3-dimetilaminopropil)-carbodiimida (EDC) y *N*-hidroxisuccinimida (NHS) de acuerdo con las instrucciones del proveedor. El polipéptido diana se diluye con acetato de sodio 10 nM, pH 4,8, hasta 5 µg/ml (~ 0,2 µM) antes de la inyección a un caudal de 5 µl/minuto para conseguir aproximadamente 10 unidades de respuesta (UR) de la proteína acoplada. Después de la inyección del polipéptido diana, se inyecta etanolamina 1 M para bloquear los grupos que no han reaccionado. Para las mediciones de la cinética, se inyectan diluciones en serie de factor dos de la molécula de unión (0,78 nM a 500 nM) en PBS con Tween 20 (PBST) al 0,05 % a 25 °C a un caudal de aproximadamente 25 µl/min. Las tasas de asociación (k_{on}) y las tasas de disociación (k_{off}) se calculan usando un modelo simple de unión de Langmuir de uno a uno (BIAcore Evaluation Software versión 3.2) por ajuste simultáneo de los sensogramas de asociación y disociación. La constante de disociación en equilibrio (K_d) se calcula como la relación k_{off}/k_{on} . Véase, por ejemplo, Chen, Y., *et al.*, (1999) J. Mol. Biol. 293:865-881. Si la tasa de asociación de un anticuerpo excede $10^6 \text{ M}^{-1} \text{ s}^{-1}$ por el ensayo de resonancia de plasmón superficial anterior, entonces la tasa de asociación puede determinarse usando la técnica de inactivación fluorescente que mide el aumento o disminución en la intensidad de emisión de fluorescencia (excitación = 295 nm; emisión = 340 nm, 16 nm de paso de banda) a 25 °C de un anticuerpo 20 nM (forma Fab) en PBS, pH 7,2, en presencia de concentraciones crecientes de antígeno medidas en un espectrómetro, tal como un espectrómetro equipado con detención de flujo (Aviv Instruments) o un espectrómetro SLM-Aminco de la serie 8000 (ThermoSpectronic) con una cubeta agitada.

Un "liposoma" es una vesícula pequeña compuesta de diversos tipos de lípidos, fosfolípidos y/o tensioactivo, que es útil para el suministro de un agente, por ejemplo, un fármaco a un mamífero. Los componentes del liposoma se disponen habitualmente en una formación de bicapa, similar a la disposición de los lípidos de las membranas biológicas.

La palabra "marcador" cuando se usa en este documento se refiere a un compuesto o composición detectable. El marcador puede ser detectable por sí mismo (por ejemplo, marcadores radioisotópicos o marcadores fluorescentes) o, en el caso de un marcador enzimático, puede catalizar la alteración química de un compuesto o composición sustrato que produce un producto detectable. Los radionúclidos que pueden servir como marcadores detectables incluyen, por ejemplo, I-131, I-123, I-125, Y-90, Re-188, Re-186, At-211, Cu-67, Bi-212 y Pd-109.

Una molécula biológica "aislada", tal como un ácido nucleico, polipéptido o anticuerpo, es una que se ha identificado y separado y/o recuperado de al menos un componente de su entorno natural.

Una referencia a "aproximadamente" un valor o parámetro en este documento incluye (y describe) realizaciones que

se refieren al valor o parámetro *per se*. Por ejemplo, una descripción que se refiera a "aproximadamente X" incluye la descripción de "X".

Se entiende que los aspectos y realizaciones de la realización descritos en este documento incluyen "que consiste en" y/o "que consiste esencialmente en" aspectos y realizaciones.

II. Técnicas generales para realizar composiciones y métodos de la invención

En este documento se proporcionan variaciones nucleotídicas asociadas con lupus. Estas variaciones proporcionan biomarcadores para lupus, y/o predisponen o contribuyen al desarrollo, persistencia y/o progresión del lupus. Por consiguiente, la invención divulgada en este documento es útil en diversos entornos, por ejemplo, en métodos y composiciones que se refieren al diagnóstico y el tratamiento del lupus.

Detección de variaciones genéticas

El ácido nucleico, de acuerdo con cualquier de los métodos anteriores, puede ser ADN genómico; ARN transcrito de ADN genómico; o ADNc generado de ARN. El ácido nucleico puede obtenerse de un vertebrado, por ejemplo, un mamífero. Un ácido nucleico se dice que "deriva de" una fuente particular si se obtiene directamente de esa fuente o si es una copia de un ácido nucleico encontrado en esa fuente.

El ácido nucleico incluye copias del ácido nucleico, por ejemplo, copias que se producen por amplificación. La amplificación puede ser deseable en ciertos, por ejemplo, para obtener una cantidad deseada de material para detectar variaciones. Por ejemplo, un polinucleótido asociado a PRO o parte del mismo puede amplificarse a partir de material de ácido nucleico. Los amplicones entonces pueden someterse a un método de detección de variaciones, tal como los descritos a continuación, para determinar si una variación está presente en el amplicón.

Pueden detectarse variaciones por ciertos métodos conocidos para los expertos en la materia. Dichos métodos incluyen, aunque sin limitación, secuenciación de ADN; ensayos de prolongación de cebador, incluyendo ensayos de incorporación de nucleótidos específica de alelo y ensayos de prolongación de cebador específico de alelo (por ejemplo, PCR específica de alelo, reacción en cadena de ligamiento (LCR) específica de alelo y gap-LCR); ensayos de hibridación de oligonucleótidos específicos de alelo (por ejemplo, ensayos de ligamiento de oligonucleótidos); ensayos de protección de escisión en que se usa protección contra agentes de escisión para detectar bases emparejadas incorrectamente en dúplex de ácido nucleico; análisis de unión de la proteína MutS; análisis electroforético que compara movilidad de moléculas de ácido nucleico variantes y de tipo silvestre; electroforesis en gel con gradiente desnaturizante (DGGE), como en, por ejemplo, Myers *et al.* (1985) *Nature* 313:495; análisis de escisión por RNasa en pares de bases emparejadas incorrectamente; análisis de escisión química o enzimática de ADN heterodúplex; espectrometría de masas (por ejemplo, MALDI-TOF); análisis genético de un bit (GBA); ensayos de 5' nucleasa (por ejemplo, TaqMan®); y ensayos que emplean balizas moleculares. Algunos de estos métodos se analizan en más detalle a continuación.

La detección de variaciones en ácidos nucleicos diana puede conseguirse por clonación molecular y secuenciación de los ácidos nucleicos diana usando técnicas bien conocidas en la técnica. Como alternativa, pueden usarse técnicas de amplificación tales como la reacción en cadena de la polimerasa (PCR) para amplificar secuencias de ácido nucleico diana directamente de una preparación de ADN genómico de tejido tumoral. La secuencia de ácido nucleico de las secuencias amplificadas entonces puede determinarse y pueden identificarse las variaciones a partir de las mismas. Las técnicas de amplificación son bien conocidas en la técnica, por ejemplo, la reacción en cadena de la polimerasa se describe en Saiki *et al.*, *Science* 239:487, 1988; patentes de Estados Unidos n.º 4.683.203 y 4.683.195.

La reacción en cadena de la ligasa, que se conoce en la técnica, también puede usarse para amplificar secuencias de ácido nucleico diana. Véase, por ejemplo, Wu *et al.*, *Genomics* 4:560-569 (1989). Además, también puede usarse una técnica conocida como PCR específica de alelo para detectar variaciones (por ejemplo, sustituciones). Véase, por ejemplo, Ruano y Kidd (1989) *Nucleic Acids Research* 17:8392; McClay *et al.* (2002) *Analytical Biochem.* 301:200-206. En ciertas realizaciones de esta técnica, se usa un cebador específico de alelo donde el nucleótido 3' terminal del cebador es complementario a (es decir, es capaz de formar pares de bases específicamente con) una variación particular en el ácido nucleico diana. Si la variación particular no está presente, no se observa un producto de amplificación. También puede usarse el sistema de mutación refractario a amplificación (ARMS) para detectar variaciones (por ejemplo, sustituciones). ARMS se describe, por ejemplo, en la publicación de solicitud de patente europea n.º 0332435 y en Newton *et al.*, *Nucleic Acids Research*, 17:7, 1989.

Otros métodos útiles para detectar variaciones (por ejemplo, sustituciones) incluyen, aunque sin limitación, (1) ensayos de incorporación de nucleótidos específica de alelo, tal como ensayos de prolongación de una única base (véase, por ejemplo, Chen *et al.* (2000) *Genome Res.* 10:549-557; Fan *et al.* (2000) *Genome Res.* 10:853-860; Pastinen *et al.* (1997) *Genome Res.* 7:606-614; y Ye *et al.* (2001) *Hum. Mut.* 17:305-316); (2) ensayos de prolongación de cebador específico de alelo (véase, por ejemplo, Ye *et al.* (2001) *Hum. Mut.* 17:305-316; y Shen *et al.* *Genetic Engineering News*, vol. 23, 15 de marzo de 2003), incluyendo PCR específica de alelo; (3) ensayos de 5'

nucleasa (véase, por ejemplo, De La Vega *et al.* (2002) *BioTechniques* 32:S48-S54 (que describe el ensayo TaqMan®); Ranade *et al.* (2001) *Genome Res.* 11:1262-1268; y Shi (2001) *Clin. Chem.* 47:164-172); (4) ensayos que emplean balizas moleculares (véase, por ejemplo, Tyagi *et al.* (1998) *Nature Biotech.* 16:49-53; y Mhlanga *et al.* (2001) *Methods* 25:463-71); y (5) ensayos de ligamiento de oligonucleótidos (véase, por ejemplo, Grossman *et al.* (1994) *Nuc. Acids Res.* 22:4527-4534; publicación de solicitud de patente n.º US 2003/0119004 A1; publicación internacional PCT n.º WO 01/92579 A2; y patente de Estados Unidos n.º 6.027.889).

Las variaciones también pueden detectarse por métodos de detección de emparejamientos incorrectos. Los emparejamientos incorrectos son dúplex de ácido nucleico hibridados que no son un 100 % complementarios. La ausencia de complementariedad total puede deberse a eliminaciones, inserciones, inversiones o sustituciones. Un ejemplo de un método de detección de emparejamientos incorrectos es el ensayo de detección de reparación de emparejamiento incorrectos (MRD) descrito, por ejemplo, en Faham *et al.*, *Proc. Natl Acad. Sci. USA* 102:14717-14722 (2005) y Faham *et al.*, *Hum. Mol. Genet.* 10:1657-1664 (2001). Otro ejemplo de una técnica de escisión de emparejamiento incorrecto es el método de protección de RNasa, que se describe en detalle en Winter *et al.*, *Proc. Natl. Acad. Sci. USA*, 82:7575, 1985, y Myers *et al.*, *Science* 230:1242, 1985. Por ejemplo, un método de la invención puede implicar el uso de una ribosonda marcada que es complementaria al ácido nucleico diana de tipo silvestre humano. La ribosonda y el ácido nucleico diana derivado de la muestra tisular se atemperan (hibridan) juntos y posteriormente se digieren con la enzima RNasa A que es capaz de detectar algunos emparejamientos incorrectos en una estructura de ARN dúplex. Si se detecta un emparejamiento incorrecto por RNasa A, escinde en el sitio del emparejamiento incorrecto. Por tanto, cuando la preparación de ARN hibridado se separa en una matriz de gel electroforético, si se ha detectado un emparejamiento incorrecto y se ha escindido por RNasa A, se observará un producto de ARN que es más pequeño que el ARN dúplex de longitud completa para la ribosonda y el ARNm o ADN. La ribosonda no tiene que ser la longitud completa del ácido nucleico diana, sino que puede ser una parte del ácido nucleico diana, con la condición de que abarque la posición sospechosa de tener una variación.

De una manera similar, pueden usarse sondas de ADN para detectar emparejamientos incorrectos, por ejemplo, a través de escisión enzimática o química. Véase, por ejemplo, Cotton *et al.*, *Proc. Natl. Acad. Sci. USA*, 85:4397, 1988; y Shenk *et al.*, *Proc. Natl. Acad. Sci. USA*, 72:989, 1975. Como alternativa, pueden detectarse emparejamientos incorrectos por desplazamientos en la movilidad electroforética de dúplex emparejados incorrectamente respecto a dúplex emparejados. Véase, por ejemplo, Cariello, *Human Genetics*, 42:726, 1988. Con las ribosondas o las sondas de ADN, el ácido nucleico diana sospechoso de comprender una variación puede amplificarse antes de la hibridación. También pueden detectarse cambios en el ácido nucleico diana usando hibridación de Southern, especialmente si los cambios son reordenamientos importantes, tales como eliminaciones e inserciones.

Las sondas de polimorfismos de longitud de fragmento de restricción (RFLP) para el ácido nucleico o los genes marcadores adyacentes pueden usarse para detectar variaciones, por ejemplo, inserciones o eliminaciones. Las inserciones y eliminaciones también pueden detectarse por clonación, secuencias y amplificación de un ácido nucleico diana. El análisis de polimorfismos de conformación monocatenaria (SSCP) también puede usarse para detectar variaciones de cambio de bases de un alelo. Véase, por ejemplo, Orita *et al.*, *Proc. Natl. Acad. Sci. USA* 86:2766-2770, 1989, y *Genomics*, 5:874-879, 1989.

Composiciones de la invención

La invención proporciona composiciones de polinucleótidos aislados que comprenden un polinucleótido o fragmento del mismo que comprende un SNP. En una realización, el polinucleótido es un polinucleótido asociado a PRO.

En particular, la invención proporciona composiciones que comprenden conjuntos únicos y/o combinaciones de SNP que pueden usarse como un perfil o característica genética indicativa de un sujeto en riesgo de desarrollar lupus, o indicativa de la enfermedad o síntoma o afección del mismo. Los polimorfismos divulgados en este documento son útiles como biomarcadores para evaluar el riesgo de desarrollar lupus, así como para dianas para el diseño de reactivos de diagnóstico. En algunas realizaciones, el SNP no está asociado con un gen. En otras realizaciones el SNP está asociado con un gen y puede estar localizado en una región intergénica o intragénica y, más particularmente, puede estar localizado en una región codificante o no codificante. Los genes asociados con un SNP de la presente invención pueden estar asociados con un gen desconocido o pueden estar asociados con un gen conocido, por ejemplo, ITGAM o BLK.

Los SNP identificados en este documento proporcionan dianas para el desarrollo de agentes terapéuticos para su uso en el diagnóstico y tratamiento de pacientes con lupus identificado genéticamente, incluyendo el diagnóstico y el tratamiento dirigido de subpoblaciones de pacientes con lupus que muestran una característica genética distinta que comprende uno o más de los SNP de la presente invención. Por ejemplo, en una realización, los genes que contienen las variaciones genéticas identificadas en este documento, y el ácido nucleico (por ejemplo, ADN o ARN) asociado con estos genes y proteínas codificadas por estos genes, pueden usarse como dianas para el desarrollo de agentes terapéuticos (por ejemplo, compuestos de molécula pequeña, anticuerpos, agentes antisentido/iARN, etc.) o usarse directamente como agentes terapéuticos (por ejemplo, proteínas terapéutica, etc.) para el tratamiento del lupus.

Por consiguiente, en un aspecto, la divulgación proporciona un conjunto de uno o más SNP que forman una característica genética única para evaluar el riesgo de desarrollar lupus. En un aspecto, la característica genética única comprende aproximadamente 1-10, 10-20, 20-30, 30-40 o 40-50 SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

En un aspecto, la característica genética única comprende 1 o más SNP, 3 o más SNP, 3 o más SNP, 4 o más SNP, 5 o más SNP, 6 o más SNP, 7 o más SNP, 8 o más SNP, 9 o más SNP, 10 o más SNP, 11 o más SNP, 12 o más SNP, 13 o más SNP, 14 o más SNP, 15 o más SNP, 16 o más SNP, 17 o más SNP, 18 o más SNP, 19 o más SNP o 20 o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10. En un aspecto, los SNP de la característica genética se seleccionan de la tabla 6. En otro aspecto los SNP se seleccionan del grupo que consiste en rs9888739, rs13277113, rs7574865, rs2269368, rs6889239, rs2391592 y rs21177770. En otro aspecto, los SNP se seleccionan del grupo que consiste en rs2187668, rs10488631, rs7574865, rs9888739, rs13277113, rs2431697, rs6568431, rs10489265, rs2476601, rs2269368, rs1801274, rs4963128, rs5754217, rs6445975, rs3129860, rs10516487, rs6889239, rs2391592 y rs21177770.

En otra realización, la divulgación proporciona un polinucleótido aislado (por ejemplo, ADN o ARN) o fragmento del mismo que es de al menos aproximadamente 10 nucleótidos de longitud, donde el polinucleótido o fragmento del mismo comprende: (a) una variación genética en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) seleccionado de los SNP expuestos en las figuras 1-17 y las tablas 1-10, o (b) el complemento de (a). En una realización, el polinucleótido aislado es ADN genómico que comprende un polimorfismo de un único nucleótido (SNP) seleccionado de los SNP expuestos en cualquiera de las figuras 1-17 y las tablas 1-10. En otra realización, el polinucleótido aislado es un ARN que comprende un polimorfismo de un único nucleótido (SNP) seleccionado de cualquiera de los expuestos en las figuras 1-17 y las tablas 1-10.

En una realización de la divulgación, la variación genética en la región anterior del sitio de inicio de la transcripción de la tirosina cinasa linfocitaria B (BLK) y C8orf13 (cromosoma 8p23.1) está asociada con riesgo de enfermedad en la serie de casos/controles tanto estadounidenses como suecos (rs13277113, OR = 1,39, meta P = 1×10^{-10}) y también con niveles alterados de ARNm en líneas de linfocitos B. En otra realización, las variantes en la región de integrina alfa M (ITGAM) e integrina Alfa X (ITGAX) (cromosoma 16p11.2) están asociadas con SLE en la muestra combinada (rs11574637, OR = 1,33, meta P = 3×10^{-11}). En una exploración exhaustiva de todo el genoma en SLE, los autores de la presente invención han identificado y después han confirmado a través de replicación, dos nuevos loci genéticos: a) un alelo de región promotora que se correlaciona con expresión reducida de BLK y expresión aumentada de C8orf13 y b) SNP (o variantes) dentro de la región ITGAM/ITGAX que están en fuerte desequilibrio del ligamiento con dos alelos no sinónimos comunes de ITGAM.

En una realización, el polinucleótido o fragmento del mismo es de al menos aproximadamente 10 nucleótidos de longitud, como alternativa de al menos aproximadamente 15 nucleótidos de longitud, como alternativa de al menos aproximadamente 20 nucleótidos de longitud, alternativamente de al menos aproximadamente 30 nucleótidos de longitud, alternativamente de al menos aproximadamente 40 nucleótidos de longitud, alternativamente de al menos aproximadamente 50 nucleótidos de longitud, alternativamente de al menos aproximadamente 60 nucleótidos de longitud, alternativamente de al menos aproximadamente 70 nucleótidos de longitud, alternativamente de al menos aproximadamente 80 nucleótidos de longitud, alternativamente de al menos aproximadamente 90 nucleótidos de longitud, alternativamente de al menos aproximadamente 100 nucleótidos de longitud, alternativamente de al menos aproximadamente 110 nucleótidos de longitud, alternativamente de al menos aproximadamente 120 nucleótidos de longitud, alternativamente de al menos aproximadamente 130 nucleótidos de longitud, alternativamente de al menos aproximadamente 140 nucleótidos de longitud, alternativamente de al menos aproximadamente 150 nucleótidos de longitud, alternativamente de al menos aproximadamente 160 nucleótidos de longitud, alternativamente de al menos aproximadamente 170 nucleótidos de longitud, alternativamente de al menos aproximadamente 180 nucleótidos de longitud, alternativamente de al menos aproximadamente 190 nucleótidos de longitud, alternativamente de al menos aproximadamente 200 nucleótidos de longitud, alternativamente de al menos aproximadamente 250 nucleótidos de longitud, alternativamente de al menos aproximadamente 300 nucleótidos de longitud, alternativamente de al menos aproximadamente 350 nucleótidos de longitud, alternativamente de al menos aproximadamente 400 nucleótidos de longitud, alternativamente de al menos aproximadamente 450 nucleótidos de longitud, alternativamente de al menos aproximadamente 500 nucleótidos de longitud, alternativamente de al menos aproximadamente 600 nucleótidos de longitud, alternativamente de al menos aproximadamente 700 nucleótidos de longitud, alternativamente de al menos aproximadamente 800 nucleótidos de longitud, alternativamente de al menos aproximadamente 900 nucleótidos de longitud, alternativamente de al menos aproximadamente 1000 nucleótidos de longitud y como alternativa de aproximadamente la longitud completa de la secuencia codificante. En cualquiera de estas realizaciones, el fragmento o polinucleótido de longitud completa también puede incluir parte o toda la región flanqueante de origen natural de un SNP. En este contexto, el término "aproximadamente" significa la longitud de la secuencia de nucleótidos de referencia más o menos un 10 % de esa longitud de referencia.

En otra realización, la secuencia del polinucleótido comprende una variación genética dentro de una región de desequilibrio del ligamiento, por ejemplo, como se expone en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una

realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En otra realización, se proporciona el complemento de cualquiera de los polinucleótidos anteriores. En otra realización, se proporciona un PRO codificado por cualquiera de los polinucleótidos anteriores.

En una realización, el polinucleótido aislado proporcionado en este documento está marcado de forma detectable, por ejemplo, con un radioisótopo, un agente fluorescente o un agente cromogénico. En otra realización, un polinucleótido aislado es un cebador. En otra realización, un polinucleótido aislado es un oligonucleótido, por ejemplo, un oligonucleótido específico de alelo. En otra realización, un oligonucleótido puede ser, por ejemplo, de 7-60 nucleótidos de longitud, 9-45 nucleótidos de longitud, 15-30 nucleótidos de longitud o 18-25 nucleótidos de longitud. En otra realización, un oligonucleótido puede ser, por ejemplo, PNA, morfolino-fosforamidatos, LNA o 2'-alcoialcoxi. En este documento se proporcionan oligonucleótidos que son útiles, por ejemplo, como sondas de hibridación para la detección de variaciones genéticas.

En una realización, la divulgación proporciona una composición que comprende una pluralidad de polinucleótidos capaces de hibridar específicamente con al menos 1, 2, 3, 4 o 5 polinucleótidos asociados a PRO, comprendiendo cada polinucleótido asociado a PRO una variación genética en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10, o complementos de dichos polinucleótidos asociados a PRO. En una realización, los polinucleótidos se proporcionan como una matriz, chip génico o conjunto de genes (por ejemplo, un conjunto de genes o fragmentos de los mismos, proporcionados por separado o como una mezcla). En otra realización, se proporciona un oligonucleótido específico de alelo que hibrida con una región de un polinucleótido asociado a PRO que comprende una variación genética (por ejemplo, una sustitución). En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. El oligonucleótido específico de alelo, cuando se hibrida con una región del polinucleótido asociado a PRO, comprende un nucleótido que forma pares de bases con la variación genética. En otra realización, se proporciona el complemento de un oligonucleótido específico de alelo. En otra realización, se proporciona una micromatriz que comprende un oligonucleótido específico de alelo o su complemento. En otra realización, un oligonucleótido específico de alelo o su complemento es un cebador específico de alelo. En una realización, el oligonucleótido específico de alelo comprende una variación genética en una secuencia polinucleotídica asociada a PRO, donde el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En otra realización, se proporciona el complemento de cualquiera de los polinucleótidos anteriores.

Un oligonucleótido específico de alelo puede usarse junto con un oligonucleótido de control que es idéntico al oligonucleótido específico de alelo, excepto porque el nucleótido que forma pares específicamente con la variación genética está remplazado con un oligonucleótido que forma pares específicamente con el nucleótido correspondiente presente en el polinucleótido asociado a PRO de tipo silvestre. dichos oligonucleótidos pueden usarse en ensayos de unión competitiva en condiciones de hibridación para permitir que los oligonucleótidos distingan entre un polinucleótido asociado a PRO que comprende una variación genética y un polinucleótido asociado a PRO que comprende el nucleótido correspondiente de tipo silvestre.

Usando métodos rutinarios basados en, por ejemplo, la longitud y la composición de bases de los oligonucleótidos, los expertos en la materia pueden llegar a las condiciones de hibridación adecuadas en las que (a) un oligonucleótido específico de alelo se unirá preferentemente a un polinucleótido asociado a PRO que comprende una variación genética respecto a un polinucleótido asociado a PRO de tipo silvestre, y (b) el oligonucleótido de control se unirá preferentemente a un polinucleótido asociado a PRO de tipo silvestre respecto a un polinucleótido asociado a PRO que comprende una variación genética. Las condiciones a modo de ejemplo incluyen condiciones de alta rigurosidad, por ejemplo, condiciones de hibridación de EDTA en fosfato salino convencional (SSPE) 5x y NaDodSO₄ (SDS) al 0,5 % a 55 °C, seguido por lavado con SSPE 2x y SDS al 0,1 % a 55 °C o a temperatura ambiente. En otra realización, se proporciona un agente de unión que se une preferentemente a un PRO que comprende una variación de aminoácido, respecto a un PRO de tipo silvestre. En una realización, la variación de aminoácido es cualquiera producida por una variación genética en una posición nucleotídica correspondiente a un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10 (incluyendo, por ejemplo, cualquier SNP específico en cualquiera de estas figuras o tablas). En otra realización, el agente de unión es un anticuerpo.

Métodos de uso

La invención también proporciona diversas composiciones adecuadas para su uso en la realización de los métodos de la invención. En una realización, la divulgación comprende al menos una molécula de ácido nucleico útil para detectar una o más variaciones genéticas como se divulga en las figuras 1-17 y las tablas 1-10. Dicha molécula de ácido nucleico puede usarse en los métodos de la presente invención, por ejemplo, para la detección, ensayo y tratamiento de lupus. En algunas realizaciones, la molécula de ácido nucleico está adherida a un sustrato sólido

como se describe en este documento.

En otra realización, la divulgación proporciona matrices que pueden usarse en los métodos de la presente invención. En una realización, una matriz de la divulgación comprende moléculas de ácido nucleico individuales o colecciones de moléculas de ácido nucleico útiles para detectar una o más variaciones genéticas. Por ejemplo, una matriz de la invención puede comprender una serie de oligonucleótidos específicos de alelo individuales situados de forma concreta o conjuntos de oligonucleótidos específicos de alelo. Varias técnicas son bien conocidas en la técnica para adherir ácidos nucleicos a un sustrato sólido tal como un portaobjetos de vidrio. Un método es incorporar bases modificadas o análogos que contengan un resto reactivo que sea capaz de adherirse a un sustrato sólido, tal como un grupo amina, un derivado de un grupo amina u otro grupo con una carga positiva, en las moléculas de ácido nucleico que se sintetizan. El producto sintetizado entonces está en contacto con un sustrato sólido, tal como un portaobjetos de vidrio recubierto con un aldehído u otro grupo reactivo. El aldehído u otro grupo reactivo formará un enlace covalente con el resto reactivo en el producto amplificado, que quedará covalentemente unido al portaobjetos de vidrio. otros métodos, tales como los que usando química superficial de aminopropil silicano también son conocidos en la técnica.

Una muestra biológica, de acuerdo con cualquiera de los métodos anteriores, puede obtenerse usando ciertos métodos conocidos para los expertos en la materia. Las muestras biológicas pueden obtenerse de animales vertebrados y, en particular, mamíferos. A menudo se usa biopsia tisular para obtener un trozo representativo de tejido tumoral. Como alternativa, pueden obtenerse células tumorales indirectamente en forma de tejidos o líquidos que se sabe o se cree que contienen las células tumorales de interés. Por ejemplo, pueden obtenerse muestras de lesiones de cáncer broncopulmonar por resección, broncoscopia, aspiración con aguja fina, cepillados bronquiales o de esputo, líquido pleural o sangre. Las variaciones en los ácidos nucleicos diana (o polipéptidos codificados) pueden detectarse a partir de una muestra de tumor o de otras muestras corporales tales como orina, esputo o suero. (Las células cancerosas se desechan de tumores y aparecen en dichas muestras corporales). Cribando puede dichas muestras corporales, conseguirse un diagnóstico fácil y simple para enfermedades tales como cáncer. Además, el progreso del tratamiento puede controlarse más fácilmente ensayando dichas muestras corporales para variaciones en los ácidos nucleicos diana (o polipéptidos codificados). Adicionalmente, se conocen en la técnica métodos para enriquecer una preparación tisular para las células tumorales. Por ejemplo, el tejido puede aislarse de secciones en parafina o de criostato. Las células cancerosas también pueden separarse de células normales por citometría de flujo o microdissección de captura láser.

Posterior a la determinación de que un sujeto, o la muestra tisular o celular comprende una variación genética divulgada en este documento, se contempla que puede administrarse una cantidad eficaz de un agente terapéutico contra el lupus apropiado al sujeto para tratar la afección de lupus en el sujeto. El diagnóstico en mamíferos de las diversas afecciones patológicas descritas en este documento puede hacerse por el experto en la materia. Hay técnicas de diagnóstico disponibles en la técnica que permiten, por ejemplo, el diagnóstico o detección de lupus en un mamífero.

Un agente terapéutico contra el lupus puede administrarse de acuerdo con métodos conocidos, tales como administración intravenosa, como un bolo o por infusión continua durante un periodo de tiempo, por vía intramuscular, intraperitoneal, intracerebromedular, subcutánea, intraarticular, intrasinobial, intratecal, oral, tópica o por inhalación. Opcionalmente, la administración puede realizarse mediante infusión con mini bomba usando diversos dispositivos disponibles en el mercado.

Las dosificaciones y posología eficaces para la administración de agentes terapéuticos contra el lupus pueden determinarse de forma empírica, y hacer dichas determinaciones pertenecen a las habilidades de la materia. Pueden emplearse dosificaciones únicas o múltiples. Por ejemplo, una dosificación o cantidad eficaz de inhibidor de interferón usado en solitario puede variar de aproximadamente 1 mg/kg a aproximadamente 100 mg/kg de peso corporal o más por día. Puede realizarse la escala de las dosificaciones entre especies de una manera conocida en la técnica, por ejemplo, como se divulga en Mordenti et al., Pharmaceut. Res., 8:1351 (1991).

Cuando se emplea administración *in vivo* de un agente terapéutico contra el lupus, las cantidades de dosificación normales pueden variar de aproximadamente 10 ng/kg hasta 100 mg/kg de peso corporal del mamífero o más por día, preferiblemente, de aproximadamente 1 µg/kg/día a 10 mg/kg/día, dependiendo de la vía de administración. Se proporcionan directrices en cuanto a las dosificaciones y métodos particulares de suministro en la bibliografía; véanse, por ejemplo, las patentes de Estados Unidos n.º 4.657.760; 5.206.344; o 5.225.212. Se prevé que diferentes formulaciones serán eficaces para diferentes compuestos de tratamiento y diferentes trastornos, que la administración dirigida a un órgano o tejido, por ejemplo, puede necesitar el suministro de una manera diferente a la de otro órgano o tejido.

Se contempla que pueden emplearse otros tratamientos adicionales en los métodos. El uno o más tratamientos diferentes pueden incluir, aunque sin limitación, la administración de esteroides u otros regímenes de tratamiento habituales para el trastorno en cuestión. Se contempla que dichos otros tratamientos pueden emplearse como un agente diferente de, por ejemplo, un agente terapéutico contra el lupus dirigido.

La invención también proporciona métodos de detección de la presencia de lupus, detectando una variación en un PRO o polinucleótido asociado a PRO derivado de una muestra biológica. En una realización, la muestra biológica se obtiene de un mamífero sospechoso de tener lupus.

5 La divulgación también proporciona métodos de determinación del genotipo de una muestra biológica detectando si una variación genética está presente en un polinucleótido asociado a PRO derivado de la muestra biológica. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En dicha realización, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En otra realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En otra realización, se sabe que la muestra biológica comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación. En otra realización, la muestra biológica es una línea celular, por ejemplo, una línea celular primaria o inmortalizada. En dicha realización, el genotipado proporciona una base para clasificar o subclasificar la enfermedad.

20 La invención también proporciona métodos de identificación de células en una muestra biológica de un mamífero que se sabe que comprenden o son sospechosas de comprender un PRO o polinucleótido asociado a PRO que comprende una variación, detectando la variación en un PRO o polinucleótido asociado a PRO derivado de las células de la muestra biológica. En una realización, la variación es una variación genética. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En otra realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

La divulgación también proporciona métodos de diagnóstico del lupus en un mamífero, detectando la presencia de una variación en un PRO o polinucleótido asociado a PRO derivado de una muestra biológica obtenida del mamífero, donde la muestra biológica se sabe que comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación. La divulgación también proporciona métodos para ayudar al diagnóstico del lupus en un mamífero, detectando la presencia de una variación en un PRO o polinucleótido asociado a PRO derivado de una muestra biológica obtenida del mamífero, donde se sabe que la muestra biológica comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación. En una realización, la variación es una variación genética. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En otra realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

50 Pueden usarse diversos algoritmos conocidos en la técnica y descritos en este documento para evaluar el riesgo de desarrollar lupus y la respuesta al tratamiento. Las variantes asociadas con un fenotipo pueden interactuar de una manera aditiva dependiente de la dosis alélica. En algunas realizaciones de la invención, puede usarse un algoritmo basado en un esquema de estratificación para evaluar el riesgo de desarrollar lupus, la gravedad de la enfermedad y la respuesta al tratamiento. Los casos de lupus pueden estratificarse en grupos basándose en el número de alelos de riesgo portados. En una realización, el alelo de riesgo se define como el alelo enriquecido en casos de lupus respecto a los controles de los loci. Por ejemplo, en una realización, cuando se enumera un total de 19 alelos de 18 loci, entonces el número posible máximo de alelos de riesgo es igual a 38. Como se describe en este documento, pueden determinarse los casos de lupus estratificados por el número de alelos de riesgo y terciles de la distribución resultante. Los terciles de casos de lupus entonces pueden examinarse para diferencias en la gravedad de la enfermedad y el riesgo y la respuesta al tratamiento. En otra realización, se proporciona un método para predecir si un sujeto con lupus responderá a un agente terapéutico dirigido a PRO o polinucleótido asociado a PRO determinando si el sujeto comprende una variación en un PRO o un polinucleótido asociado a PRO, donde la presencia de una variación en un PRO o polinucleótido asociado a PRO indica que el sujeto responderá al agente terapéutico. En una realización, la variación es una variación genética. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de

las figuras 1-17 y las tablas 1-10. En otra realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización el SNP está en una región codificante del gen.

La divulgación también abarca métodos de detección de la ausencia o presente en un sujeto, o muestra obtenida del mismo, de una variación genética en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10 mediante (a) poner en contacto el ácido nucleico del sujeto o la muestra con cualquiera de los polinucleótidos descritos anteriormente en condiciones adecuadas para la formación de un completo de hibridación entre el ácido nucleico y el polinucleótido; y (b) detectar si el polinucleótido forma pares de bases específicamente con el ácido nucleico en la posición nucleotídica. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

La divulgación también proporciona métodos de detección de ausencia o presencia de una variación genética en un ácido nucleico asociado con un PRO mediante (a) poner en contacto el ácido nucleico con un oligonucleótido específico de alelo que es específico para la variación genética en condiciones adecuadas para la hibridación del oligonucleótido específico de alelo con el ácido nucleico; y (b) detectar la ausencia o presencia de hibridación específica de alelo. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En otra realización, el oligonucleótido específico de alelo es un cebador específico de alelo.

La divulgación también proporciona métodos para evaluar la predisposición de un sujeto a desarrollar lupus detectando la presencia o ausencia en el sujeto de una variación en un PRO o polinucleótido asociado a PRO, donde la presencia de una variación en un PRO o polinucleótido asociado a PRO indica que el sujeto está predispuesto a desarrollar lupus. En una realización, la variación es una variación genética. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En otra realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

La divulgación también proporciona métodos de subclasificación del lupus en un mamífero, comprendiendo el método detectar la presencia de una variación en un polinucleótido asociado a PRO en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en cualquiera de las figuras 1-17 y las tablas 1-10 en una muestra biológica obtenida del mamífero, donde se sabe que la muestra biológica comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación. En una realización, la variación es una variación genética. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En una realización, la subclasificación se caracteriza por implicación tisular/orgánica (por ejemplo, nefritis lúpica), género y/o etnia.

En una realización de los métodos de detección de la invención, la detección comprende realizar un proceso seleccionado de un ensayo de prolongación de cebador; un ensayo de prolongación de cebador específico de alelo; un ensayo de incorporación de nucleótidos específica de alelo; un ensayo de hibridación de oligonucleótido

específico de alelo; un ensayo de 5' nucleasa; un ensayo que emplea balizas moleculares; y un ensayo de ligamiento de oligonucleótido.

La divulgación también proporciona métodos de identificación de un agente terapéutico eficaz para tratar el lupus en una subpoblación de paciente, comprendiendo el método correlacionar la eficacia del agente con la presencia de una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) en la subpoblación de pacientes, donde el SNP es uno de los enumerados en las figuras 1-17 y las tablas 1-10, identificando de ese modo el agente como eficaz para tratar el lupus en dicha subpoblación de pacientes. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

Los métodos de la invención proporcionan información útil para determinar las etapas de intervención clínica apropiadas, si las hay y según lo apropiado. Por lo tanto, en una realización de un método de la divulgación, el método comprende además una etapa de intervención clínica basada en los resultados de la evaluación de la presencia o ausencia de una variación en un PRO o polinucleótido asociado a PRO como se divulga en este documento. Por ejemplo, la intervención apropiada puede implicar etapas profilácticas y de tratamiento, o uno o más ajustes de cualquiera de las etapas profilácticas o de tratamiento del momento basándose en la información genética obtenida por un método de la invención.

Como sería evidente para los expertos en la materia, en cualquier método de la invención, aunque la detección de la presencia de una variación indicaría positivamente una característica de una enfermedad (por ejemplo, presencia o subtipo de una enfermedad), la ausencia de detección de una variación también sería informativa, proporcionando la caracterización recíproca de la enfermedad.

La divulgación también proporciona métodos de amplificación de un ácido nucleico que comprende un polinucleótido asociado a PRO o se proporciona un fragmento del mismo, donde el polinucleótido asociado a PRO o fragmento del mismo comprende una variación genética. En una realización, el método comprende (a) poner en contacto el ácido nucleico con un cebador que hibrida con una secuencia 5' o 3' de la variación genética y (b) prolongar el cebador para generar un producto de amplificación que comprende la variación genética. En una realización, el método comprende además poner en contacto el producto de amplificación con un segundo cebador que hibrida con una secuencia 5' o 3' de la variación genética y prolongar el segundo cebador para generar un segundo producto de amplificación. En una de dichas realizaciones, el método comprende además amplificar el producto de amplificación y el segundo producto de amplificación, por ejemplo, por reacción en cadena de la polimerasa.

En algunas realizaciones, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP de la presente invención. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

Otros métodos adicionales de la invención incluyen métodos de tratamiento del lupus en un mamífero, que comprenden las etapas de obtener un tejido o una muestra celular del mamífero, examinar el tejido o las células para la presencia o ausencia de una variación como se divulga en este documento, y tras determinar la presencia o ausencia de la variación en dicho tejido o muestra celular, administrar una cantidad eficaz de un agente terapéutico apropiado a dicho mamífero. Opcionalmente, los métodos comprenden administrar una cantidad eficaz de un agente terapéutico contra el lupus dirigido y, opcionalmente, un segundo agente terapéutico (por ejemplo, esteroides, etc.) a dicho mamífero.

En una realización, se proporciona un método de tratamiento del lupus, comprendiendo el método administrar al sujeto una cantidad eficaz de un antagonista o agonista de PRO. En una realización, el sujeto muestra variación en un PRO o polinucleótido asociado a PRO. En una realización, la variación es una variación genética. En una realización, la variación genética está en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una de dichas realizaciones, la variación genética comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora)

comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

La invención también proporciona métodos de tratamiento de una afección de lupus en un sujeto en el que se sabe que hay una variación genética presente en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para tratar la afección. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación es un SNP en un polinucleótido asociado a PRO que codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

La divulgación también proporciona métodos de tratamiento de un sujeto que tiene una afección de lupus, comprendiendo el método administrar al sujeto un agente terapéutico que se sabe que es eficaz para tratar la afección en un sujeto que tiene una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación es un SNP en un polinucleótido asociado a PRO que codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

La divulgación también proporciona métodos de tratamiento de un sujeto que tiene una afección de lupus, comprendiendo el método administrar al sujeto un agente terapéutico que previamente ha demostrado ser eficaz para tratar dicha afección en al menos un estudio clínico, donde el agente ese administró a al menos cinco sujetos humanos que tenían cada uno una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación es un SNP en un polinucleótido asociado a PRO que codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En una realización, los al menos cinco sujetos tenían dos o más SNP diferentes en total para el grupo de al menos cinco sujetos. En una realización, los al menos cinco sujetos tenían el mismo SNP para el grupo completo de al menos cinco sujetos.

La divulgación también proporciona métodos de tratamiento de un sujeto con lupus que es de una subpoblación específica de pacientes con lupus, que comprende administrar al sujeto una cantidad eficaz de un agente terapéutico que está aprobado como agente terapéutico para dicha subpoblación, donde la subpoblación se caracteriza al menos en parte por la asociación con una variación genética en una posición nucleotídica correspondiente a un SNP enumerado en las figuras 1-17 y las tablas 1-10. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación es una SNP en un polinucleótido asociado a PRO que codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En una realización, la subpoblación es de ascendencia europea. En una realización, la divulgación proporciona un método que comprende fabricar un agente terapéutico contra el lupus y envasar el agente con instrucciones para administrar el agente a un sujeto que tiene o se cree que tiene lupus y que tiene una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación es un SNP en un polinucleótido asociado a PRO que codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen.

La divulgación también proporciona métodos de especificación de un agente terapéutico para su uso en una subpoblación de pacientes con lupus, comprendiendo el método proporcionar instrucciones para administrar el

agente terapéutico a una subpoblación de pacientes caracterizada por una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en la figura 5-10. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación es un SNP en un polinucleótido asociado a PRO que codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En una realización, la subpoblación es de ascendencia europea.

La divulgación también proporciona métodos para comercializar un agente terapéutico para su uso en una subpoblación de pacientes con lupus, comprendiendo el método informar a una audiencia determinada acerca del uso del agente terapéutico para tratar a la subpoblación de pacientes caracterizada por la presencia, en pacientes de dicha subpoblación, de una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en la figura 5-10. En una realización, la variación comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, la variación es un SNP en un polinucleótido asociado a PRO que codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento (por ejemplo, como se expone en las figuras 1-17 y las tablas 1-10). En una realización, la variación genética está en ADN genómico que codifica un gen (o su región reguladora), donde el gen (o su región reguladora) comprende un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10. En una realización, el SNP está en una región no codificante del gen. En una realización, el SNP está en una región codificante del gen. En una realización de cualquiera de los métodos anteriores que comprenden el uso de un agente terapéutico, dicho agente comprende un agente terapéutico contra el lupus como se divulga en este documento.

La divulgación también proporciona métodos para modular la señalización a través del receptor de linfocitos B en un sujeto en el que se sabe que hay una variación genética presente en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para modular la señalización a través del receptor de linfocitos B.

La divulgación también proporciona métodos para modular la diferenciación de células Th17 en un sujeto en el que se sabe que hay variación genética presente en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para modular la diferenciación de células Th17.

Kits

En una realización, de la divulgación, se proporcionan kits. En una realización, un kit comprende cualquiera de los polinucleótidos descritos en este documento, opcionalmente con una enzima. En una realización, la enzima es al menos una enzima seleccionada de una nucleasa, una ligasa y una polimerasa.

En una realización, la divulgación proporciona un kit que comprende una composición de la invención e instrucciones para usar la composición para detectar el lupus determinando si el genoma de un sujeto comprende una variación genética como se divulga en este documento. En una realización, la composición de la divulgación comprende una pluralidad de polinucleótidos capaces de hibridar específicamente con al menos 1, 2, 3, 4 o 5 polinucleótidos asociados a PRO, comprendiendo cada polinucleótido asociado a PRO una variación genética en una posición nucleotídica correspondiente a la posición de un SNP expuesto en cualquiera de las figuras 1-17 y las tablas 1-10, o complementos de dichos polinucleótidos asociados a PRO. En una realización, la composición de la divulgación comprende polinucleótidos que codifican al menos una parte de un PRO. En una realización, la composición de la invención comprende cebadores de ácido nucleico capaces de unirse a y lograr la polimerización (por ejemplo, amplificación) de al menos una parte de un polinucleótido asociado a PRO. En una realización, la composición de la divulgación comprende un agente de unión (por ejemplo, cebador, sonda) que detecta específicamente un polinucleótido asociado a PRO (o complemento del mismo) (o producto génico correspondiente). En una realización, la composición de la divulgación comprende un agente de unión que se une específicamente a al menos una parte de un PRO. En una realización, la divulgación proporciona un artículo de fabricación que comprende un agente terapéutico, combinado con instrucciones para usar el agente para tratar a un paciente con lupus que tiene una variación en un polinucleótido asociado a PRO como se divulga en este documento.

Para su uso en las aplicaciones descritas o sugeridas anteriormente, también se proporcionan por la invención kits o artículos de fabricación. Dichos kits pueden comprender un medio de vehículo que está compartimentado para recibir en confinamiento cerrado uno o más medios de recipiente tales como viales, tubos y similares, comprendiendo cada uno de los medios de recipiente uno de los elementos diferentes a usar en el método. Por ejemplo, uno de los medios de recipiente puede comprender una sonda que está o puede estar marcada de forma detectable. Dicha sonda puede ser un polinucleótido específico para un polinucleótido asociado a PRO. Cuando el kit utiliza hibridación de ácido nucleico para detectar el ácido nucleico diana, el kit también puede tener recipientes que contienen uno o más nucleótidos para la amplificación de la secuencia de ácido nucleico diana y/o un recipiente

que comprende un medio indicador, tal como una proteína de unión a biotina, tal como avidina o estreptavidina, unido a una molécula indicadora, tal como un marcador enzimático, fluorescente o radioisotópico.

El kit de la divulgación normalmente comprenderá el recipiente descrito anteriormente y uno o más recipientes diferentes que comprenden materiales deseables desde un punto de vista comercial y del usuario, incluyendo tampones, diluyentes, filtros, agujas, jeringas y prospectos con instrucciones para su uso. Puede haber una etiqueta presente en el recipiente para indicar que la composición se usa para un tratamiento específico o aplicación no terapéutica, y también puede indicar directrices para su uso *in vivo* o *in vitro*, tales como los descritos anteriormente.

Los kits de la divulgación tienen varias realizaciones. Una realización típica es un kit que comprende un recipiente, una etiqueta en dicho recipiente y una composición contenida dentro de dicho recipiente; donde la composición incluye un agente de detección para un PRO o polinucleótido asociado a PRO, la etiqueta en dicho recipiente indica que la composición puede usarse para evaluar la presencia del PRO o polinucleótido asociado en PRO en al menos un tipo de célula de mamífero e instrucciones para usar el agente de detección para evaluar la presencia del PRO o polinucleótido asociado a PRO en al menos un tipo de célula de mamífero. El kit puede comprender además un conjunto de instrucciones y materiales para preparar una muestra tisular y aplicar el anticuerpo y la sonda a la misma sección de una muestra tisular. Por ejemplo, un kit puede comprender un recipiente, una etiqueta en dicho recipiente y una composición contenida dentro de dicho recipiente; donde la composición incluye un polinucleótido que hibrida con un complemento de un polinucleótido asociado a PRO en condiciones rigurosas, la etiqueta en dicho recipiente indica que la composición puede usarse para evaluar la presencia de un polinucleótido asociado a PRO en al menos un tipo de célula de mamífero e instrucciones para usar el polinucleótido para evaluar la presencia de ARN o ADN asociado a PRO en al menos un tipo de célula de mamífero.

Otros componentes opcionales en el kit incluyen uno o más tampones (por ejemplo, tampón de bloqueo, tampón de lavado, tampón de sustrato, etc.), otros reactivos tales como sustratos (por ejemplo, cromógeno) que está alterado químicamente por un marcador enzimático, solución de recuperación de epítipo, muestras de control (controles positivos y/o negativos), uno o más portaobjetos de control, etc.

Métodos de comercialización

La divulgación de este documento también abarca un método para comercializar un agente terapéutico contra el lupus o una composición farmacéuticamente aceptable del mismo, que comprende promover, instruir y/o especificar a un grupo destinatario, el uso del agente o composición farmacéutica del mismo para tratar a un paciente o población de pacientes con lupus del que se ha obtenido una muestra que muestra la presencia de una variación genérica como se divulga en este documento.

La comercialización generalmente consigue su comunicación a través de un medio no personal en que se identifica el auspiciador y se controla el mensaje. La comercialización con los fines de este documento incluye publicidad, relaciones públicas, colocación del producto, patrocinio, suscripción y promoción de ventas. Este término también incluye anuncios informativos patrocinados que aparecen en cualquiera de los medios de comunicación escritos diseñados para reclamar a un público masivo para persuadir, informar, promover, motivar o modificar de otro modo el comportamiento hacia un patrón favorable de adquisición, soporte o aprobación de la invención de este documento.

La comercialización del método de diagnóstico de este documento puede conseguirse por cualquier medio. Los ejemplos de medios de comercialización usados para suministrar estos mensajes incluyen televisión, radio, películas, revistas, periódicos, internet y vallas publicitarias, incluyendo anuncios, que son mensajes que aparecen en los medios de difusión.

El tipo de comercialización usada dependerá de muchos factores, por ejemplo, de la naturaleza de la audiencia determinada a alcanzar, por ejemplo, hospitales, compañías de seguros, clínicas, doctores, enfermeras y pacientes, así como consideraciones de coste y las leyes y regulaciones jurisdiccionales relevantes que rigen la comercialización de medicamentos y agentes de diagnóstico. La comercialización puede ser individualizada o personalizada basándose en las caracterizaciones del usuario definidas por la interacción del servicio y/u otros datos tales como demografía de usuarios y localización geográfica.

Los siguientes son ejemplos de los métodos y composiciones de la invención. Se entiende que pueden ponerse en práctica otras diversas realizaciones, dada la descripción general proporcionada anteriormente.

Ejemplos

La información bibliográfica para las referencias citadas (e indicadas por número) en los ejemplos 1-3 se proporcionan al final del ejemplo 3. La información bibliográfica para las referencias citadas (e indicadas por número) en los ejemplos 4-6 se proporcionan al final del ejemplo 6.

Ejemplo 1. Materiales y métodos para una exploración de asociación de todo el genoma en lupus eritematoso diseminado

Este ejemplo describe materiales y métodos emprendidos para realizar una exploración de todo el genoma para SLE en una muestra grande que comprende 1311 casos de SLE y 3340 controles. De 500 000 variantes, que capturaron una variación común entre un 85 % estimado del genoma humano, se genotiparon 24 y se ensayaron para una asociación a SLE.

Sujetos

Se genotiparon las muestras de los casos de SLE de las siguientes colecciones: a) 338 sujetos de la Autoimmune Biomarkers Collaborative Network (ABCoN), un depósito patrocinado por NIH/NIAMS,²⁵ b) 141 sujetos del Multiple Autoimmune Disease Genetics Consortium (MADGC),²⁶ c) 613 sujetos de la University of California San Francisco (UCSF) Lupus Genetics Project^{10,27} y d) 335 sujetos del University of Pittsburgh Medical Center (UPMC)²⁸ más 8 muestras recogidas en el The Feinstein Institute for Medical Research. Todos los casos de SLE son autodenominados caucásicos. El diagnóstico de SLE (cumplimiento de cuatro o más de los criterios definidos por el American College of Rheumatology (ACR)²⁹) se confirmó en todos los casos por revisión del registro médico (94 %) o a través de documentación escrita de los criterios por los reumatólogos a cargo (6 %). Se recuperaron y tabularon los datos clínicos en cada institución. La figura 4 muestra los recuentos y porcentajes para cada uno de los once criterios de clasificación del ACR para SLE.²⁹

Se examinó un total de 3583 muestras de control en los análisis de asociación. Como parte de este proyecto, se seleccionaron 1861 muestras de control y después se genotiparon para la colección del New York Cancer Project (NYCP)³⁰, basándose en la etnia autodenominada, el género y la edad. Además, se obtuvieron datos de genotipo de 1722 muestras de control de autodenominados caucásicos de la base de datos disponible al público iControlDB <www.illumina.com/pages.ilmn?ID=231>.

Para la replicación, se genotiparon muestras de ADN de una colección independiente de 793 pacientes suecos de SLE (todos los cuales cumplían cuatro o más de los criterios de clasificación para SLE como se define por el ACR) y 857 individuos suecos sanos de control. Los pacientes eran de clínicas de reumatología en Lund, Uppsala, Karolinska (Solna) y Umea University Hospitals⁷. Los comités de revisión institucional de todas las instituciones colaboradoras aprobaron estos estudios y todos los participantes dieron su consentimiento informado.

Genotipado

Las muestras de control del NYCP (N=1861) se genotiparon en el Illumina HumanHap550 Genotyping BeadChip³¹ en el The Feinstein Institute. Se genotiparon 1465 muestras (464 casos, 1001 controles) en el HumanHap550v1 chip y se genotiparon 1875 muestras (1015 casos, 860 controles) en el HumanHap550v3 chip. Los datos de genotipo de 1452 de estas muestras de control se presentaron a iControlDB y quedaron disponibles al público antes de la publicación. Adicionalmente, se obtuvo un conjunto independiente de 1722 muestras caucásicas genotipadas usando el HumanHap550 BeadChip de los estudios 66 y 67 del iControlDB <www.illumina.com/pages.ilmn?ID=231>. Las muestras de casos se genotiparon en The Feinstein Institute en fases en serie: la serie 1 consistía en 479 de ABCoN y MADGC, la serie 2 incluía los 613 casos de UCSF y la serie 3 estaba compuesta de 387 casos de UPMC y The Feinstein Institute. Los 545 080 polimorfismos de un único nucleótido (SNP) presentes en ambas versiones HumanHap550 se hicieron avanzar en el análisis. Las muestras de casos y de control con tasas de éxito promedio < 80 % a través del chip se volvieron a genotipar.

En la colección de replicación sueca, los SNP rs11574637 y rs13277113 se genotiparon usando ensayos homogéneos de prolongación de cebador de una única base con detección de polarización de la fluorescencia en la plataforma de tecnología de SNP en Uppsala <www.genotyping.se> y reactivos de Perkin-Elmer.³² La tasa de éxito de genotipo en las muestras fue de un 96 % y la reproducibilidad fue de un 100 % de acuerdo con los ensayos duplicados de un 4,6 % de los genotipos. Se genotipó un árbol genealógico CEPH de tercera generación con 20 miembros en paralelo con las muestras del estudio, y no se observó desviación de la herencia mendeliana para cualquiera de los SNP.

Filtros de calidad de datos

Las muestras con una tasa de éxito promedio de ≤ 95 % (N=42) o cuando el género informado del individuo era discordante con el género observado (N=21) se excluyeron del análisis. La identidad por estado (IBS) a través del genoma se estimó para cada muestra, y las muestras se examinaron para la relación críptica. Una muestra de cada par estimado como duplicado o familiar de 1^{er}-3^{er} grado se eliminó ($P_i\text{-hat} > 0,10$ y $Z1 \geq 0,15$, N=161). Tres de estos pares estaban compuestos de tanto un caso como de un control; se eliminó el control. Los SNP con una frecuencia en los casos de < 1 % (N=21 644) o un HWE $P \leq 1 \times 10^{-6}$ en los controles (N=2819) se eliminaron del análisis. Los SNP con valores faltantes > 5 % (N=6074) se eliminaron. Los SNP se ensayaron para la probabilidad de una diferencia significativa en valores faltantes entre los casos y los controles; los SNP con $P \leq 1 \times 10^{-5}$ (N=7646) se eliminaron. Los SNP también se ensayaron para los efectos de lote: por ejemplo, entre muestras ABCoN y todos los

demás casos; los SNP con $P < 1 \times 10^{-9}$ ($N=13$) se eliminaron.

Los valores atípicos de la población se detectaron usando EIGENSTRAT³³. Se excluyeron muestras de más de 6 desviaciones típicas de la media a lo largo de cualquiera de los 10 componentes principales superiores del análisis ($N=141$). Los datos de las muestras de las 3340 muestras de control restantes se asignaron aleatoriamente a cada serie de caso de SLE de forma proporcionada, produciendo una relación de control:caso de ~2,5 (tabla 1).

La serie 1 consistía en 411 casos y 1047 controles, la serie 2 estaba compuesta de 595 casos y 1516 controles y la serie 3 estaba compuesta de 305 casos y 777 controles. Globalmente, un 93 % de los casos eran mujeres y un 62 % de los controles eran mujeres. No se apreciaron diferencias significativas en las frecuencias alélicas entre hombres y mujeres.

Los SNP con $> 2\%$ de pérdida de datos en al menos una serie y cuando los datos perdidos estaban distribuidos de forma desigual entre los casos y los controles (valores faltantes diferenciales, $P < 1 \times 10^{-3}$) se eliminaron ($N= 3323$). Los SNP en la región pseudoautosómica del cromosoma X ($N=13$) no mostraron asociación significativa y se excluyeron del análisis adicional. El filtrado de la muestra y el marcador se realizó usando módulos analíticos dentro del programa informático PLINK³⁴. Para cada serie, se hizo avanzar un total de 502 033 SNP a análisis posteriores.

Análisis de datos

La asociación de todos los SNP a la susceptibilidad a SLE se calculó usando tablas de contingencia 2x2. Después se calculó un factor de inflación de control (λ_{gc}) genómico para cada serie de muestras.³⁵ El factor de inflación de control genómico es una medida basada en la mediana de ji cuadrado que refleja si el volumen de la distribución se adapta a la hipótesis nula ($\lambda_{gc}=1,0$). Un valor de $\lambda_{gc} > 1$ indica una elevación de la estadística de asociación de ji cuadrado promedio debido a artefactos técnicos sistémicos o la presencia de estratificación de la población. Después de eliminar los datos de baja calidad para minimizar los artefactos técnicos, se apreciaron evidencias de inflación para cada serie: 1,14, 1,18 y 1,11, respectivamente, para la serie 1, 2 y 3. Para corregir la presencia de estratificación de la población, se calcularon los componentes principales para cada serie usando un subconjunto de SNP en EIGENSTRAT. Los SNP con MAF de casos $< 2\%$ (5011), HWE P de control $\leq 1 \times 10^{-4}$ (1792) o datos perdidos $> 1\%$ (50414) se eliminaron, así como los SNP en regiones de patrones LD anómalos debido a variación estructural en los cromosomas 6 (de 24-36 Mb), 8 (8-12 Mb), 11 (42-58 Mb) y 17 (40-43 Mb). Los 440 202 SNP restantes se usaron para calcular los componentes principales. En cada serie, los cuatro primeros componentes principales se usaron para ajustar la estadística de asociación para los 502 033 SNP. Después del ajuste para la estratificación de la población, el λ_{gc} para cada serie se aproximó a 1,0 (véase la tabla 1). La estadística de asociación corregida para cada serie se combinó por la fusión ponderada del valor Z que incorpora λ_{gc} (figura 12). Los 50 loci superiores se muestran en la figura 5. Adicionalmente, la estadística para todos los SNP que pasan los filtros QC de cada serie y la estadística de asociación combinada se han resumido en la tabla (no mostrada).

Para ensayar la heterogeneidad entre los tres estudios de casos-controles para las variantes más asociadas, se ejecutó el ensayo de Breslow-Day implementado en PLINK para el SNP con la mejor asociación en cada una de estas regiones: HLA DRB, STAT4, IRF5, BLK y ITGAM/ITGAX. No se detectó heterogeneidad significativa (cada $P > 0,2$).

Las asociaciones entre los SNP individuales y los subfenotipos se calcularon para el conjunto de datos combinado (figura 9), usando el ensayo de heterogeneidad de Mantel-Haenszel y la relación de probabilidad combinada implementada en Stata 9.2 (www.stata.com/). Los valores P calculados no se ajustaron para ensayo múltiple, ya que se sabe que los criterios de la ACR están correlacionados y una corrección simple de Bonferroni de $\alpha=0,05/11=0,0045$ probablemente sería excesivamente conservativa.

Análisis de expresión génica

Las mediciones de expresión génica de líneas de linfocitos B transformadas con virus de Epstein-Barr de 210 individuos HapMap sanos no relacionados de un conjunto de datos disponible al público (proyecto GENEVAR, www.sanger.ac.uk/humgen/genevar/) se examinaron para una correlación con variantes asociadas significativamente con SLE.³⁶ Específicamente, se examinó la intensidad media de fluorescencia de 4 mediciones de sondas para *BLK* (GI_33469981-S), *C8orf13* (GI_32698772-S), *ITGAM* (GI_6006013-S), *IT-GAX* (GI_34452172-S), *ACTB* (beta-actina, GI_5016088-S) y *GAPDH* (GI_7669491-S) de 60 residentes de Estados Unidos con ascendencia noreuropea y de Europa occidental (CEU), 60 individuos Yoruba (YRI), 45 chinos Han de Beijing (CHB) y 45 individuos japoneses de Tokio (JPT). Los datos de expresión para *BLK*, *C8orf13*, *GAPDH* y *ACTB* se estratificaron por el genotipo de rs13277113 (obtenido de HapMap (www.hapmap.org)), y se midió la significación de la expresión diferencial por un ensayo t bilateral asumiendo una varianza igual. Asimismo, los datos de expresión para *ITGAM*, *ITGAX*, *GAPDH* y *ACTB* se estratificaron por el genotipo en rs11574637 y se ensayaron para la significación usando un ensayo- t . Los datos de expresión normalizados en una escala logarítmica entre las poblaciones HapMap como se describe por el proyecto GENEVAR produjeron resultados similares a la intensidad media de fluorescencia.

La asociación de la expresión de *BLK* y *C8orf13* con la variación cis genética en un conjunto independiente de 400

linfocitos B transformados con EBV se obtuvo por el examen y la recolección de datos de un estudio publicado recientemente (www.sph.umich.edu/csg/liang/asthma/).³⁷ Específicamente, la asociación de un sustituto para rs13277113 (rs4840568) con los niveles de expresión de BLK (sonda 206255_at) y C8orf13 (sonda 226614_s_at) se midió como se describe por Dixon et al.³⁷

Ejemplo 2 identificación de C8orf13/BLK y ITGAM/ITGAX como novedosos loci de susceptibilidad

Análisis de asociación de todo el genoma

Un total de 502 033 SNP polimórficos en los chips Illumina pasaron los filtros de control de calidad y se ensayaron para la asociación a SLE de un modo estratificado usando 3 series de casos-contróles (tabla 1). Se calculó una estadística de asociación combinada por la adición de los valores Z convertidos de la estadística de ensayo ji cuadrado, corregida por EIGENSTRAT, ponderada para el tamaño de la serie y ajustado para el λ_{gc} residual de cada serie (véase métodos).

Se muestra una comparación de los valores P de metanálisis observados respecto a los valores P para una distribución nula en la figura 1. Se observó una desviación significativa de la distribución nula en la cola de la distribución (figura 1A, rombos negros), que puede indicar la presencia de verdaderas asociaciones positivas. Se apreció una fuerte asociación a SLE para tres loci de riesgo establecidos. En la región de HLA de clase II, rs2187668 es un factor de pronóstico casi perfecto del alelo DRB1*0301³⁸ y fue la variante más fuertemente asociada con SLE en el análisis combinado ($P = 3 \times 10^{-21}$). 157 SNP adicionales de la región HLA, muchos de los cuales están correlacionados con el alelo DRB1*0301, han observado valores P menores de 5×10^{-7} (figura 1B). Se observó una fuerte asociación con variantes ligadas al haplotipo de riesgo bien validado de factor 5 regulador de interferón (*IRF5*) (por ejemplo, rs10488631, $P = 2 \times 10^{-11}$).⁷⁻⁹ Además, se observó una asociación con *STAT4* (rs7574865, $P = 9 \times 10^{-14}$). Se presentó una asociación de *STAT4* con SLE y también artritis reumatoide recientemente.¹⁰ El conjunto de datos de SLE en esta ocasión solapa con los del informe previo¹⁰, e incluye 341 casos y 2905 controles adicionales que no se incluyeron en el análisis previo. Además, los valores P para los SNP de *STAT4* superiores presentados en esta ocasión se han corregido para la estratificación de la población.

Después de eliminar las variantes en HLA, *IRF5* y *STAT4* del análisis de ji esperado frente al observado, no se eliminó la desviación de los valores P de la distribución nula (figura 1A, círculos), lo que sugiere la presencia de loci adicionales de SLE. Como se muestra en la figura 1B, múltiples SNP cerca del gen de la tirosina cinasa de linfocitos B (*BLK*) y en la región que contiene los genes de integrina alfa M (*TGAM*) e integrina alfa X (*ITGAX*) estaban muy asociados con SLE con el análisis combinado. Ninguno de estos genes o regiones se ha implicado previamente en la susceptibilidad a SLE.

BLK/C8orf13

Varias variantes del brazo corto del cromosoma 8 (8p23.1) se asociaron con SLE (figura 2, tabla 2, figura 8). El alelo "A" de rs13277113 estaba muy enriquecido en los casos de SLE estadounidenses respecto a los controles ($P = 8 \times 10^{-8}$, OR combinado = 1,39, 95 % C.I. = 1,26-1,54). para confirmar esta observación inicial, se tipó una colección independiente de 793 casos de SLE y 857 controles coincidentes de Suecia para rs13277113 y también se observó una asociación convincente del alelo minoritario "A" con SLE ($P = 3,6 \times 10^{-4}$, OR = 1,33, 95 % C.I. = 1,13-1,55; tabla 2). Un metanálisis de rs13277113 usando las muestras tanto estadounidenses como suecas mostró un $P = 1,4 \times 10^{-10}$, que sobrepasa el umbral de significación riguroso en todo el genoma de asociación $P < 5 \times 10^{-8}$.³⁹

El rs13277113 se sitúa en el intervalo entre dos genes transcritos en direcciones opuestas: *BLK* - una tirosina cinasa de la familia de src que emite señales posteriores al receptor de linfocitos B, y *C8orf13* - un gen expresado de forma ubicua de función desconocida (figura 2). Ninguna variante de la región codificante conocida de *BLK* o *C8orf13* está en el desequilibrio del ligamiento (LD) con rs13277113.

Se ha demostrado que la variación genética común se correlaciona con niveles de expresión génica en *cis*.^{8,36,37,40} Para determinar si los SNP promotores asociados podrían influir en la expresión del ARNm de *BLK* y/o *C8orf13*, se consultó un conjunto de datos de expresión génica generado a partir de líneas celulares de linfocitos B transformados con virus de Epstein-Barr de las 210 muestras HapMap no relacionadas.³⁶ Sorprendentemente, el alelo "A" de riesgo de rs13277113 se asoció con niveles inferiores de expresión de ARNm de *BLK* (figura 2B). Los homocigotos para el alelo A mostraron niveles ~ 50 % inferiores de expresión que los homocigotos para el alelo G, y los heterocigotos A/G tenían niveles intermedios. Curiosamente, la expresión del gen *C8orf13* también se correlacionó con el haplotipo de riesgo, pero en la dirección opuesta. El alelo A de rs13277113 se asoció con mayor expresión de *C8orf13* en las líneas transformadas, mientras que el alelo G se asoció significativamente con expresión inferior (figura 2C). De nuevo, los heterocigotos A/G mostraron niveles intermedios de expresión. La expresión de varios ARNm de control (por ejemplo, beta-actina, GAPDH) no varió en las líneas celulares basándose en el genotipo en rs13277113 (figura 6) y se observaron diferencias alélicas constantes en la expresión de BLK en todas las poblaciones HapMap (figura 7). Estos resultados se confirmaron analizando un conjunto de datos independiente de expresión génica y SNP de todo el genoma en 400 líneas de linfocitos B transformadas que no son

HapMap.³⁷ En este conjunto de datos, un marcador correlacionado con rs13277113 (rs4840568, $r^2 = 0,77$) se asoció con expresión disminuida de *BLK* ($P = 8,9 \times 10^{-27}$, sonda 206255_at) y también expresión aumentada de *CBorf13* ($P = 4,6 \times 10^{-35}$, sonda 226614_s_at).

Múltiples sitios de unión al factor de transcripción conservados, incluyendo los motivos para IRF1, PPARG y un elemento de respuesta estimulado por interferón, están localizados en la región 5' de *BLK* y *CBorf13*. Sin embargo, ni rs13277113 ni las variantes correlacionadas ($r^2 > 0,5$) alteraron los sitios de unión del factor de transcripción conocidos u otros motivos de ácido nucleico funcionales conocidos. Se concluye que rs13277113, o la variación fuertemente asociada con rs13277113, altera el nivel de la expresión de ARNm de *BLK* y *C8orf13*.

ITGAM/ITGAX

Las variantes dentro del grupo de los genes de la cadena alfa de integrina en el cromosoma 16 también se asociaron significativamente con SLE (figura 3, tabla 2). Se observó una asociación reproducible del alelo "C" de rs11574637 entre las 3 series de SLE ($P = 5 \times 10^{-7}$, OR = 1,30, 95 % C.I. = 1,17-1,45). De forma importante, el alelo "C" de rs11574637 mostró un fuerte enriquecimiento similar en la serie de replicación sueca ($P = 4 \times 10^{-7}$, OR = 1,59, 95 % C.I. = 1,33-1,91; tabla 2) y el metanálisis mostró un P combinado = 3×10^{-11} . Se concluyó que variación ligada a rs11574637 marca un alelo de riesgo de SLE confirmado, y que el locus ITGAM/ITGAX contribuye a la patogénesis de SLE.

El rs11574637 es parte de un gran bloque de SNP correlacionados que abarca ~ 150 kb de varios genes codificantes incluyendo *ITGAM* y la parte 5' de *ITGAX* (figura 3A). Tanto *ITGAM* como *ITGAX* se expresan a niveles detectables en linfocitos B transformados con EBV, sin embargo, rs11574637 no se correlacionaban significativamente con niveles de expresión de ARNm de ningún gen (datos no mostrados). Es de interés potencial que el SNP rs11574637 está correlacionado con 2 variantes no sinónimas de *ITGAM*. En la población de control, se correlaciona una variante Pro1146Ser (rs1143678, $P = 2,5 \times 10^{-5}$) con un r^2 de 0,85 para la variante rs11574637 asociada a la enfermedad. El alelo "C" de rs1143678 y el alelo 1146Ser forma un haplotipo en un 18,2 % de los cromosomas de control; el alelo "C" también está presente en un 2 % diferente del haplotipo que carece del alelo 1146Ser. Un segundo alelo no sinónimo (rs1143683, Ala858Val) no se genotipó directamente en el presente estudio, sino que se correlacionó fuertemente con Pro1146Ser ($r^2 = 0,85$ en HapMap CEU). Se requerirán estudios adicionales para determinar si las variantes no sinónimas de *ITGAM* o uno o más alelos adicionales están subyacentes a la asociación dentro de la región ITGAM/ITGAX.

Asociaciones con características clínicas de SLE

Finalmente, se examinaron las asociaciones entre los dos SNP superiores, rs11574637 (*BLK*) y rs13277113 (*ITGAM*), y la presencia de criterios de ACR individuales, usando las series de casos combinadas 1-3 (figura 9 y véase métodos). La asociación más fuerte fue una relación inversa entre el alelo minoritario rs11574637 y la presencia de artritis, OR = 0,73 (95 % CI = 0,59-0,91, $P = 0,0045$). Ambas variantes estaban moderadamente asociadas con criterios hematológicos: rs11574637, OR = 1,21 (95 % CI = 1,00-1,47, $P = 0,04$) y rs13277113, OR = 1,23 (95 % CI = 1,03-1,46, $P = 0,02$). No se observaron otras asociaciones significativas.

Análisis

El actual esfuerzo describe los resultados de un estudio exhaustivo de asociación de todo el genoma realizado en SLE. Estudiando una gran cantidad de casos de SLE - 1311 - y un grupo incluso mayor de controles - 3340, se detectaron los alelos principales que contribuyen al riesgo a SLE. Las fuertes señales observadas en la región de HLA, *IRF5* y *STAT4* sirvieron como controles positivos para el experimento y confirmaron que estos loci están entre los factores genéticos más importantes de esta enfermedad.

La tirosina cinasa *BLK* de la familia de src es un nuevo gen candidato interesante para SLE. La expresión de *BLK* está muy restringida al linaje de linfocitos B.⁴¹ La expresión de *Blk* en ratones se observa por primera vez en el ciclo de células pro-B tardías, continua a través del desarrollo de linfocitos B y posteriormente está regulado por disminución en linfocitos B plasmáticos.⁴² Un ratón con el *Blk* inactivado no tiene fenotipo manifiesto⁴³, y no se han realizado estudios funcionales en linfocito B humanos. Sin limitarse a teoría alguna, *BLK* es una de las tirosinas cinasas que transducen las señales posteriores al receptor de linfocitos B y, quizás, tiene una función redundante en ratones, dada la ausencia de un fenotipo en los que tienen el gen inactivado. Hay precedentes de diferencias principales en las especies en la función de las cinasas asociadas al receptor de linfocitos B. Por ejemplo, la deficiencia de tirosina cinasa (*BTK*) de Bruton en seres humanos da lugar a agammaglobulinemia asociada al cromosoma X y una ausencia completa de linfocitos B.⁴⁴ Sin embargo, la deficiencia de *Btk* en ratones está asociada con un fenotipo mucho más leve, con producción de linfocitos B maduros que están funcionalmente alterados.⁴⁵

La señalización a través del receptor de linfocitos B es importante para establecer el repertorio de linfocitos B a través de la inducción de anergia, eliminación y edición del receptor durante el desarrollo de linfocitos B.^{46,47} Como se muestra en esta ocasión, el alelo de riesgo en *BLK* está asociado con expresión reducida del ARNm de *BLK* en líneas de linfocitos B transformados. Sin limitare a teoría alguna, los niveles alterados de proteína de BLK podrían

influir en los mecanismos de tolerancia en linfocito B, que predisponen a los individuos a autoinmunidad sistémica. Se ha demostrado recientemente un mecanismo similar para *Ly108*, uno de los loci genéticos principales en el modelo de ratón NZM2410 para lupus.⁴⁸ Por consiguiente, en una realización de la invención, un experto en la materia puede usar la información proporcionada en este documento para evaluar el efecto del haplotipo de riesgo sobre la expresión del gen expresado de forma ubicua *C8orf13*.

Un segundo locus identificado en esta exploración es *ITGAM/ITGAX*. Aunque *ITGAX* no se excluye de la consideración basada en el LD fuerte en la región que se extiende hasta la parte 5' *ITGAX*, los datos sugieren que *ITGAM* puede ser el gen relevante en la región. *ITGAM* (también conocido como CD11b, Mac-1 y el receptor de tipo 3 del complemento) es una molécula de cadena alfa de integrina bien caracterizada que se expresa por diversos tipos celulares mieloides, incluyendo células dendríticas, macrófagos, monocitos y neutrófilos.⁴⁹⁻⁵¹ *ITGAM* forma un heterodímero con *ITGB2* (CD18) y media la adhesión entre los tipos celulares en el sistema inmunitario y la adhesión de células mieloides al endotelio.⁵² Los ratones deficientes para *ITGAM* muestran una progresión potenciada de la enfermedad e inflamación en varios modelos de autoinmunidad,⁵³⁻⁵⁵ incluyendo lupus y los datos recientes sugieren que *ITGAM* puede funcionar normalmente suprimiendo la diferenciación de Th17,⁵⁶ una ruta que se ha ligado a la inducción de autoinmunidad. Es de interés que la expresión de *CD11b* se ha presentado elevada en los neutrófilos de pacientes con SLE activo.⁵⁷ El alelo de riesgo para *ITGAM* con sus dos alelos no sinónimos altamente correlacionados puede predisponer a una función y/o regulación de la expresión alterada de la proteína, contribuyendo de ese modo a autoinmunidad sistémica.

En resumen, los datos actuales identifican dos nuevos loci de susceptibilidad para SLE: *BLK/C8orf13* en el cromosoma 8 y *ITGAM/ITGAX* en el cromosoma 16. Los genes candidatos más probables dentro de esto dos loci son *BLK* y *ITGAM*. La identificación de estos genes proporciona nuevas perspectivas importantes sobre la base genética de SLE y también sugiere nuevas dianas potenciales para tratamiento.

Ejemplo 3 una exploración de asociación de todo el genoma en lupus eritematoso diseminado (SLE) e identificación de novedosos loci correlacionados con SLE

En este ejemplo, el conjunto de datos inicial consistía en los casos y los controles del estudio de asociación de todo el genoma descrito anteriormente en los ejemplos 1 y 2, con los genotipos de los chips Illumina HumanHap550v1 chips y los chips Illumina HumanHap550v3. El conjunto de datos de los chips Illumina HumanHap550v1 consistía en 555352 SNP en cada uno de los 464 casos y 1962 controles. El conjunto de datos de los chips Illumina HumanHap550v3 consistía en los 561466 SNP en cada uno de los 971 casos y 1621 controles. Para cada conjunto de datos se aplicaron filtros de control de calidad de forma similar a la manera descrita anteriormente en los ejemplos 1 y 2. El conjunto de datos resultante de los chips HumanHap550v1 consistía en 534523 SNP en cada uno de los 422 casos y 1881 controles. El conjunto de datos resultantes de los chips HumanHap550v3 consistía en 549273 SNP en cada uno de los 929 casos y 1558 controles.

El conjunto de datos anterior de los chips Illumina HumanHap550v1 se combinó con el conjunto de datos anterior de los chips HumanHap550v3. El conjunto de datos resultante consistía en 564307 SNP en cada uno de los 1351 casos y 3439 controles. Este conjunto de datos se combinó con los genotipos de los estudios de cáncer de mama y de próstata CGEMS: 553820 SNP en cada una de las 4527 muestras, usadas como controles. El conjunto de datos resultante consistía en 570099 SNP en cada uno de los 1351 casos y 7966 controles. Se aplicaron filtros de control de calidad de forma similar a la manera descrita anteriormente en los ejemplos 1 y 2. El conjunto de datos resultante consistía en 446856 SNP en cada uno de los 1351 caso y 7966 controles.

El conjunto de datos anterior se usó para atribuir probabilidades de genotipo para cada SNP CEU polimórfico en el HapMap en fase II, mediante el programa IMPUTE (www.stats.ox.ac.uk/~marchini/software/gwas/impute.html). Se usó el tamaño de población eficaz recomendado (-Ne 11418).

La asociación entre el estado de SLE y cada SNP atribuido se calculó con el programa SNPTTEST (www.stats.ox.ac.uk/~marchini/software/gwas/snptest.html). Se excluyeron los valores atípicos de la población; se determinaron con el programa EIGENSTRAT, de una manera similar a la descrita anteriormente en los ejemplos 1 y 2. Se ensayaron modelos de frecuencia tanto aditivos como generales.

Referencias

1. Hochberg MC. The epidemiology of systemic lupus erythematosus. In: Wallace DJ, Hahn BH, eds. Dubois' Lupus Erythematosus. 5.^a ed. Baltimore: Williams and Wilkins; 1997.
2. Wakeland EK, Liu K, Graham RR, Behrens TW. Delineating the genetic basis of systemic lupus erythematosus. *Immunity* 2001;15(3):397-408.
3. Nath SK, Kilpatrick J, Harley JB. Genetics of human systemic lupus erythematosus: the emerging picture. *Curr Opin Immunol* 2004;16(6):794-800.
4. Goldberg MA, Arnett FC, Bias WB, Shulman LE. Histocompatibility antigens in systemic lupus erythematosus. *Arthritis Rheum* 1976;19(2):129-32.
5. Graham RR, Ortmann WA, Langefeld CD, *et al*. Visualizing human leukocyte antigen class II risk haplotypes in

human systemic lupus erythematosus. *Am J Hum Genet* 2002; 71(3):543-53.

6. Graham RR, Ortmann W, Rodine P, *et al.* Specific combinations of HLA-DR2 and DR3 class II haplotypes contribute graded risk for disease susceptibility and autoantibodies in human SLE. *Eur J Hum Genet* 2007; 15(8):823-30.

7. Sigurdsson S, Nordmark G, Goring HH, *et al.* Polymorphisms in the tyrosine kinase 2 and interferon regulatory factor 5 genes are associated with systemic lupus erythematosus. *Am J Hum Genet* 2005; 76(3):528-37.

8. Graham RR, Kozyrev SV, Baechler EC, *et al.* A common haplotype of interferon regulatory factor 5 (IRF5) regulates splicing and expression and is associated with increased risk of systemic lupus erythematosus. *Nat Genet* 2006; 38(5):550-55.

9. Graham RR, Kyogoku C, Sigurdsson S, *et al.* Three functional variants of IFN regulatory factor 5 (IRF5) define risk and protective haplotypes for human lupus. *Proc Natl Acad Sci USA*. 2007; 104(16):6758-63.

10. Remmers EF, Plenge RM, Lee AT, *et al.* STAT4 and the risk of rheumatoid arthritis and systemic lupus erythematosus. *N Engl J Med* 2007; 357(10):977-86.

11. Ronnblom L, Alm GV. A pivotal role for the natural interferon alpha-producing cells (plasmacytoid dendritic cells) in the pathogenesis of lupus. *J Exp Med* 2001; 194(12):F59-63.

12. Baechler EC, Gregersen PK, Behrens TW. The emerging role of interferon in human systemic lupus erythematosus. *Curr Opin Immunol* 2004; 16(6):801-07.

13. Banchereau J, Pascual V. Type I interferon in systemic lupus erythematosus and other autoimmune diseases. *Immunity* 2006; 25(3):383-92.

14. Miyagi T, Gil MP, Wang X, Louten J, Chu WM, Biron CA. High basal STAT4 balanced by STAT1 induction to control type 1 interferon effects in natural killer cells. *J Exp Med* 2007; publicación electrónica; 10 de septiembre.

15. A haplotype map of the human genome. *Nature* 2005; 437(7063):1299-320.

16. Dewan A, Liu M, Hartman S, *et al.* HTRA1 promoter polymorphism in wet age-related macular degeneration. *Science* 2006; 314(5801):989-92.

17. Genome-wide association study of 14 000 cases of seven common diseases and 3 000 shared controls. *Nature* 2007; 447(7145):661-78.

18. Matarin M, Brown WM, Scholz S, *et al.* A genome-wide genotyping study in patients with ischaemic stroke: initial analysis and data release. *Lancet neurology* 2007; 6(5):414-20.

19. Moffatt MF, Kabesch M, Liang L, *et al.* Genetic variants regulating ORMDL3 expression contribute to the risk of childhood asthma. *Nature* 2007; 448(7152):470-73.

20. Plenge RM, Seielstad M, Padyukov L, *et al.* TRAF1-C5 as a Risk Locus for Rheumatoid Arthritis - A Genomewide Study. *N Engl J Med* 2007.

21. Saxena R, Voight BF, Lyssenko V, *et al.* Genome-wide association analysis identifies loci for type 2 diabetes and triglyceride levels. *Science* 2007; 316(5829):1331-36.

22. Scott LJ, Mohlke KL, Bonnycastle LL, *et al.* A genome-wide association study of type 2 diabetes in Finns detects multiple susceptibility variants. *Science* 2007; 316(5829):1341-45.

23. Scuteri A, Sanna S, Chen WM, *et al.* Genome-Wide Association Scan Shows Genetic Variants in the FTO Gene Are Associated with Obesity-Related Traits. *PLoS Genet* 2007; 3(7):e115.

24. Pe'er I, de Bakker PI, Maller J, Yelensky R, Altshuler D, Daly MJ. Evaluating and improving power in whole-genome association studies using fixed marker sets. *Nat Genet* 2006; 38(6):663-67.

25. Bauer JW, Baechler EC, Petri M, *et al.* Elevated serum levels of interferon-regulated chemokines are biomarkers for active human systemic lupus erythematosus. *PLoS medicine* 2006; 3(12):e491.

26. Criswell LA, Pfeiffer KA, Lum RF, *et al.* Analysis of families in the multiple autoimmune disease genetics consortium (MADGC) collection: the PTPN22 620W allele associates with multiple autoimmune phenotypes. *Am J Hum Genet* 2005; 76(4):561-71.

27. Seligman VA, Suarez C, Lum R, *et al.* The Fcγ receptor IIIA-158F allele is a major risk factor for the development of lupus nephritis among Caucasians but not non-Caucasians. *Arthritis Rheum* 2001; 44(3):618-25.

28. Demirci FY, Manzi S, Ramsey-Goldman R, *et al.* Association of a common interferon regulatory factor 5 (IRF5) variant with increased risk of systemic lupus erythematosus (SLE). *Ann Hum Genet* 2007; 71(Pt 3):308-11.

29. Hochberg MC. Updating the American College of Rheumatology revised criteria for the classification of systemic lupus erythematosus. *Arthritis Rheum* 1997; 40(9):1725.

30. Mitchell MK, Gregersen PK, Johnson S, Parsons R, Vlahov D. The New York Cancer Project: rationale, organization, design, and baseline characteristics. *J Urban Health* 2004; 81(2):301-10.

31. Gunderson KL, Steemers FJ, Ren H, *et al.* Whole-genome genotyping. *Methods Enzymol* 2006; 410:359-76.

32. Hsu TM, Chen X, Duan S, Miller RD, Kwok PY. Universal SNP genotyping assay with fluorescence polarization detection. *Biotechniques* 2001; 31(3):560.

33. Price AL, Patterson NJ, Plenge RM, Weinblatt ME, Shadick NA, Reich D. Principal components analysis corrects for stratification in genome-wide association studies. *Nat Genet* 2006;38(8): 904-09.

34. Purcell S, Neale B, Todd-Brown K, *et al.* PLINK: a tool set for whole-genome association and population-based linkage analyses. *Am J Hum Genet* 2007; 81(3):559-75.

35. Devlin B, Roeder K, Wasserman L. Genomic control for association studies: a semiparametric test to detect excess-haplotype sharing. *Biostatistics* 2000;1(4): 369-87.

36. Stranger BE, Nica AC, Forrest MS, *et al.* Population genomics of human gene expression. *Nat Genet* 2007; 39(10):1217-24.

37. Dixon AL, Liang L, Moffatt MF, *et al.* A genome-wide association study of global gene expression. *Nat Genet* 2007; 39(10):1202-7.

38. de Bakker PI, McVean G, Sabeti PC, *et al.* A high-resolution HLA and SNP haplotype map for disease association studies in the extended human MHC. *Nat Genet* 2006; 38(10):1166-72.
39. Hirschhorn JN, Daly MJ. Genome-wide association studies for common diseases and complex traits. *Nature reviews* 2005; 6(2):95-108.
40. Cheung VG, Spielman RS, Ewens KG, Weber TM, Morley M, Burdick JT. Mapping determinants of human gene expression by regional and genome-wide association. *Nature* 2005; 437(7063):1365-69.
41. Dymecki SM, Zwollo P, Zeller K, Kuhajda FP, Desiderio SV. Structure and developmental regulation of the Blymphoid tyrosine kinase gene blk. *J Biol Chem* 1992; 267(7):4815-23.
42. Wasserman R, Li YS, Hardy RR. Differential expression of the blk and ret tyrosine kinases during B lineage development is dependent on Ig rearrangement. *J Immunol* 1995; 155(2):644-51.
43. Texido G, Su IH, Mecklenbrauker I, *et al.* The B-cell-specific Src-family kinase Blk is dispensable for B-cell development and activation. *Mol Cell Biol* 2000;20(4): 1227-33.
44. Tsukada S, Saffran DC, Rawlings DJ, *et al.* Deficient expression of a B cell cytoplasmic tyrosine kinase in human X-linked agammaglobulinemia. *Cell* 1993; 72(2):279-90.
45. Khan WN, Alt FW, Gerstein RM, *et al.* Defective B cell development and function in Btk-deficient mice. *Immunity* 1995; 3(3):283-99.
46. Cornall RJ, Goodnow CC. B cell antigen receptor signalling in the balance of tolerance and immunity. *Novartis Foundation symposium* 1998; 215:21-30.
47. Nemazee D, Weigert M. Revising B cell receptors. *J Exp Med* 2000; 191(11):1813-7.
48. Kumar KR, Li L, Yan M, *et al.* Regulation of B cell tolerance by the lupus susceptibility gene Ly108. *Science* 2006; 312(5780):1665-9.
49. Abbas AR, Baldwin D, Ma Y, *et al.* Immune response in silico (IRIS): immune-specific genes identified from a compendium of microarray expression data. *Genes Immun* 2005; 6(4):319-31.
50. Hynes RO. Integrins: versatility, modulation, and signaling in cell adhesion. *Cell* 1992; 69(1):11-25.
51. Lu H, Smith CW, Perrard J, *et al.* LFA-1 is sufficient in mediating neutrophil emigration in Mac-1-deficient mice. *J Clin Invest* 1997; 99(6):1340-50.
52. Dunne JL, Collins RG, Beaudet *et al.*, Ballantyne CM, Ley K. Mac-1, but not LFA-1, uses intercellular adhesion molecule-1 to mediate slow leukocyte rolling in TNF-alpha-induced inflammation. *J Immunol* 2003; 171(11):6105-10 y las tablas 1-6.
53. Hammerberg C, Katiyar SK, Carroll MC, Cooper KD. Activated complement component 3 (C3) is required for ultraviolet induction of immunosuppression and antigenic tolerance. *J Exp Med* 1998; 187(7):1133-38.
54. Sohn JH, Bora PS, Suk HJ, Molina H, Kaplan HJ, Bora NS. Tolerance is dependent on complement C3 fragment iC3b binding to antigen-presenting cells. *Nat Med* 2003; 9(2):206-12.
55. Watts GM, Beurskens FJ, Martin-Padura I, *et al.* Manifestations of inflammatory arthritis are critically dependent on LFA-1. *J Immunol* 2005;174(6): 3668-75.
56. Ehrchiou D, Xiong Y, Xu G, Chen W, Shi Y, Zhang L. CD11b facilitates the development of peripheral tolerance by suppressing Th17 differentiation. *J Exp Med* 2007; 204(7):1519-24.
57. Buyon JP, Shadick N, Berkman R, *et al.* Surface expression of Gp 165/95, the complement receptor CR3, as a marker of disease activity in systemic Lupus erythematosus. *Clin Immunol Immunopathol* 1988; 46(1):141-49.
58. Giglio S, Broman KW, Matsumoto N, *et al.* Olfactory receptor-gene clusters, genomic-inversion polymorphisms, and common chromosome rearrangements. *Am J Hum Genet* 2001; 68(4):874-83.
59. Sugawara H, Harada N, Ida T, *et al.* Complex low-copy repeats associated with a common polymorphic inversion at human chromosome 8p23. *Genomics* 2003; 82(2):238-44.

Ejemplo 4 exploración de asociación de todo el genoma en 1310 casos de SLE y 7859 controles

Métodos: información de muestra y genotipado de casos de SLE y controles

La selección y genotipado de las muestras de casos de SLE se describieron previamente (1). En resumen, se genotiparon muestras de ADN de a) 338 sujetos de la Autoimmune Biomarkers Collaborative Network (ABCoN), un depósito fundado por el NIH/NIAMS (2), b) 141 sujetos del Multiple Autoimmune Disease Genetics Consortium (MADGC) (3), (ABCON + MADGC = serie de casos 1), c) 613 sujetos del University of California San Francisco (UCSF) Lupus Genetics Project (4, 5) (serie de casos 2) y d) 335 sujetos del University of Pittsburgh Medical Center (UPMC) (6) más 8 muestras recogidas en The Feinstein Institute for Medical Research (serie de casos 3) usando la matriz Illumina 550K. Todos los casos de SLE eran norteamericanos de ascendencia europea, determinada por autodenominación. El diagnóstico de SLE (cumplimiento de cuatro o más de los criterios definidos del American College of Rheumatology (ACR) (7)) se confirmó en todos los casos por revisión del registro (94 %) o a través de documentación escrita de los criterios de los reumatólogos a cargo (6 %). Los datos clínicos para estas series de casos se presentaron en otra parte (4, 3, 2, 6, 5).

Se examinó un total de 8147 muestras de control genotipadas usando la matriz Illumina 550K en los análisis de asociación. Se usaron tres fuentes para los controles (todos norteamericanos de ascendencia europea: 1861 muestras de la colección del New York Health Project (NYHP) (8); 1722 muestras de la base de datos disponible al público iControlDB (www.illumina.com/pages.ilmn?ID=231); y 4564 muestras del proyecto disponible al público Cancer Genetics Markers of Susceptibility (CGEMS) (<http://cgems.cancer.gov/>). El genotipado de las muestras NYHP se describió previamente (1).

Filtros de calidad de datos de genotipo

Se realizó un filtrado de muestras y SNP usando módulos analíticos dentro de los programas informáticos PLINK (9) y EIGENSTRAT (10), como se describe a continuación.

a) Casos de SLE, muestras NYCP y muestras iControlDB

Se usó la matriz de SNP Illumina 550K SNP, versión 1 (HH550v1) para genotipar 464 casos y 1962 controles y se usó la matriz de SNP Illumina 550K SNP, versión 3 (HH550v3) para genotipar 971 casos y 1621 controles como se describe (1). Las muestras en las que el género indicado no coincidía con el género observado (HH550v1: 10, HH550v3: 11) y las muestras con > 5 % de genotipos perdidos (HH550v1: 25, HH550v3: 21) se excluyeron del análisis. La relación críptica entre los casos de SLE y los controles se determinó por la estimación de la identidad por estado (IBS) a través del genoma para todas las posibles combinaciones de muestras por pares. Una muestra de cada par estimado como duplicado o familiar de 1^{er}-3^{er} grado se excluyó ($P_i\text{-hat} \geq 0,10$ y $Z1 \geq 0,15$; HH550v1: 88, HH550v3: 73). Los SNP con HWE $P \leq 1 \times 10^{-6}$ en los controles (HH550v1: 3176, HH550v3: 2240) y los SNP con > 5 % de datos perdidos (HH550v1: 12605, HH550v3: 7137) se eliminaron. Los SNP se ensayaron para una diferencia significativa en la frecuencia de datos perdidos entre los casos y los controles, y los SNP con $P \leq 1 \times 10^{-5}$ en el ensayo de valores faltantes diferencial implementado en PLINK se eliminaron (HH550v1: 5027, HH550v3: 2804). Los SNP también se ensayaron para una diferencia de frecuencia alélica significativa entre los géneros; todos los SNP tenían $P \geq 1 \times 10^{-9}$ en los controles. Los datos se examinaron para la presencia de los efectos de lote (por ejemplo, entre muestras ABCoN y todos los demás casos) y los SNP con una diferencia de frecuencia alélica con un $P < 1 \times 10^{-9}$ se excluyeron (HH550v1: 18, HH550v3: 10). Las variantes con genotipos haploides heterocigóticos se establecieron a pérdida (HH550v1: 2305, HH550v3: 875). Además, se eliminaron las variantes con una frecuencia alélica minoritaria < 0,0001 removed (HH550v1: 97, HH550v3: 57).

b) Muestras CGEMS

Para las 2277 muestras de cáncer de próstata y, por separado, 2287 muestras de cáncer de mama, los genotipos haploides heterocigóticos se establecieron a pérdida (próstata: 2717, mama: 0). Las muestras en las que el género presentado no coincidía con el género observado (próstata: 0, mama: 2) y las muestras con > 5 % de datos perdidos (próstata: 15, mama: 1) se excluyeron. Las muestras se ensayaron para la relación críptica, como se describe anteriormente y se eliminó una muestra de cada par estimado como duplicado o familiar de 1^{er}-3^{er} grado ($P_i\text{-hat} \geq 0,10$ y $Z1 \geq 0,15$; próstata: 12, mama: 7). Se eliminaron los SNP con un MAF < 0,0001 (próstata: 3254, mama: 2166).

c) Todas las muestras

Se aplicaron filtros adicionales de calidad de datos al conjunto de datos combinado que consistía en todos los casos de SLE y controles. Se eliminaron los SNP con > 5 % de datos perdidos (N=65 421) y las muestras con > 5 % de datos perdidos (N=0). Se realizó un ensayo para muestras duplicadas usando 957 SNP independientes con MAF $\geq 0,45$ y se no se encontraron muestras duplicadas. Se eliminaron los SNP con HWE $P \leq 1 \times 10^{-6}$ en los controles (N=2174) y los SNP con > 2 % de datos perdidos (N=5522). Se ensayaron los SNP para una diferencia significativa en la proporción de datos perdidos entre los casos y los controles y se eliminaron los SNP con un diferencial de exceso de datos perdidos ($P \leq 1 \times 10^{-5}$, N=16 080). Los SNP se ensayaron para una diferencia significativa entre los géneros y todos los SNP tuvieron $P \geq 1 \times 10^{-9}$ en los controles. También se examinaron los SNP para la presencia de los efectos de lote; en particular, entre las muestras de cáncer de mama CGEMS y todos los demás controles, y entre las muestras de cáncer de próstata CGEMS y todos los demás controles y se eliminaron los SNP con $P < 1 \times 10^{-9}$ (N=73). Después de la aplicación de los filtros de calidad anteriores, permanecieron 480 831 SNP.

Los casos y los controles se ensayaron para la presencia de datos atípicos de la población usando EIGENSTRAT. Se excluyeron los SNP con MAF < 2 % en los casos (N=16 068), HWE $P \leq 1 \times 10^{-4}$ en los controles (N=977) o > 1 % de datos perdidos (N=17 029); los SNP en regiones de patrones de LD anómalos debido a variación estructural en los cromosomas 6 (de 24-36 Mb), 8 (8-12 Mb), 11 (42-58 Mb) y 17 (40-43 Mb); y los SNP en la región pseudoautosómica del cromosoma X (N=12) con el fin de determinar los componentes principales (EIGENSTRAT) de variación para detectar los datos atípicos de la población. Las muestras con más de 6 desviaciones típicas de la media a lo largo de cualquiera de los 10 componentes principales superiores se eliminaron (N=148).

El conjunto de datos final tenía 1310 casos, 7859 controles y 480 831 SNP, y el factor de inflación de control genómico (λ_{gc}) (11) fue de 1,06 después de la aplicación de los filtros de calidad de datos anteriores.

Atribución de genotipos no observados

El desequilibrio del ligamiento extenso presente en el genoma humano permite la interferencia de variantes no tipadas en ciertas situaciones con un alto grado de confianza. IMPUTE, un programa para atribuir genotipos no observados en estudios de casos-controles de todo el genoma basados en un conjunto de haplotipos conocidos (haplotipos HapMap de fase II, www.hapmap.org), se usó en el análisis

(www.stats.ox.ac.uk/~marchini/software/gwas/impute.html).

Atribución de los casos GNE y los controles NYCP, iDB y CGEMS

Después de los filtros de control de calidad, había 1310 casos GNE, controles 3344 NYCP e iDB, 4515 controles CGEMS y 446 856 SNP. El programa IMPUTE (v0.3.1) se ejecutó con el haplotipo CEU, leyendas y archivos de mapa incluidos alineados para NCBI construcción 35. El tamaño de población eficaz se estableció al valor recomendado de 11418. No se usó ningún archivo de hebra; se activó la comprobación de alineación de hebras en IMPUTE. Los casos, los controles NYCP e iDB, y los controles CGEMS se atribuyeron por separado, y cada cromosoma se atribuyó por separado en su totalidad. Se atribuyeron 2 562 708 SNP.

Se usó SNPTTEST (v1.1.3) para hacer ensayos de asociación en genotipos tanto reales como atribuidos. Para los SNP que ya estaban genotipados, se usaron los genotipos reales. El ensayo de asociación fue el ensayo de Cochran-Armitage para un efecto genético aditivo, con la opción "apropiado" para tener en cuenta completamente la inexactitud de los genotipos. Se mantuvieron únicamente los SNP con un valor de información por encima de 0,50 (es decir, $\text{frequentist_add_proper_info} > 0,50$) (2 481 907 SNP [97 %]).

Resultados. Se presenta una lista no redundante de los loci de SLE asociados con SLE ($P < 1 \times 10^{-5}$) en el análisis de 1310 casos y 7859 controles en la tabla 1. Se generó la lista ordenada por rango presentando la variante única con el valor P más bajo en un intervalo de +/- 100 kb de la generada presentando la variante única con el valor P más bajo en un intervalo de +/- 100 kb del análisis de 2,3 millones de SNP como se describe anteriormente.

Tabla 1: se muestran loci asociados con SLE ($P \leq 1 \times 10^{-5}$) en el alelo de 1310 casos y 7859 controles. La lista ordenada por rango se generó presentando la variante única con el valor P más bajo en un intervalo de +/- 100 kb del análisis de 2,3 millones de SNP como se describe. El SNP (dbSNP id), cromosoma, posición (posición de los pares de bases en la construcción 35 del genoma humano), la frecuencia de alelos minoritarios en los casos de SLE y controles, el valor P de SNPTTEST (en un modelo aditivo, que corrige la precisión de atribución), el valor de información de atribución (una estimación de la precisión de atribución) y la relación de probabilidad (con intervalos de confianza del 95 %).

Frecuencia alélica							
SNP	Cromosoma	Posición (construcción 35)	Casos (N=1310)	Controles (N=7859)	P	Valor de información de atribución	Proporción de probabilidad (95 % c.i.)
rs2187668	6	32713862	0,190	0,117	2,49E-24	1,00	1,76 (1,58-1,97)
rs13236009	7	128257124	0,175	0,111	2,37E-20	0,96	0,59 (0,52-0,66)
rs11889341	2	191769248	0,308	0,230	1,55E-19	0,97	1,49 (1,36-1,64)
rs6565228	16	31236781	0,041	0,029	1,40E-11	0,69	0,71 (0,56-0,91)
rs2736345	8	11389894	0,335	0,278	3,80E-09	0,96	1,31 (1,2-1,44)
rs6889239	5	150437964	0,300	0,251	1,21E-07	1,00	0,78 (0,72-0,86)
rs2391592	7	27983196	0,531	0,472	2,34E-07	0,92	0,79 (0,72-0,86)
rs2177770	2	141630291	0,086	0,064	2,54E-07	0,72	1,37 (1,15-1,63)
rs12039904	1	169943930	0,283	0,238	8,66E-07	0,95	1,26 (1,15-1,39)
rs4591368	2	71637413	0,008	0,004	1,03E-06	0,57	0,5 (0,3-0,84)
rs12882608	14	82621870	0,205	0,170	1,10E-06	0,90	0,79 (0,71-0,89)
rs3024493	1	203332363	0,187	0,148	1,96E-06	0,96	0,76 (0,68-0,85)
rs11678272	2	42061217	0,313	0,270	2,38E-06	0,98	0,81 (0,74-0,89)
rs874952	2	65520176	0,121	0,157	2,77E-06	1,00	1,35 (1,19-1,53)
rs2053482	8	98289410	0,076	0,054	3,01E-06	0,93	0,7 (0,6-0,83)
rs2431697	5	159812556	0,389	0,438	3,50E-06	1,00	1,22 (1,12-1,33)

rs6679677	1	114015850	0,107	0,079	3,55E-06	0,91	0,72 (0,62-0,83)
rs12445476	16	84548770	0,158	0,196	4,11E-06	0,99	1,3 (1,16-1,46)
rs2208384	1	232216137	0,212	0,252	4,43E-06	0,99	1,26 (1,14-1,39)
rs6879995	5	158447777	0,353	0,304	4,47E-06	0,92	0,8 (0,73-0,88)
rs10502821	18	39720893	0,003	0,001	4,54E-06	0,89	0,18 (0,07-0,45)
rs3790565	1	67523377	0,225	0,187	4,84E-06	1,00	0,79 (0,71-0,87)
rs2024831	6	14822843	0,218	0,185	5,17E-06	0,92	0,81 (0,73-0,9)
rs2066943	4	85247083	0,013	0,026	5,68E-06	0,61	0,5 (0,34-0,74)
rs12986652	2	180300843	0,112	0,087	5,69E-06	0,96	0,75 (0,66-0,86)
rs1196592	18	34246756	0,356	0,312	5,77E-06	1,00	0,82 (0,75-0,89)
rs7759216	6	106695307	0,425	0,377	5,84E-06	1,00	1,22 (1,12-1,33)
rs17484292	1	180031707	0,056	0,040	5,97E-06	0,75	0,7 (0,57-0,85)
rs1579289	5	107837653	0,134	0,168	6,44E-06	0,94	1,31 (1,15-1,48)
rs11970105	6	49429182	0,149	0,117	6,65E-06	0,89	0,76 (0,67-0,86)
rs10082917	12	40285103	0,390	0,346	6,93E-06	0,94	1,21 (1,1-1,32)
rs7006016	8	29655999	0,110	0,087	8,02E-06	0,80	1,29 (1,11-1,5)
rs11757479	6	114612099	0,059	0,040	8,63E-06	0,84	0,67 (0,55-0,81)
rs4968210	17	7398076	0,414	0,368	9,64E-06	0,94	0,83 (0,76-0,9)

Ejemplo 5. Metanálisis de loci de riesgo de SLE presentados en la exploración de asociación de GNE

Métodos

Examen de la bibliografía de SLE y los criterios para el loci de SLE confirmados

Un total de 16 alelos cumplían uno de los criterios descritos a continuación para los loci de riesgo de SLE confirmados (tabla 2).

1) Loci de riesgo de SLE con al menos 2 informes independientes de $P \leq 1 \times 10^{-5}$.

Se examinó la bibliografía para loci con 2 informes independientes en cohortes de SLE no solapantes con un $P \leq 1 \times 10^{-5}$. La investigación de la bibliografía representa publicaciones anteriores a abril de 2008. Se requirió la variante idéntica (o sustituto con $r^2 > 0,3$) que muestra asociación a SLE con la misma dirección de efecto. Un total de 7 alelos cumplió los requisitos, incluyendo HLA-DRB1*0301 (HLA-DR3,(18, 19)), HLA-DRB1*1501 (HLA-DR2,(18, 19)), proteína tirosina fosfatasa no receptora de tipo 22 (PTPN22, (20, 21)), factor 5 regulador de interferón (IRF5, (22, 23)), transductor de señales y activador de la transcripción 4 (STAT4, (5, 21)), tirosina cinasa de linfocitos B (BLK, (21, 1)) e integrina alfa M (ITGAM, (1, 24)). El alelo idéntico o mejor sustituto ($r^2 > 0,85$) en la exploración de asociación de todo el genoma de 1310 casos de SLE y 7859 controles descrito en esta ocasión se hizo avanzar en el análisis (tabla 2).

2) Loci de riesgo de SLE con un único informe de $P \leq 1 \times 10^{-5}$.

Se realizó una investigación de bibliografía para loci de riesgo de SLE con un $P \leq 1 \times 10^{-5}$ presentado en una única publicación desde abril de 2008 y se identificó un total de 18 loci.

En 13 de los loci, se genotipó la variante idéntica o sustituto casi perfecto ($r^2 > 0,9$) en la exploración de genoma de 1310 casos de SLE y 7859 controles descrita anteriormente (tabla 4). Se realizó un metanálisis usando la metodología descrita a continuación para los 13 loci, y 8 de los loci consiguieron un $P \leq 5 \times 10^{-8}$. Los loci (marcados

por un único gen dentro del locus) que consiguen significancia de todo el genoma incluyen la proteína 1 transformante de tumor de la pituitaria (*PTTG1*), APG5 de autofagia de tipo 5 (*ATG5*), proteína rA9 de tipo SR de unión a CTD (*KIAA1542*), enzima E2L3 de conjugación a ubiquitina (*UBE2L3*), dominio PX que contiene serina/treonina cinasa (*PXK*), fragmento Fc de IgG, receptor IIa de baja afinidad (*FCGR2A*), superfamilia 4 del factor de necrosis tumoral (ligando) (*TNFSF4*) y la proteína estructural de linfocitos B con repeticiones de anquirina 1 (*BANK1*). La variante que alcanza significancia en todo el genoma en el metanálisis se hizo avanzar en el análisis (tabla 5, tabla 2). En los 5 loci restantes, la variante presentada o sustituto casi perfecto ($r^2 > 0,9$) no se genotipó en la exploración de asociación de todo el genoma de 1310 casos de SLE y 7859 controles SLE (tabla 2). Sin embargo, una variante en la cinasa 1 asociada al receptor de interleucina-1 (*IRAK1*) obtuvo un $P \leq 1 \times 10^{-4}$ observado y se hizo avanzar en el análisis (tabla 1).

Metanálisis

La estadística de asociación corregida para cada serie se combinó por la suma de los valores Z ponderados para el tamaño de la cohorte.

Tabla 2. Estadística de asociación para 16 alelos de riesgo de SLE confirmados en nuestro GWAS de 1310 casos de SLE y 7859 controles. Los alelos están por valor P.

Locus	Cromosoma	SNP	Frecuencia de alelo		Caso	Control	Valor P	Proporción de probabilidad (95% CI)
			Posición* (Mb)	Alelo minoritario				
<i>HLA-DR3</i>	6p21.32	rs2187668 rs1048863	32,714 128,18	T	0,190	0,117	$9,5 \times 10^{-25}$	1,76 (1,58-1,97)
<i>IRF5</i>	7q32.1	1	8 191,79	C	0,170	0,109	$1,4 \times 10^{-19}$	1,68 (1,50-1,89)
<i>STAT4</i>	2q32.2	rs7574865	0	T	0,312	0,235	$2,5 \times 10^{-14}$	1,48 (1,34-1,64)
<i>ITGAM</i>	16p11.2	rs9888739 rs1327711	31,221	T	0,175	0,127	$2,3 \times 10^{-11}$	1,46 (1,31-1,63)
<i>BLK</i>	8p23.1	3	11,387 159,81	A	0,294	0,242	$1,7 \times 10^{-8}$	1,30 (1,19-1,43)
<i>PTTG1</i>	5q33.3	rs2431697 rs6568431	3 106,69	C	0,389	0,438	$3,3 \times 10^{-6}$	0,82 (0,75-0,89)
<i>ATG5</i>	6q21	rs1048926	5 169,96	A	0,423	0,376	$5,5 \times 10^{-6}$	1,22 (1,12-1,32)
<i>TNFSF4</i>	1q25.1	5	8 114,09	C	0,278	0,238	$8,7 \times 10^{-6}$	1,24 (1,09-1,30)
<i>PTPN22</i>	1p13.2	rs2476601	0 152,71	A	0,116	0,089	$8,9 \times 10^{-6}$	1,35 (1,18-1,54)
	Xq28	rs2269368	1 158,29	T	0,175	0,141	$1,1 \times 10^{-5}$	1,29 (1,15-1,45)
<i>FCGR2A</i>	1q23.3	rs1801274	3	A	0,463	0,500	$4,1 \times 10^{-4}$	0,86 (0,79-0,94)
	11p15.5	rs4963128	0,580	T	0,303	0,333	$3,1 \times 10^{-3}$	0,87 (0,80-0,96)
<i>UBE2L3</i>	22q11.21	rs5754217	20,264	T	0,215	0,192	$6,4 \times 10^{-3}$	1,15 (1,04-1,27)
<i>PXK</i>	3p14.3	rs6445975	58,345	G	0,305	0,281	0,010	1,13 (1,03-1,23)
<i>HLA-DR2</i>	6p21.32	rs3129860 rs1051648	32,509 103,10	A	0,160	0,147	0,092	1,10 (0,98-1,24)
<i>BANK1</i>	4q24	7	8	A	0,288	0,304	0,096	0,93 (0,85-1,01)

*Las posiciones son de NCBI construcción 35.

Tabla 3. Loci de riesgo de SLE en al menos 2 informes independientes con el mismo SNP (o sustituto con $r^2 > 0,3$) con $P \leq 1 \times 10^{-5}$.

Locus	Cromosoma	Informe 1			Informe 2			GNE GWAS	
		Alelo	Valor P	Referencia	Alelo	r^2 para el alelo en el informe 1	Valor P	Referencia adicionales	SNP
<i>PTPN22</i>	1p13.2	rs2476601	$1,0 \times 10^{-5}$	(20)	rs2476601	1,00	$5,2 \times 10^{-6}$	(21)	rs2476601
<i>STAT4</i> <i>HLA-DR2</i>	2q32.2	rs7574865	$1,9 \times 10^{-9}$	(5)	rs7574865	1,00	$2,8 \times 10^{-9}$	(1)	rs7574865
		DRB1*1501	$1,0 \times 10^{-5}$	(18)	DRB1*501	1,00	$1,0 \times 10^{-7}$	(26)	rs3129860
<i>HLA-DR3</i>	6p21.32	DRB1*0301	$1,0 \times 10^{-6}$	(18)	DRB1*0301	1,00	$1,0 \times 10^{-5}$	(21), (1), (26), (27)	rs2187668
<i>IRF5</i>	7q32.1	rs2004640	$5,2 \times 10^{-8}$	(22)	rs2004640	1,00	$4,4 \times 10^{-16}$	(23)	rs10488631
<i>BLK</i>	8p23.1	rs13277113	$1,0 \times 10^{-10}$	(1)	rs6985109	0,33	$2,5 \times 10^{-11}$	(21)	rs13277113
<i>ITGAM</i>	16p11.2	rs1143679	$6,9 \times 10^{-22}$	(24)	rs11574637	-	$3,0 \times 10^{-11}$	(1)	rs9888739
								(21)	
									0,86

Tabla 4. Loci de riesgo de SLE presentados únicamente una vez con $P \leq 1 \times 10^{-5}$ y en que era posible el metanálisis. Ocho loci obtuvieron un meta $P \leq 5 \times 10^{-8}$.

Locus	Cromosoma	Informe		Referencia	GNE GWAS*			Meta P
		Alelo	Valor P		SNP	r^2 para alelo en el informe	Valor P	
<i>PTTG1</i>	5q33.3	rs2431697	1,0 x 10^{-10}	(21)	rs2431697	1,00	3,3 x 10^{-6}	5,3 x 10^{-14}
<i>ATG5</i>	6q21	rs6568431	1,7 x 10^{-8}	(21)	rs6568431	1,00	5,5 x 10^{-6}	2,7 x 10^{-12}
	11p15.5	rs4963128	3,0 x 10^{-10}	(21)	rs4963128	1,00	3,1 x 10^{-3}	1,0 x 10^{-9}
<i>UBE2L3</i>	22q11.21	rs5754217	7,5 x 10^{-8}	(21)	rs5754217	1,00	6,4 x 10^{-3}	7,3 x 10^{-9}
<i>PXK</i>	3p14.3	rs6445975	7,1 x 10^{-9}	(21)	rs6445975	1,00	0,010	1,0 x 10^{-8}
<i>FCGR2A</i>	1q23.3	rs1801274	6,8 x 10^{-7}	(21)	rs1801274	1,00	4,1 x 10^{-4}	3,9 x 10^{-8}
<i>TNFSF4</i>	1q25.1	rs12039904	1,0 x 10^{-5}	(29)	rs10489265	0,91	8,7 x 10^{-6}	
<i>BANK1</i>	4q24	rs10516487	3,7 x 10^{-10}	(30)	rs10516487	1,00	0,096	--
<i>NMNAT2</i>	1q25.3	rs2022013	1,1 x 10^{-7}	(21)	rs2022013	1,00	0,15	5,1 x 10^{-6}
<i>ICA1</i>	7p21.3	rs10156091	1,9 x 10^{-7}	(21)	rs10156091	1,00	0,095	2,0 x 10^{-5}
<i>LYN</i>	8q12.1	rs7829816	5,4 x 10^{-9}	(21)	rs7829816	1,00	0,48	3,6 x 10^{-3}
<i>SCUBE1</i>	22q13.2	rs2071725	1,2 x 10^{-7}	(21)	rs2071725	1,00	0,63	8,3 x 10^{-3}
<i>ITPR3</i>	6p21.31	rs3748079	2,9 x 10^{-8}	(31)	rs3748079	1,00	0,95	--

*GNE GWAS: nuestro GWAS de 1310 casos de SLE y 7859 controles.

Tabla 5. Loci de riesgo de SLE presentados únicamente una vez con $P \leq 1 \times 10^{-5}$ y en los que no era posible el metanálisis. Únicamente *IRAK1* obtuvo un SNP con $P \leq 1 \times 10^{-5}$ en nuestro GNE GWAS.

Locus	Cromosoma	Alelo	Valor P	Referencia	Mejor SNP†	Valor P	SNP	Valor P
<i>IRAK1</i>	Xq28	rs10127175	$9,6 \times 10^{-6}$	(32)	rs2269368	$1,1 \times 10^{-5}$	rs2269368	$1,1 \times 10^{-5}$
<i>CRP</i>	1q23.2	rs3093061	$6,4 \times 10^{-7}$	(33)	rs3820099	$3,0 \times 10^{-3}$	-	-
<i>SELP</i>	1q24.2	rs3917815	$5,7 \times 10^{-6}$	(32)	rs9332628	$8,8 \times 10^{-3}$	-	-
<i>PDCD1</i>	2q37.3	rs11568821	$1,0 \times 10^{-5}$	(34)	rs3892357	0,84	-	-
<i>TYK2</i>	19p13.2	rs2304256	$2,2 \times 10^{-8}$	(22)	rs12720356	$2,7 \times 10^{-3}$	-	--

*GNE GWAS: nuestro GWAS de 1310 casos de SLE y 7859 controles.

†Mejor SNP: el SNP con el menor P en ese locus.

Ejemplo 6. Resumen*Resumen de loci de riesgo de SLE*

- 5 Se identificaron los loci de riesgo de SLE usando dos métodos principales-a) análisis de 1310 casos de SLE y 7859 controles, y b) un metanálisis con los loci de riesgo de SLE presentados previamente.

Se proporcionan una lista no redundante de las variantes con fuerte asociación a riesgo de SLE ($P < 1 \times 10^{-6}$) en la tabla 6.

Algoritmo para evaluar el riesgo de SLE y respuesta al tratamiento

Se sabe que las variantes asociadas con un fenotipo interactúan de una manera aditiva dependiente de la dosis alélica (38, 39). En una realización a modo de ejemplo, puede usarse el siguiente algoritmo para evaluar el riesgo a lupus, la gravedad de la enfermedad y la respuesta al tratamiento. Los casos de lupus pueden estratificarse en grupos basándose en el número de alelos de riesgo que portan. En esta realización a modo de ejemplo, el alelo de riesgo se identifica como el alelo enriquecido en casos de lupus respecto a los controles de los loci. Por ejemplo, en la tabla 6, hay un total de 19 alelos de 18 loci, haciendo que el número posible máximo de alelos de riesgo sea igual a 38. Pueden determinarse casos de lupus estratificados por el número de alelos de riesgo y terciles de la distribución resultante. Los terciles de casos de lupus entonces pueden examinarse para las diferencias en la gravedad de la enfermedad, el riesgo y respuesta al tratamiento.

Tabla 6: Loci de riesgo de lupus

Locus	Cromosoma	SNP	Posición* (Mb)	Valor P	Fuente
HLA-DR3	6p21.32	rs2187668	32,714	$9,5 \times 10^{-25}$	Tabla 2
IRF5	7q32.1	rs10488631	128,188	$1,4 \times 10^{-19}$	Tabla 2
STAT4	2q32.2	rs7574865	191,790	$2,5 \times 10^{-14}$	Tabla 2
ITGAM	16p11.2	rs9888739	31,221	$2,3 \times 10^{-11}$	Tabla 2
BLK	8p23.1	rs13277113	11,387	$1,7 \times 10^{-8}$	Tabla 2
PTTG1	5q33.3	rs2431697	159,813	$3,3 \times 10^{-6}$	Tabla 2
ATG5	6q21	rs6568431	106,695	$5,5 \times 10^{-6}$	Tabla 2
TNFSF4	1q25.1	rs10489265	169,968	$8,7 \times 10^{-6}$	Tabla 2
PTPN22	1p13.2	rs2476601	114,090	$8,9 \times 10^{-6}$	Tabla 2
	Xq28	rs2269368	152,711	$1,1 \times 10^{-5}$	Tabla 2
FCGR2A	1q23.3	rs1801274	158,293	$4,1 \times 10^{-4}$	Tabla 2
	11p15.5	rs4963128	0,580	$3,1 \times 10^{-3}$	Tabla 2
UBE2L3	22q11.21	rs5754217	20,264	$6,4 \times 10^{-3}$	Tabla 2
PXK	3p14.3	rs6445975	58,345	0,01	Tabla 2
HLA-DR2	6p21.32	rs3129860	32,509	0,092	Tabla 2
BANK1	4q24	rs10516487	103,108	0,096	Tabla 2
TNIP1	5	rs6889239	150,438	$2,2 \times 10^{-8}$	Tabla 1
JAZF1	7	rs2391592	27,983	$2,3 \times 10^{-7}$	Tabla 1
LRP1B	2	rs2177770	141,630	$2,5 \times 10^{-7}$	Tabla 1

* Posición cromosómica de la variante en pares de bases en NCBI construcción 35 (Hg17, mayo de 2004) del genoma humano (http://www.ncbi.nlm.nih.gov/genome/guide/human/release_notes.html#b35)

Referencias

1. G. Hom et al., N Engl J Med 358, 900 (2008).
2. J. W. Bauer et al., PLoS Med 3, e491 (2006).
3. L. A. Criswell et al., Am J Hum Genet 76, 561 (2005).
4. V. A. Seligman et al., Arthritis Rheum 44, 618 (2001).
5. E. F. Remmers et al., N Engl J Med 357, 977 (2007).
6. F. Y. Demirci et al., Ann Hum Genet 71,308 (2007).
7. M. C. Hochberg, Arthritis Rheum 40, 1725 (1997).
8. M. K. Mitchell, P. K. Gregersen, S. Johnson, R. Parsons, D. Vlahov, J Urban Health 81,301 (2004).
9. S. Purcell et al., Am J Hum Genet 81,559 (2007).
10. A. L. Price et al., Nat Genet 38, 904 (2006).
11. B. Devlin, K. Roeder, L. Wasserman, Biostatistics 1, 369 (2000).
12. R. R. Graham et al., Arthritis research 3, 299 (2001).
13. M. C. Hochberg, Arthritis and Rheumatism 40, 1725 (1997).
14. C. Wellcome Trust Case Control, Nature 447, 661 (2007).
15. S. Purcell.
16. S. Purcell et al., American Journal of Human Genetics 81,559 (2007).
17. A. L. Price et al., Nat Genet 38, 904 (2006).
18. K. Hartung, Baur, M.P., Coldewey, R., Fricke, M., Kalden, J.R., Lakomek, H.J., Peter, H.H., Schendel, D.,

- Schneider, P.M., Seuchter, S.A., Stangel, W., Deicher, H.R.G., J Clin. Invest. 90, 1346 (1992).
 19. Z. Yao et al., Eur J Immunogenet 20, 259 (1993).
 20. Y. H. Lee et al., Rheumatology (Oxford) 46, 49 (2007).
 21. J. B. Harley et al., Nat Genet 40, 204 (2008).
 5 22. S. Sigurdsson et al., Am J Hum Genet 76, 528 (2005).
 23. R. R. Graham et al., Nat Genet 38, 550 (2006).
 24. S. K. Nath et al., Nat Genet 40, 152 (2008).
 25. C. Kyogoku et al., Am J Hum Genet 75, 504 (2004).
 26. J. B. Harley, K. L. Moser, P. M. Gaffney, T. W. Behrens, Curr Opin Immunol 10, 690 (1998).
 10 27. M. M. Fernando et al., PLoS Genet 3, e192 (2007).
 28. R. R. Graham et al., Proc Natl Acad Sci U S A 104, 6758 (2007).
 29. D. S. Cunninghame Graham et al., Nature Genetics 40, 83 (2008).
 30. S. V. Kozyrev et al., Nat Genet 40, 211 (2008).
 31. T. Oishi et al., J Hum Genet 53, 151 (2008).
 15 32. C. O. Jacob et al., Arthritis Rheum 56, 4164 (2007).
 33. J. C. Edberg et al., Hum Mol Genet 17, 1147 (2008).
 34. L. Prokunina et al., Nat Genet 32, 666 (2002).
 35. S. Paabo, Nature 421, 409 (2003).
 36. K. A. Frazer et al., Nature 449, 851 (2007).
 20 37. P. I. de Bakker et al., Nat Genet 37, 1217 (2005).
 38. J. Maller et al., Nat Genet 38, 1055 (2006).
 39. G. Lettre et al., Nat Genet 40, 584 (2008).

Ejemplo 7. Resumen de secuenciación

Métodos

El ADN genómico de 192 pacientes con SLE y 96 controles sanos se amplificó en su genoma completo antes de volver a secuenciarlo. Se volvió a secuenciar el ADN genómico de todos los exones y las regiones no codificantes seleccionadas (2,5 kb de la región promotora-dirección 5' del exón 1) en la cinasa de linfocitos B (BLK), integrina alfa M (ITGAM) e integrina alfa X (ITGAX).

Se realizó una convocatoria de alelos inicial por el software proporcionado por "Polymorphic". Todos los polimorfismos codificantes, así como los alelos no codificantes comunes se verificaron manualmente para confirmar los éxitos de alelo y crear los archivos de genotipado, usados para el análisis de asociación y haplotipos.

Se proporcionan variantes de ITGAM/ITGAX en las tablas 7 y 9 y las tablas 8 y 10. Las variantes de las tablas 7 y 9 no están presentes en la base de datos dbSNP construcción 129. Las variantes de las tablas 8 y 10 se descubrieron por secuenciación de ITGAM/ITGAX y BLK.

Tabla 7: Variantes de los exones y la región promotora de ITGAM e ITGAX.

ID	Alelo minoritario	Cromosoma	Posición	Frecuencia de alelo		
				Alelo	Casos	Controles
exón8_Gln246Arg	A	16	31192218	A	0,0026	0,005
exón8_Glu247Lys	G	16	31192220	G	0,0026	0,005
237T>C_ex17	C	16	31243375	C	0,175	0,174
186G>A_ex20	A	16	31244226	G	0,652	0,62
exón26_Gly1003Glu	G	16	31248726	G	0,0026	0,005
3'UTR; inserción de 10 pb	Inserción [GAGTGTGTGC]	16	31250691	G	0,434	NaN
no codificante 1a 329T>C	C	16	31250736	C	0,201	NaN

Tabla 8: Variantes de los exones y la región promotora de ITGAM e ITGAX

ID	Alelo minoritario	Cromosoma	Posición	Frecuencia de alelo		
				Alelo	Casos	Controles
rs3764327	T	16	31180630	C	0,736	0,677
rs1143679	A	16	31184312	A	0,144	0,109
rs35314490	A	16	31190665	A	0,155	0,078
rs9939679	C	16	31195622	C	0,16	0,109
rs11861251	C	16	31196897	C	0,17	0,115
rs1143683	T	16	31244389	C	0,827	0,823
rs41321249	A	16	31248925	A	0	0,021
rs7188189	T	16	31250109	C	0,914	0,844
rs1143678	T	16	31250506	C	0,825	0,823
rs4594268	T	16	31250744	T	0,246	NaN
rs9933520	G	16	31250887	G	0,178	0,172

rs3087796	G	16	31251154	A	0,738	0,635
rs41523147	C	16	31251171	C	0,026	0,005
rs4597342	T	16	31251270	C	0,729	0,661
rs11574633	C	16	31274819	C	0,16	0,115
rs2230429	G	16	31282036	G	0,348	0,271
rs12448775	T	16	31292149	T	0,042	0,031
rs41419150	inserción [CTTTA]	16	31251462-31251462	T	0,176	0,102

Tabla 9: Variantes de los exones y la región promotora de BLK

ID	Alelo minoritario	Cromosoma	Posición	Frecuencia de alelo		
				Alelo	Casos	Controles
1120_C>T	T	8	11387925	T	0,021	0,005
434_C>T	T	8	11404452	T	0,144	0,12
ex5_112T>C	T	8	11443842	T	0,516	0,417
ex9_121T>C	T	8	11451476	C	0,829	0,798
ex11_6G>A	G	8	11456175	A	0,882	0,818
975_G>A	A	8	11456175	G	0,89	0,818
exón6_Trp131Arg	T	8	11445099	T	0,0034	0
exón8_Pro237Pro	T	8	11450340	T	0,0026	0,005
exón10_Thr325Lys	C	8	11452900	C	0,0034	0
exón13_Arg474Arg	T	8	11458929	T	0,0034	0

Tabla 10: Variantes de los exones y la región promotora de BLK

ID	Alelo minoritario	Cromosoma	Posición	Frecuencia de alelo		
				Alelo	Casos	Controles
rs10097015	T	8	11458793	T	0,419	0,358
rs1042689	T	8	11459203	T	0,377	0,323
rs1042701	A	8	11459455	G	0,558	0,51
rs11784016	T	8	11404079	C	0,717	0,667
rs1382567	C	8	11388309	T	0,521	0,484
rs1382568	C	8	11388630	C	0,317	0,198
rs1382568	G	8	11388631	A	0,524	0,49
rs2250788	A	8	11389466	G	0,838	0,771
rs2251056	C	8	11386986	A	0,84	0,781
rs2409782	C	8	11404502	C	0,238	0,234
rs2736344	T	8	11388088	C	0,861	0,776
rs2898289	A	8	11455795	A	0,38	0,307
rs4629826	C	8	11404447	G	0,937	0,917
rs4840568	A	8	11388429	A	0,327	0,214
rs4841557	A	8	11452981	A	0,416	0,328
rs4841558	C	8	11453006	C	0,413	0,328
rs4841561	T	8	11456182	T	0,384	0,307
rs4841561	T	8	11456183	T	0,369	0,307
rs55758736	A	8	11442984	A	0,016	0,021
rs56185487	A	8	11443008	A	0,005	0
rs7843987	C	8	11459540	T	0,555	0,521
rs922483	T	8	11389322	T	0,359	0,224
rs9694294	C	8	11388131	G	0,846	0,776

Ejemplo 8

Sujetos y diseño del estudio

Se realizó un estudio de asociación de todo el genoma para SLE. Se genotiparon 1079 casos de SLE y 1411 controles con el Illumina HumanHap550 Genotyping BeadChip (555 352 SNP). Los casos de SLE eran de tres cohortes distintas. Las muestras de control se eligieron basándose en el tipado de HLA, etnia, género y edad disponibles. La mayoría de los controles (todos menos 277) se eligieron de modo que la frecuencia de los haplotipos de HLA DR2 y DR3 coincidieran con los encontrados en SLE.

Ha habido tres versiones de Illumina HumanHap550. El número de SNP compartidos entre la versión 1 y la versión 3 es 545 080; únicamente estos SNP se analizaron. La versión 1 se usó para todas las muestras de la cohorte 1 y la

cohorte 2 y las 1001 muestras de control. La versión 3 se usó para todas las muestras de la cohorte 3 y 410 muestras de control.

Se repitieron los chips con tasas de éxito promedio < 80 %. Después de completar todas las repeticiones, se eliminaron las muestras con tasas de éxito < 90 %.

Las muestras se dividieron inicialmente en dos grupos para el análisis. El primer grupo (grupo 1) consistía en todas las muestras de la cohorte 1 y la cohorte 2 (466 casos) y 724 muestras de control. El segundo grupo (grupo 2) consistía en todas las muestras de la cohorte 3 (613 casos) y las 687 muestras de control restantes.

Filtrado en el grupo 1

Las muestras se comprobaron para la concordancia entre el género determinado por el genotipo y los registros clínicos; se encontró una discrepancia en 10 muestras (3 casos, 7 controles), que se eliminaron del análisis adicional.

Las muestras entonces se ensayaron para mezcla intercontinental usando el programa STRUCTURE (se puede acceder al enlace en línea escribiendo "pritch.bsd.uchicago.edu/structure" con ".html" como sufijo) (esencialmente como se describe en Pritchard et al., Genetics (2000), 155:945-959; Falush et al., Genetics (2003), 164:1567-1587; Falush et al., Molecular Ecology Notes (2007), doi:10.1111/j.1471-8286.2007.01758.x). El HumanHap550 incluye un "panel de ensayo de ADN" de 276 SNP que son ideales para determinar el porcentaje de ascendencia para las poblaciones CEU, YRI y CHB+JPT del proyecto HapMap. (CHB y JPT no podían discriminarse usando estos SNP.) Unos 274 de los 276 SNP en el panel de ensayo de ADN se genotiparon en todas las poblaciones HapMap; se ejecutó STRUCTURE con genotipos de estos 274 SNP en el conjunto que consistía en las muestras restantes del grupo 1 (463 casos, 717 controles) más una muestra de cada árbol genealógico en el proyecto HapMap (es decir, 20 muestras CEPH de Utah (CEU), 30 muestras Yoruba (YRI), 45 muestras chinas Han (CHB) y 44 muestras japonesas (JPT)). Las muestras HapMap se incluyeron como controles positivos y para ayudar al algoritmo de agrupación. Se ejecutó STRUCTURE independientemente tres veces con los mismos parámetros: usando el modelo de ascendencia mezclada y el modelo de frecuencia alélica correlacionado sin información previa de la población, suponiendo tres poblaciones, con 30 000 etapas de inicialización seguidas por 100 000 etapas de MonteCarlo de cadenas de Markov. Las tres ejecuciones obtuvieron coeficientes muy similares de ascendencia para cada muestra y cada muestra HapMap obtuvo > 93,0 % de ascendencia con su origen geográfico; cada muestra CEU obtuvo > 97,0 % de ascendencia CEU. Las muestras que obtuvieron < 90,0 % de ascendencia CEU en cualquier de las tres ejecuciones (28 casos, 24 controles) se eliminaron del análisis adicional.

Para las muestras restantes (435 casos, 693 controles), los SNP con tasas de éxito < 95 % (23 275 SNP (4 %)) se eliminaron del análisis adicional. Después, los SNP con probabilidad de Hardy-Weinberg $\leq 0,001$ en los controles (15 622 SNP (3 %)) se eliminaron del análisis adicional.

Filtrado en el grupo 2

Las muestras no se comprobaron explícitamente para la concordancia entre el género determinado por el genotipo y los registros clínicos.

Los SNP con tasas de éxito < 95 % (34 998 SNP (6 %)) se eliminaron del análisis adicional.

Las muestras entonces se ensayaron para mezcla intercontinental usando STRUCTURE, como se describe anteriormente. Las muestras que obtuvieron < 90,0 % de ascendencia CEU en cualquiera de las tres ejecuciones (21 casos, 24 controles) se eliminaron del análisis adicional.

Para las muestras restantes (592 casos, 663 controles), los SNP con probabilidad de Hardy-Weinberg $\leq 0,001$ en los controles (22 202 SNP (4 %)) se eliminaron del análisis adicional.

Combinación de los grupos 1 y 2

Los grupos 1 y 2 se combinaron para el análisis final.

Las muestras restantes (435 casos, 693 controles) en el grupo 1 se combinaron con las muestras restantes (592 casos, 663 controles) en el grupo 2 para producir el grupo final (1027 casos, 1356 controles). Únicamente los SNP restantes tanto en el grupo 1 como en el grupo 2 (496 458 SNP) se analizaron adicionalmente.

Todas las muestras sin discrepancias de género (1076 casos, 1404 controles) se comprobaron para observar si podían ser duplicadas o relacionadas. Inicialmente, todos los pares de muestras se compararon entre 800 SNP propagados en todo el genoma. Los candidatos duplicados y relacionados entonces se comprobaron entre 540 000+ SNP. Se detectaron tres grupos de valores atípicos. El primer grupo (20 pares) obtuvo >95 % de identidad entre cada par y se consideró duplicado. El segundo grupo (17 pares) obtuvo 67-77 % de identidad entre cada par y se

consideró relacionado. El tercer grupo (5 pares) obtuvo 58-63 % de identidad entre cada par y se consideró relacionado. (La identidad promedio entre las muestras fue de 51-55 %). Globalmente, se eliminaron 39 muestras (29 casos, 10 controles) del grupo final.

5 Los SNP en el ADN mitocondrial (19 SNP) se eliminaron del análisis adicional.

El grupo principal resultante (998 casos, 1346 controles, 496 439 SNP) se usó en el análisis siguiente.

El mismo análisis también se realizó en subconjuntos específicos del grupo principal: subconjunto 1:

10 mujeres únicamente (907 casos, 967 controles) y subconjunto 2: casos con nefritis lúpica, y todos los controles (286 casos, 1346 controles)

Análisis y resultados

15 Todos los SNP en el grupo principal se analizaron usando EIGENSTRAT, que es un programa que se describe esencialmente en Price et al., Nature Genetics (2006), 38:904 - 909 (puede accederse al enlace en línea escribiendo "genepath.med.harvard.edu/~reich/EIGENSTRAT" con ".html" como sufijo), que también corrige la estratificación de poblaciones. Los 10 componentes principales superiores se usaron para eliminar los valores atípicos durante 5 rondas y después corregir la estratificación. La estadística de ji cuadrado de EIGENSTRAT se calculó entonces y se calculó la probabilidad unilateral de la distribución de ji cuadrado con la función CHIDIST de Microsoft Excel con un grado de libertad.

25 Para determinar las regiones candidatas superiores, primero se redujo la cantidad de SNP candidatos usando un umbral de valor P: para el subconjunto 1 (mujeres) y el subconjunto 2 (nefritis), los SNP con un $P > 2,0 \times 10^{-5}$ se eliminaron del análisis adicional y para el grupo principal (998 casos, 1346 controles), los SNP con un $P > 7,0 \times 10^{-5}$ se eliminaron del análisis adicional. En el subconjunto de mujeres, permanecieron 19 SNP. En el subconjunto de nefritis, permanecieron 35 SNP. En el grupo principal, permanecieron 47 SNP. Entonces, se determinó la región de desequilibrio del ligamiento (LD) que contenía cada SNP examinando los diagramas de LD utilizando el programa HelixTree (puede accederse al enlace en línea escribiendo "www.goldenhelix.com/pharmhelixfeatures" con ".html" como sufijo) (Golden Helix, Montana, EE. UU.). El algoritmo EM se usó para calcular D' y r^2 usando únicamente los genotipos de estos casos y controles. Las regiones se identificaron a ojo, usando $D' \geq 0,9$ como límite.

35 Una vez que se identificó cada región, los genes en cada región se examinaron con un explorador de genoma establecido en la técnica (por ejemplo, el UCSC Genome Browser, esencialmente descrito en Kuhn et al., Nucleic Acids Res. (2007), 35(edición de la base de datos):D668-73; puede accederse al enlace en línea escribiendo "genome.ucsc" con "edu" como sufijo, conjunto de marzo de 2006). Se examinó la expresión génica inmunoespecífica, determinada en el estudio IRIS (Abbas et al., Genes and Immunity (2005), 6:319-331, incluyendo su material suplementario en línea). Se identificaron las regiones candidatas superiores, por ejemplo, por la presencia de genes inmunoespecíficos en una región. En el subconjunto de nefritis, se eligieron 11 regiones que contenían 20 SNP candidatos como probables de contener al menos un alelo de riesgo para SLE (figura 12). En el subconjunto de mujeres, se eligieron 6 regiones adicionales que contenían 9 SNP candidatos (figura 13). En el grupo principal, se eligieron 6 regiones adicionales que contienen 8 SNP candidatos (figura 14). Se eligió un total 23 regiones que contenían 37 SNP candidatos. Debe apreciarse que los SNP se enumeraron según el grupo de estudio que tenía el resultado más fuerte para los SNP, por tanto, no se muestran aciertos duplicados entre los grupos de estudio. además, los aciertos en la región de MHC no se incluyeron. Se determinaron las regiones LD identificadas basándose en los datos de las figuras 12-14 y se resumen en la figura 15-17, respectivamente.

50 Los siguientes párrafos numerados (párr.) contienen enunciados adicionales de diversos aspectos de la presente divulgación -

1. Un método de evaluación de un sujeto está en riesgo de desarrollar lupus, comprendiendo el método detectar, en una muestra biológica obtenida de dicho sujeto, la presencia de una característica genética indicativa de riesgo de desarrollar lupus, en el que dicha característica genética comprende un conjunto de uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.
2. El método del párr. 1, en el que dicho conjunto de SNP comprende aproximadamente 1-10, 10-20, 20-30, 30-40 o 40-50 SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.
3. El método del párr. 1, en el que dicho conjunto de SNP comprende 2 o más SNP, 3 o más SNP, 4 o más SNP, 5 o más SNP, 6 o más SNP, 7 o más SNP, 8 o más SNP, 9 o más SNP, 10 o más SNP, 11 o más SNP, 12 o más SNP, 13 o más SNP, 14 o más SNP, 15 o más SNP, 16 o más SNP, 17 o más SNP, 18 o más SNP, 19 o más SNP o 20 o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.
4. El método del párr. 1, en el que dicho conjunto de SNP comprende 1-19 SNP seleccionados de la tabla 6.
5. El método del párr. 1, en el que dicho conjunto de SNP comprende un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10.
6. El método del párr. 1, en el que dicho conjunto de SNP comprende un SNP de ITGAM seleccionado de

cualquiera de los SNP de ITGAM expuestos en las tablas 7-10.

7. El método del párr. 6, en el que dicho conjunto de SNP comprende además un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10.

8. El método del párr. 1, en el que dicho conjunto de SNP comprende uno o más SNP seleccionados del siguiente grupo de SNP: rs2187668, rs10488631, rs7574865, rs9888739, rs13277113, rs2431697, rs6568431, rs10489265, rs2476601, rs2269368, rs1801274, rs4963128, rs5754217, rs6445975, rs3129860, rs10516487, rs6889239, rs2391592 y rs2177770.

9. Un método de diagnóstico de lupus en un sujeto, comprendiendo el método detectar, en una muestra biológica obtenida de dicho sujeto, la presencia de una característica genética indicativa de lupus, en el que dicha característica genética comprende un conjunto de uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

10. Un polinucleótido aislado que comprende (a) un polinucleótido asociado a PRO o fragmento del mismo que es de al menos aproximadamente 10 nucleótidos de longitud, en el que el polinucleótido asociado a PRO o fragmento del mismo comprende una variación genética en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10, o (b) el complemento de (a).

11. El polinucleótido aislado del párr. 10 en el que la variación genética está en ADN genómico que codifica un gen (o su región reguladora) que comprende un polimorfismo de un único nucleótido (SNP) enumerado de las figuras 1-17 y las tablas 1-10.

12. El polinucleótido aislado del párr. 11, en el que el SNP está en una región no codificante del gen.

13. El polinucleótido aislado del párr. 11, en el que el SNP está en una región codificante del gen.

14. El polinucleótido aislado del párr. 10, en el que el polinucleótido aislado es un cebador.

15. El polinucleótido aislado del párr. 10, en el que el polinucleótido aislado es un oligonucleótido.

16. Un oligonucleótido que es (a) un oligonucleótido específico de alelo que hibrida con una región de un polinucleótido asociado a PRO que comprende una variación genética en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10, o (b) el complemento de (a).

17. El oligonucleótido del párr. 16, en el que el SNP está en ADN genómico que codifica un gen (o su región reguladora) que comprende un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10.

18. El oligonucleótido del párr. 17 en el que el SNP está en una región no codificante del gen.

19. El oligonucleótido del párr. 17, en el que el SNP está en una región codificante del gen.

20. El oligonucleótido del párr. 16, en el que el oligonucleótido específico de alelo es un cebador específico de alelo.

21. Un kit que comprende el oligonucleótido del párr. 16 y, opcionalmente, al menos una enzima.

22. El kit del párr. 21, en el que la al menos una enzima es una polimerasa.

23. El kit del párr. 21, en el que la al menos una enzima es una ligasa.

24. Una micromatriz que comprende el oligonucleótido del párr. 16.

25. Un método de detección de la ausencia o presencia de una variación en un polinucleótido asociado a PRO en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en las figuras 1-17 y las tablas 1-10, comprendiendo el método (a) poner en contacto un ácido nucleico sospechoso de comprender la variación con un oligonucleótido específico de alelo que es específico para la variación en condiciones adecuadas para la hibridación del oligonucleótido específico de alelo con el ácido nucleico; y (b) detectar la presencia o ausencia de hibridación específica de alelo.

26. El método del párr. 25, en el que la variación comprende un SNP expuesto en las figuras 1-17 y las tablas 1-10.

27. Un método de amplificación de un ácido nucleico que comprende una variación en un polinucleótido asociado a PRO en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en las figuras 1-17 y las tablas 1-10, comprendiendo el método (a) poner en contacto el ácido nucleico con un cebador que hibrida con el ácido nucleico en una secuencia 3' de la variación y (b) prolongar el cebador para generar un producto de amplificación que comprende la variación.

28. El método del párr. 27, en el que la variación comprende un SNP expuesto en las figuras 1-17 y las tablas 1-10.

29. Un método de determinación del genotipo de una muestra biológica de un mamífero, comprendiendo el método detectar, en material de ácido nucleico derivado de la muestra biológica, la ausencia o presencia de una variación en un polinucleótido asociado a PRO en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en las figuras 1-17 y las tablas 1-10.

30. El método del párr. 29, en el que la variación comprende un SNP expuesto en las figuras 1-17 y las tablas 1-10.

31. El método del párr. 29, en el que se sabe que la muestra biológica comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación.

32. El método del párr. 29, en el que la muestra biológica es un tejido enfermo.

33. El método del párr. 29, en el que la detección comprende realizar un proceso seleccionado de un ensayo de prolongación de cebador; un ensayo de prolongación de cebador específico de alelo; un ensayo de incorporación de nucleótidos específica de alelo; un ensayo de hibridación de oligonucleótido específico de alelo; un ensayo de 5' nucleasa, un ensayo que emplea balizas moleculares; y un ensayo de ligamiento de oligonucleótidos.

34. Un método de subclasificación del lupus en un mamífero, comprendiendo el método detectar la presencia de una variación en un polinucleótido asociado a PRO en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en las figuras 1-17 y las tablas 1-10 en una muestra biológica obtenida del mamífero, en el que se sabe que la muestra biológica comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación.

35. El método del párr. 34, en el que la variación es una variación genética.

36. El método del párr. 35, en el que la variación comprende un SNP expuesto en las figuras 1-17 y las tablas 1-10.

37. El método del párr. 34, en el que la detección comprende realizar un proceso seleccionado de un ensayo de prolongación de cebador; un ensayo de prolongación de cebador específico de alelo; un ensayo de incorporación de nucleótidos específica de alelo; un ensayo de hibridación de oligonucleótido específico de alelo, un ensayo de 5' nucleasa; un ensayo que emplea balizas moleculares; y un ensayo de ligamiento de oligonucleótidos.

38. Un método para predecir si un sujeto con lupus responderá a un agente terapéutico contra el lupus, comprendiendo el método determinar si el sujeto comprende una variación en un polinucleótido asociado a PRO en una posición nucleotídica correspondiente a la posición de un polimorfismo de un único nucleótido (SNP) expuesto en las figuras 1-17 y las tablas 1-10, en el que la presencia de una variación indica que el sujeto responderá al agente terapéutico.

39. El método del párr. 38, en el que la variación es una variación genética.

40. El método del párr. 39, en el que la variación comprende un SNP expuesto en las figuras 1-17 y las tablas 1-10.

41. Un método de diagnóstico o pronóstico del lupus en un sujeto, comprendiendo el método detectar la presencia de una variación en un PRO o polinucleótido asociado a PRO derivado de una muestra biológica obtenida del sujeto, en el que:

(a) se sabe que la muestra biológica comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación;

(b) la variación comprende o está localizada en una posición nucleotídica correspondiente a un SNP expuesto en las figuras 1-17 y las tablas 1-10; y

(c) la presencia de la variación es un diagnóstico o pronóstico de lupus en el sujeto.

42. Un método de ayuda al diagnóstico o pronóstico de lupus en un sujeto, comprendiendo el método detectar la presencia de una variación en un PRO o polinucleótido asociado a PRO derivado de una muestra biológica obtenida del sujeto, en el que:

(a) se sabe que la muestra biológica comprende o es sospechosa de comprender un PRO o polinucleótido asociado a PRO que comprende la variación;

(b) la variación comprende o está localizada en una posición nucleotídica correspondiente a un SNP expuesto en las figuras 1-17 y las tablas 1-10; y

(c) la presencia de la variación es un diagnóstico o pronóstico de una afección o síntoma de lupus en el sujeto.

43. El método del párr. 41 o 42, en el que el polinucleótido asociado a PRO codifica un PRO que está codificado por una secuencia dentro de una región de desequilibrio del ligamiento.

44. El método del párr. 43, en el que región de desequilibrio del ligamiento es una de las expuestas en las figuras 1-17 y las tablas 1-10.

45. El método del párr. 41 o 42, en el que la variación está en un ADN genómico que codifica un gen o su región reguladora, y en el que el gen respectivo o su región reguladora comprende un SNP expuesto en las figuras 1-17 y las tablas 1-10.

46. El método del párr. 45, en el que el SNP está en una región no codificante del gen.

47. El método del párr. 45, en el que el SNP está en una región codificante del gen.

48. Un método de identificación de un agente terapéutico eficaz para tratar el lupus en una subpoblación de pacientes, comprendiendo el método correlacionar la eficacia del agente con la presencia de una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) en la subpoblación de pacientes, en el que el SNP es uno de los enumerados en las figuras 1-17 y las tablas 1-10, identificando de ese modo el agente como eficaz para tratar el lupus en dicha subpoblación de pacientes.

49. Un método de tratamiento de una afección de lupus en un sujeto en el que se sabe que hay una variación genética presente en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerada en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para tratar la afección.

50. Un método de tratamiento de un sujeto que tiene una afección de lupus, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para tratar la afección en un sujeto que tiene una variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerada en las figuras 1-17 y las tablas 1-10.

51. Un método de tratamiento de un sujeto que tiene una afección de lupus, comprendiendo el método administrar al sujeto un agente terapéutico que ha demostrado ser eficaz para tratar dicha afección en al menos un estudio clínico, en el que el agente se administró a al menos cinco sujetos humanos que tenían cada uno una

variación genética en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10.

52. El método del párr. 51, en el que los al menos cinco sujetos tenían dos o más SNP diferentes en total para el grupo de al menos cinco sujetos.

53. El método del párr. 51, en el que los al menos cinco sujetos tenían el mismo SNP para el grupo completo de al menos cinco sujetos.

54. Un método de tratamiento de un sujeto con lupus de una subpoblación específica de pacientes con lupus, en el que la subpoblación se caracteriza al menos en parte por la asociación con la variación genética en una posición nucleotídica correspondiente a un SNP enumerado en las figuras 1-17 y las tablas 1-10, y en el que el método comprende administrar al sujeto una cantidad eficaz de un agente terapéutico que está aprobado como agente terapéutico para dicha subpoblación.

55. El método del párr. 54, en el que la subpoblación tiene nefritis lúpica.

56. El método del párr. 54, en el que la subpoblación es femenina.

57. El método del párr. 54, en el que la subpoblación es de ascendencia europea.

58. Un método que comprende fabricar un agente terapéutico contra el lupus y envasar el agente con instrucciones para administrar el agente a un sujeto que tiene o se cree que tiene lupus y que tiene una variación genética en una posición correspondiente a un polimorfismo de único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10.

59. Un método de especificación de un agente terapéutico para su uso en una subpoblación de pacientes con lupus, comprendiendo el método proporcionar instrucciones para administrar el agente terapéutico a una subpoblación de pacientes caracterizada por una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10.

60. Un método para comercializar un agente terapéutico para su uso en una subpoblación de pacientes con lupus, comprendiendo el método informar a una audiencia determinada acerca del uso del agente terapéutico para tratar a la subpoblación de pacientes caracterizada por la presencia, en pacientes de dicha subpoblación, de una variación genética en una posición correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10.

61. Un método para modular la señalización a través del receptor de linfocitos B en un sujeto en que se sabe que hay una variación genética presente en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para modular la señalización a través el receptor de linfocitos B.

62. Un método para modular la diferenciación de células Th17 en un sujeto en que se sabe que hay una variación genética presente en una posición nucleotídica correspondiente a un polimorfismo de un único nucleótido (SNP) enumerado en las figuras 1-17 y las tablas 1-10, comprendiendo el método administrar al sujeto un agente terapéutico eficaz para modular la diferenciación de células Th17.

63. Un conjunto de SNP que comprende una característica genética indicativa del riesgo de desarrollar lupus, en el que dicho conjunto de SNP comprende uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

64. El conjunto de SNP del párr. 63, en el que dicho conjunto de SNP comprende aproximadamente 1-10, 10-20, 20-30, 30-40 o 40-50 SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

65. El conjunto de SNP del párr. 63, en el que dicho conjunto de SNP comprende uno o más SNP seleccionados del grupo que consiste en rs9888739, rs13277113, rs7574865, rs2269368, rs6889239, rs2391592 y rs2177770.

66. El conjunto de SNP del párr. 63, en el que dicho conjunto de SNP comprende 2 o más SNP, 3 o más SNP, 4 o más SNP, 5 o más SNP, 6 o más SNP, 7 o más SNP, 8 o más SNP, 9 o más SNP, 10 o más SNP, 11 o más SNP, 12 o más SNP, 13 o más SNP, 14 o más SNP, 15 o más SNP, 16 o más SNP, 17 o más SNP, 18 o más SNP, 19 o más SNP o 20 o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

67. El conjunto de SNP del párr. 63, en el que dicho conjunto de SNP comprende 1-19 SNP seleccionados de la tabla 6.

68. El conjunto de SNP del párr. 63, en el que dicho conjunto de SNP comprende un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10.

69. El conjunto de SNP del párr. 63, en el que dicho conjunto de SNP comprende un son de ITGAM seleccionado de cualquiera de los SNP de ITGAM expuestos en las tablas 7-10.

70. El conjunto de SNP del párr. 69, en el que dicho conjunto de SNP comprende además un SNP de BLK seleccionado de cualquiera de los SNP de BLK expuestos en las tablas 7-10.

71. El conjunto de SNP del párr. 63, en el que dicho conjunto de SNP comprende uno o más SNP seleccionados del siguiente grupo de SNP: rs2187668, rs10488631, rs7574865, rs9888739, rs13277113, rs2431697, rs6568431, rs10489265, rs2476601, rs2269368, rs1801274, rs4963128, rs5754217, rs6445975, rs3129860, rs10516487, rs6889239, rs2391592 y rs2177770.

72. Un conjunto de SNP que comprende una característica genética indicativa de lupus, en el que dicho conjunto de SNP comprende uno o más SNP seleccionados de cualquiera de los SNP expuestos en las figuras 1-17 y las tablas 1-10.

REIVINDICACIONES

1. Un método de diagnóstico del lupus o de evaluación del riesgo de desarrollar lupus en un sujeto de una subpoblación específica de paciente con lupus, comprendiendo el método:
 - a) detectar en una muestra biología obtenida del sujeto la presencia de una característica genética indicativa de lupus o riesgo de desarrollar lupus, en donde dicha característica genética comprende la ausencia de un alelo minoritario del polimorfismo de un único nucleótido (SNP) rs4963128 y la presencia de un alelo minoritario del SNP rs2269368; y
 - b) determinar que el sujeto tiene lupus o está en riesgo de desarrollar lupus cuando la característica genética está presente.
2. El método de la reivindicación 1, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs5754217, rs7574865, rs3129860, rs10489265 y rs6445975; y/o la ausencia de un alelo minoritario de rs2431697.
3. El método de la reivindicación 1, en el que la característica genética comprende además de presencia de un alelo minoritario de uno o más de rs5754217 y rs7574865.
4. El método de la reivindicación 3, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs10488631 y rs1143679; y/o la ausencia de un alelo minoritario de rs1801274.
5. El método de la reivindicación 1, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs3129860 y rs10489265.
6. El método de la reivindicación 5, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs6568431, rs2187668, rs13277113 y rs2476601; y/o la ausencia de un alelo minoritario de rs10516487.
7. El método de la reivindicación 1, en el que la característica genética comprende además la presencia de un alelo minoritario de rs6445975; y/o la ausencia de un alelo minoritario de rs2431697.
8. Un agente terapéutico eficaz para tratar el lupus, para su uso en un método de tratamiento del lupus en un sujeto de una subpoblación específica de pacientes con lupus, en donde la presencia de la característica genética en el sujeto se determina de acuerdo con el método de la reivindicación 1, identificando de ese modo para el tratamiento al sujeto como parte de una subpoblación específica de pacientes con lupus.
9. El agente terapéutico para el uso de la reivindicación 8, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs5754217, rs7574865, rs3129860, rs10489265 y rs6445975; y/o la ausencia de un alelo minoritario de rs2431697.
10. El agente terapéutico para el uso de la reivindicación 8, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs5754217 y rs7574865.
11. El agente terapéutico para el uso de la reivindicación 10, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs10488631 y rs1143679; y/o la ausencia de un alelo minoritario de rs1801274.
12. El agente terapéutico para el uso de la reivindicación 8, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs3129860 y rs10489265.
13. El agente terapéutico para el uso de la reivindicación 12, en el que la característica genética comprende además la presencia de un alelo minoritario de uno o más de rs6568431, rs2187668, rs13277113 y rs2476601; y/o la ausencia de un alelo minoritario de rs10516487.
14. El agente terapéutico para el uso de reivindicación 8, en el que la característica genética comprende además la presencia de un alelo minoritario de rs6445975; y/o la ausencia de un alelo minoritario de rs2431697.
15. Una composición que comprende una pluralidad de moléculas de ácido nucleico cada una con capacidad de hibridación específica de alelo a un conjunto de regiones diana de un polinucleótido, comprendiendo dichas regiones alelos minoritarios de los SNP definidos en una cualquiera de las reivindicaciones 1 a 7.

Figura 1

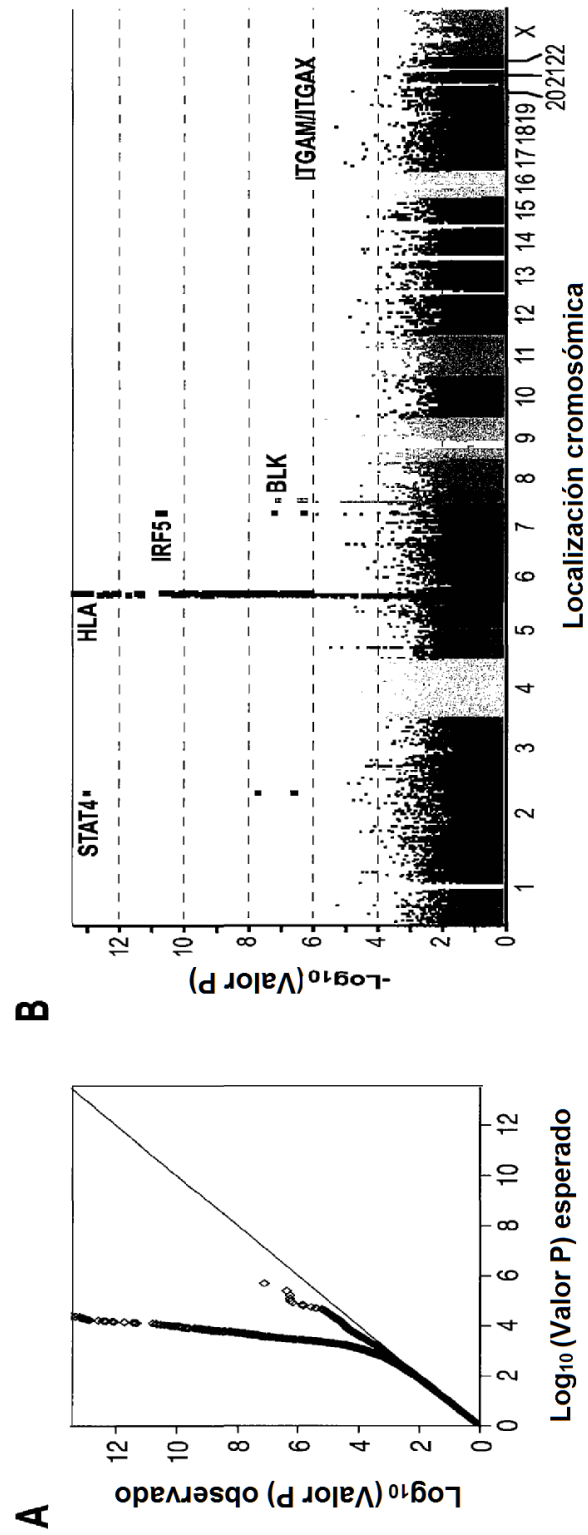


Figura 2

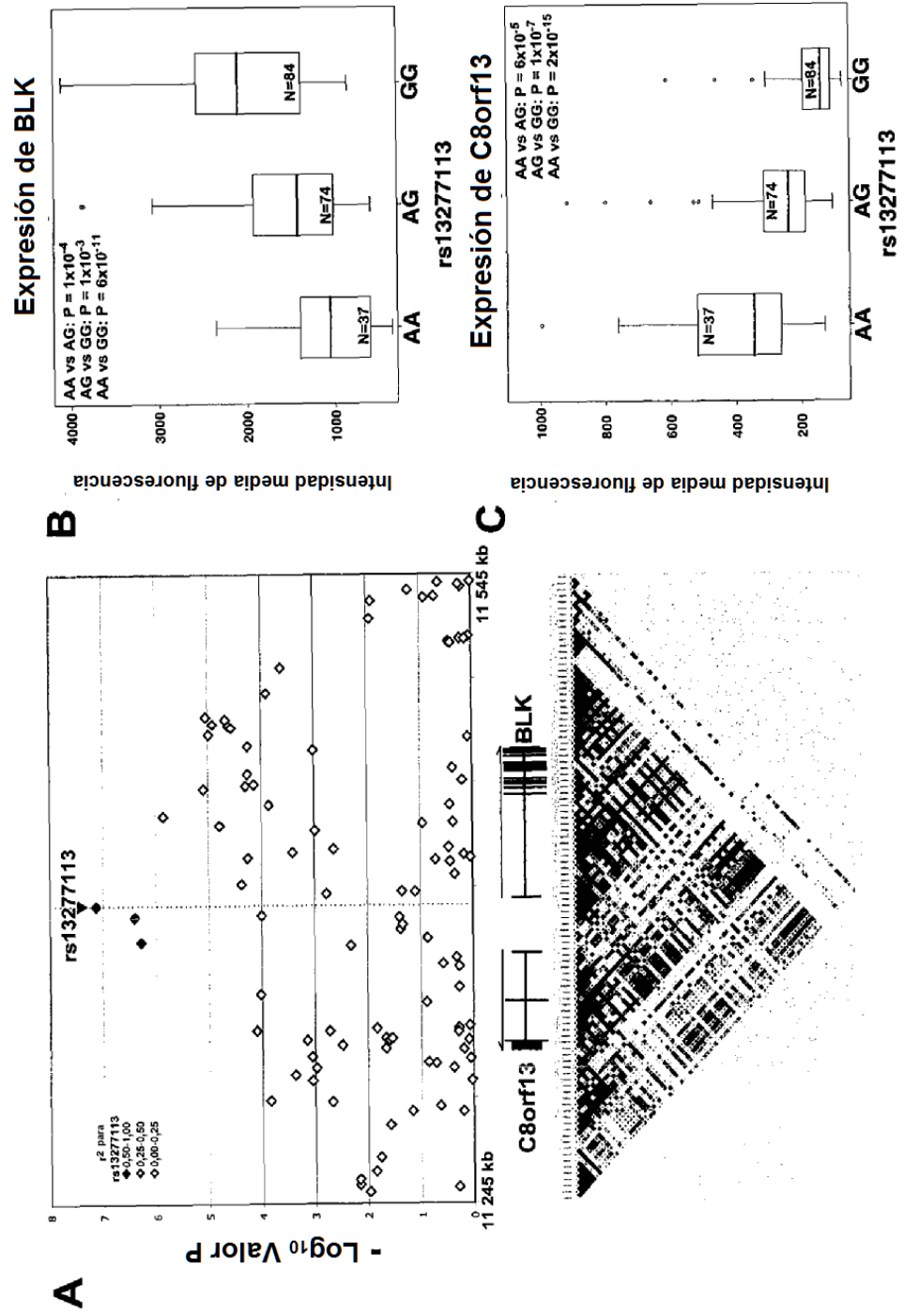


Figura 3

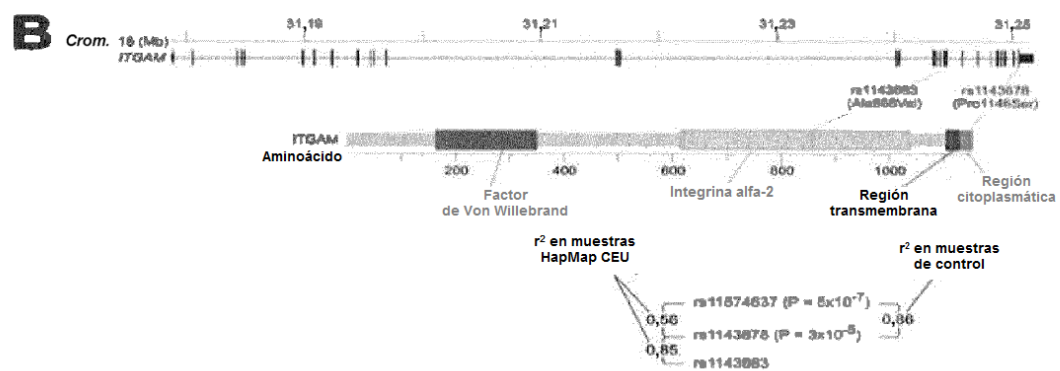
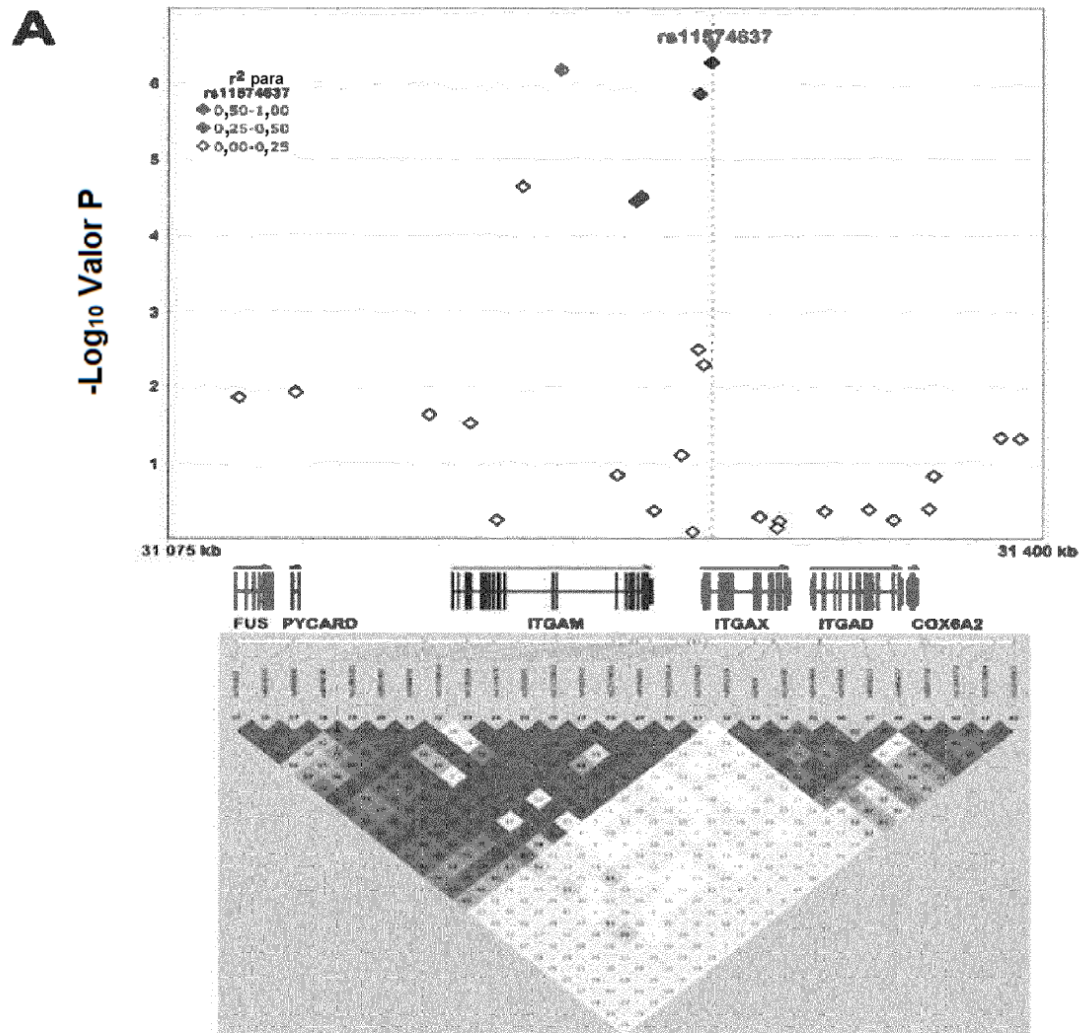


Figura 4

Frecuencia de características clínicas en las series de SLE 1-3 y los casos suecos

Fenotipos*	Todos los sujetos GWAS	Serie 1	Serie 2	Serie 3	Casos suecos
AutoAb antinucleares	1 193/1 301 (91,7%)	385/409 (94,1%)	541/595 (90,9%)	292/297 (98,3%)	758/781 (97,1%)
Artritis	1 023/1 303 (78,5%)	310/411 (75,4%)	444/595 (74,6%)	269/297 (90,6%)	594/781 (76,1%)
Fotosensibilidad	928/1 301 (71,3%)	288/411 (70,1%)	469/595 (78,8%)	171/295 (58,0%)	528/781 (67,61%)
Trastorno inmunológico	867/1 303 (66,5%)	313/411 (76,2%)	336/595 (56,5%)	218/297 (73,4%)	511/781 (65,4%)
Trastorno hemático	795/1 301 (61,1%)	274/411 (66,7%)	368/595 (61,9%)	153/295 (51,9%)	450/781 (57,6%)
Eritema vespertino	635/1 270 (50,0%)	235/411 (57,2%)	274/595 (46,1%)	126/264 (47,7%)	452/781 (57,9%)
Úlceras bucales	564/1 302 (43,3%)	222/411 (54,0%)	183/595 (30,8%)	159/296 (53,7%)	180/781 (23,0%)
Serositis	499/1 299 (38,4%)	195/410 (47,6%)	175/595 (29,4%)	129/294 (43,9%)	345/781 (44,2%)
Trastorno renal	365/1 302 (28,0%)	137/411 (33,3%)	139/595 (23,4%)	89/296 (30,1%)	239/781 (30,6%)
Trastorno neurológico	125/1 301 (9,6%)	38/411 (9,3%)	59/595 (9,9%)	28/295 (9,5%)	88/781 (11,3%)
Eritema discoide	120/1 270 (9,5%)	67/411 (16,3%)	39/595 (6,6%)	14/264 (5,3%)	196/781 (25,1%)

*Véase <http://www.rheumatology.org/publications/classification/SLE/sle.asp> para las definiciones de fenotipo

Figura 5

50 loci superiores con SLE en una exploración de genoma completa en 1311 casos y 3340 controles																				
CHR	Banda	SNP *	Posición (coord. hg19)	Locus	Serie 1: 411 casos 1047 controles			Serie 2: 595 casos 1515 controles			Serie 3: 305 casos 777 controles			Combinado: 1311 casos 3340 controles						
					Frecuencia de alelo minoritario			Frecuencia de alelo minoritario			Frecuencia de alelo minoritario			Frecuencia de alelo minoritario						
					Alelo	Caso	Control	Alelo	Caso	Control	Alelo	Caso	Control	Alelo	Caso	Control				
6	6p21.32	rs2187668	32 713 863	HIA-DOA1	A	0.18	0.11	27.1	A	0.19	0.12	29.0	A	0.21	0.11	34.6	A	0.19	0.11	27.1E-21
2	2q32.2	rs7574865	191 760 140	STAT4	T	0.31	0.22	27.2	T	0.30	0.24	14.3	T	0.33	0.24	16.1	T	0.31	0.23	6.96E-14
7	7q32.1	rs10486631	128 188 135	IRF5/IRF3	C	0.17	0.12	12.1	C	0.17	0.11	23.9	C	0.17	0.11	13.0	C	0.17	0.11	1.65E-11
6	6p23.1	rs13277113	11 390 599	BLK/Cbrf13	A	0.29	0.24	5.3	A	0.28	0.22	12.8	A	0.31	0.22	12.9	A	0.29	0.23	7.57E-08
16	16p11.2	rs11574637	31 276 378	ITGAM/ITGA2	C	0.20	0.18	1.7	C	0.24	0.19	13.4	C	0.25	0.20	10.9	C	0.23	0.19	9.35E-07
9	9p34.13	rs11243676	132 126 322	NTNG2	A	0.05	0.07	6.7	A	0.05	0.05	2.7	A	0.04	0.07	13.3	A	0.05	0.07	2.65E-06
5	5p12	rs979233	42 411 001	GHR	T	0.47	0.46	12.2	T	0.47	0.46	3.5	T	0.49	0.44	16.5	T	0.48	0.45	3.85E-06
16	16p22.3	rs17063644	66 993 224	SOC5	A	0.04	0.03	4.0	A	0.04	0.03	2.8	A	0.05	0.03	14.5	A	0.05	0.03	6.05E-06
1	1p31.1	rs12141351	71 986 495	NEGR1	A	0.04	0.02	5.7	A	0.04	0.02	7.4	A	0.04	0.02	7.5	A	0.04	0.02	6.99E-06
17	17p11.2	rs1019412	29 512 627	ACCN1	T	0.26	0.23	2.5	T	0.26	0.22	3.4	T	0.31	0.23	15.1	T	0.27	0.23	1.12E-05
7	7p14.3	rs129604	31 712 622	PDE1C	T	0.51	0.47	3.7	T	0.49	0.47	4.7	T	0.55	0.47	11.3	T	0.51	0.47	1.27E-05
2	2p36.3	rs4675077	227 063 536	IRS1	C	0.17	0.14	0.6	C	0.18	0.15	6.0	C	0.19	0.13	14.9	C	0.18	0.14	1.37E-05
4	4p21.23	rs1921880	85 314 378	MGC11324	A	0.08	0.14	9.0	A	0.11	0.13	1.3	A	0.09	0.13	9.7	A	0.09	0.13	1.49E-05
12	12p14.2	rs7656933	61 631 288	AVPR1A	G	0.01	0.02	5.2	G	0.01	0.02	7.1	G	0.01	0.02	6.8	G	0.01	0.02	1.59E-05
20	20q13.12	rs6094681	45 387 700	PRKCBP1	T	0.13	0.10	6.9	T	0.10	0.08	3.9	T	0.12	0.08	7.5	T	0.12	0.09	1.83E-05
17	17p11.2	rs4404112	19 769 646	TNFRSF13B	G	0.18	0.17	1.9	G	0.20	0.18	3.9	G	0.23	0.15	14.6	G	0.20	0.17	1.69E-05
2	2p31.1	rs1389354	172 837 045	DLX2	A	0.12	0.17	15.3	A	0.15	0.16	2.1	A	0.12	0.18	3.8	A	0.13	0.18	1.91E-05
5	5p33.3	rs254850	156 598 310	UBI CP1	T	0.21	0.17	6.8	T	0.24	0.18	15.9	T	0.21	0.19	9.8	T	0.22	0.18	2.97E-05
10	10p14	rs11255111	7 479 732	SEMBT2	A	0.02	0.04	3.0	A	0.02	0.03	6.8	A	0.01	0.04	0.3	A	0.02	0.03	2.97E-05
5	5p13.1	rs682054	42 163 692	FBXO4	C	0.29	0.28	2.2	C	0.30	0.30	6.2	C	0.34	0.28	10.4	C	0.31	0.29	2.80E-05
12	12p24.23	rs1961707	117 843 445	KIAA1653	T	0.29	0.34	6.3	T	0.29	0.34	13.1	T	0.31	0.34	2.3	T	0.32	0.34	2.90E-05
8	8p23.1	rs656319	9 861 822	MSRA	T	0.45	0.44	0.7	T	0.47	0.43	7.9	T	0.49	0.40	12.5	T	0.47	0.42	3.10E-05
8	8p12.3	rs11783343	62 855 576	ASPH	T	0.11	0.17	16.0	T	0.17	0.17	6.0	T	0.14	0.19	7.5	T	0.14	0.17	3.37E-05
1	1p21.1	rs485239	144 352 419	ACPB	A	0.19	0.16	5.1	A	0.17	0.16	2.9	A	0.18	0.15	8.2	A	0.18	0.15	3.46E-05
2	2p14	rs874952	65 526 177	SPRED2	A	0.13	0.16	7.2	A	0.12	0.14	5.4	A	0.12	0.15	4.4	A	0.12	0.14	3.48E-05
8	8p23.1	rs2955337	8 135 460	DEFB4	G	0.47	0.44	0.9	G	0.46	0.41	5.9	G	0.50	0.40	14.4	G	0.47	0.42	3.53E-05
8	8p12	rs4732950	29 852 131	DUSP4	T	0.17	0.14	6.9	T	0.17	0.15	3.7	T	0.18	0.13	6.6	T	0.17	0.14	3.56E-05
3	3p26.3	rs17039471	1 469 265	CNTN5	A	0.07	0.06	3.4	A	0.07	0.07	5.8	A	0.11	0.06	17.2	A	0.09	0.06	3.74E-05
14	14q21.1	rs1715960	73 135 265	RGS12	T	0.35	0.23	8.7	T	0.36	0.21	8.4	T	0.38	0.21	3.3	T	0.38	0.21	3.73E-05
18	18p11.31	rs132281933	84 144 139	SLC23A3	T	0.25	0.20	2.1	T	0.13	0.13	4.5	T	0.23	0.18	11.0	T	0.23	0.18	3.90E-05
17	17p13.1	rs109211	42 143 434	ANKK59B	G	0.23	0.26	2.9	G	0.20	0.25	6.3	G	0.24	0.25	8.5	G	0.23	0.19	4.03E-05
12	12p13.2	rs202503	12 162 900	BCL2L14	G	0.21	0.19	5.0	G	0.19	0.16	0.4	C	0.24	0.16	14.4	C	0.21	0.25	4.04E-05
14	14q12	rs1192118	25 195 688	STXBPA	A	0.41	0.47	8.6	A	0.41	0.45	5.4	A	0.40	0.47	3.8	A	0.41	0.46	4.10E-05
7	7p15.3	rs1029597	21 536 396	DNAAF1	A	0.16	0.19	3.1	A	0.15	0.17	3.4	A	0.13	0.19	11.0	A	0.15	0.18	4.42E-05
17	17p22	rs1504455	50 635 553	STXBPA	C	0.17	0.16	0.4	C	0.19	0.16	8.2	C	0.19	0.13	12.9	C	0.18	0.16	4.43E-05
8	8p23.1	rs16768938	10 054 330	MSRA	T	0.42	0.42	0.0	T	0.41	0.44	3.7	T	0.36	0.47	22.5	T	0.40	0.44	4.45E-05
12	12p23.3	rs16768938	107 168 267	KIAA0789	A	0.10	0.08	2.5	A	0.12	0.09	12.7	A	0.12	0.09	4.7	A	0.11	0.09	4.80E-05
4	4p22.3	rs17449854	122 985 334	ANKK5	C	0.06	0.05	1.1	C	0.06	0.04	4.1	C	0.09	0.04	13.9	C	0.07	0.05	4.89E-05
3	3p22.1	rs10510732	41 982 344	OIP106	A	0.13	0.17	3.3	A	0.13	0.16	5.1	A	0.12	0.17	7.7	A	0.13	0.17	5.04E-05
17	17p25.2	rs12451573	72 716 745	SEPT9	T	0.09	0.09	0.2	T	0.10	0.08	5.9	T	0.13	0.08	15.9	T	0.11	0.08	5.05E-05
8	8p23.1	rs1814795	10 546 696	RP11.1	C	0.47	0.44	2.4	C	0.47	0.43	3.9	C	0.51	0.42	11.0	C	0.48	0.43	5.24E-05
12	12p12	rs10506202	40 307 616	PDZRN4	G	0.29	0.28	1.8	G	0.31	0.28	7.5	G	0.34	0.27	8.6	G	0.31	0.28	5.59E-05
1	1p25.3	rs142737657	160 459 032	RGL1	T	0.06	0.04	1.9	T	0.06	0.04	3.9	G	0.05	0.04	11.8	G	0.06	0.04	5.63E-05
15	15q15.1	rs16023530	39 459 093	NUSAP1	C	0.29	0.25	4.1	C	0.28	0.27	6.0	C	0.29	0.24	6.5	C	0.29	0.26	5.89E-05
11	11p22.1	rs1021273	98 316 370	CNTN5	G	0.29	0.34	11.5	G	0.32	0.32	0.2	G	0.29	0.34	11.7	G	0.30	0.33	5.98E-05
6	6p22.31	rs8917541	121 303 805	C6orf184	G	0.09	0.11	2.1	G	0.08	0.10	4.6	G	0.07	0.13	10.3	G	0.08	0.11	6.05E-05
1	1p34.1	rs228081	44 515 595	C1orf184	G	0.37	0.31	5.4	G	0.36	0.32	13.2	G	0.34	0.31	1.9	G	0.36	0.31	6.29E-05
8	8p24.23	rs16906762	138 202 322	LOC51059	G	0.10	0.07	8.1	G	0.08	0.08	0.1	G	0.11	0.07	11.2	G	0.09	0.07	6.62E-05
2	2q24.2	rs12023360	162 979 571	IFIH1	G	0.45	0.50	2.6	G	0.44	0.49	5.1	G	0.44	0.52	9.0	G	0.45	0.50	6.85E-05

* Los SNP más asociados de los 50 loci superiores. Los SNP dentro de 100 kb de los SNP más asociados en una región se excluyeron únicamente se incluyen los SNP superiores de la región HLA. Para los resultados completos y detalles adicionales, véase la estadística resumida para todos los SNP (Figura 6).

El tamaño de cada caso compuesto de 6506 SNPs para cada serie.

* Los SNP más asociados de los 50 loci superiores. Los SNP dentro de 100 kb de los SNP más asociados en una región se excluyeron y únicamente se incluyen los SNP superiores de la región HLA. Para los resultados completos y detalles adicionales, véase la estadística resumida para todos los SNP (figura 8).
 A = Valor de J cuadrado corregido de EGENSTRAT para cada serie.
 * = El valor P combinado como se describe en los métodos.

Figura 6

Niveles de expresión de BLK, C8orf13 y genes de control en 210 líneas transformadas de linfocitos B de individuos HapMap											
	Genotipo [^]	N	BLK	C8orf13	ACTB	GAPDH	Intensidad media de fluorescencia *				
rs13277113	AA	37	1038,8	400,5	50972,2	27367,8	AAVAG	0,00011	6,1461E-05	0,691	0,746
rs13277113	AG	74	1574,0	275,7	51272,0	27656,2	AGvGG	0,00163	1,01146E-07	0,259	0,835
rs13277113	GG	87	2057,3	165,0	51823,1	27266,8	AAvGG	6,46678E-11	2,33428E-15	0,161	0,894
* Intensidad media de fluorescencia para cuatro mediciones para sondas para BLK (GI_33469981-S), C8orf13 (GI_32698772-S), ITGAM (GI_6006013-S), ITGAX (GI_34452172-S), ACTB (beta-actina, GI_5016088-S) y GAPDH (GI_7669491-S). Todos los datos de expresión son del proyecto GENEVAR (http://www.sanger.ac.uk/humgen/genevar/).											
^ Genotipos del proyecto internacional HapMap (www.hapmap.org)											

Figura 7

Expresión de BLK en linfocitos T transformados de las poblaciones HapMap															
			CEU				CHB+JPT				YRI		Todas las poblaciones		
	Genotipo ^	N	Expresión de BLK *		N	Expresión de BLK		N	Expresión de BLK		N	Expresión de BLK		Genotipo	P +
rs13277113	AA	2	10,17	33	9,84	2	10,56	37	9,89	AA vs AG	1,9x10-5				
rs13277113	AG	26	10,52	41	10,41	7	10,64	74	10,47	AG vs GG	4,5x10-5				
rs13277113	GG	31	11,06	9	10,63	47	10,77	87	10,86	AA vs GG	1,1x10-12				
* Valores de expresión normalizados para sondas para BLK (GI_33469981-S). Todos los datos de expresión son del proyecto GENEVAR (http://www.sanger.ac.uk/humgen/genevar/) .															
^ Significación de la expresión diferencial determinada por un ensayo T. (http://www.sanger.ac.uk/humgen/genevar/) .															
^ Genotipos del proyecto internacional HapMap (www.hapmap.org)															

Figura 8

Asociación de variantes de la región C8orf13/BLK y ITGAM/ITGAX con SLE por serie

Locus	Crom.	SNP	Posición	Serie 1:		Serie 2:		Serie 3:		Combinado:	
				411 casos, 1047 controles		595 casos, 1516 controles		305 casos, 777 controles		1311 casos, 3340 controles	
				Frecuencia del alelo minoritario	OR*	Frecuencia del alelo minoritario	OR	Frecuencia del alelo minoritario	OR	OR (95% c.i.)	χ^2
C8orf13/BLK	8	rs2736340	11,381,383	0,289	1,27	0,283	1,35	0,313	1,54	1,37 (1,24-1,51)	25,7
C8orf13/BLK	8	rs13277113	11,386,596	0,292	1,30	0,284	1,37	0,310	1,55	1,39 (1,26-1,54)	28,9
ITGAM/ITGAX	16	rs9937837	31,206,441	0,187	1,59	0,179	1,39	0,175	1,44	1,28 (1,16-1,41)	24,7
ITGAM/ITGAX	16	rs11574637	31,276,376	0,202	1,14	0,244	1,38	0,253	1,37	1,30 (1,17-1,45)	25,1

* Relación de probabilidad
^ Valor Meta P calculado por la adición de los valores Z de Eigenstrat ponderados ajustados para λ_{gc} residual de cada serie.

Figura 9

Asociación de variantes de C8orf13/BLK y ITGAM/ITGAX con los 11 criterios clínicos de ACR para SLE para las series 1-3								
			C8orf13/BLK: rs13277113			ITGAM/ITGAX: rs11574637		
Fenotipos	P*	Relación de probabilidad ^A (95% c.i.)	P de heterogeneidad de la serie de casos +	P	Relación de probabilidad (95% c.i.)	P de heterogeneidad de la serie de casos		
Trastorno inmunológico	0,84	-	-	0,12	-	-		
AutoAb antinucleares	0,43	-	-	0,21	-	-		
Trastorno renal	0,56	-	-	0,074	1,20 (0,98 - 1,47)	-		
Artritis	0,066	1,22 (0,99 - 1,51)	0,54	0,0045	0,73 (0,59 - 0,91)	0,42		
Trastorno hemático	0,024	1,23 (1,03 - 1,46)	0,16	0,044	1,21 (1,00 - 1,47)	0,22		
Eritema vespertilio	0,46	-	-	0,40	-	-		
Serositis	0,068	1,18 (0,99 - 1,40)	0,48	0,63	-	-		
Fotosensibilidad	0,073	1,19 (0,98 - 1,45)	0,84	0,46	-	-		
Trastorno neurológico	0,70	-	-	0,65	-	-		
Eritema discoide	0,097	1,27 (0,96 - 1,70)	0,44	0,63	-	-		
Úlceras bucales	0,17	-	-	0,10	-	-		
* Valor P para la asociación del SNP indicado con los 11 criterios clínicos de ACR en la muestra GWAS								
^A Relación de probabilidad e intervalo de confianza del 95 % para cada fenotipo. Únicamente se presentan las relaciones de probabilidad y los valores p de heterogeneidad cuando P < 0,1 de asociación								
Valor P para el ensayo de heterogeneidad de Mantel-Haenszel								

Figura 10

Asociación de variantes de C8orf13/BLK y ITGAM/ITGAX con los 11 criterios clínicos de ACR para SLE en 521 casos suecos.

Fenotipos	C8orf13/BLK: rs13277113		ITGAM/ITGAX: rs11574637	
	Relación de probabilidad [^]		Relación de probabilidad	
	P*	(95% c.i.)	P	(95% c.i.)
Trastorno inmunológico	0,740	0,94 (0,66-1,34)	0,626	0,91 (0,64-1,31)
AutoAb antinucleares	0,609	0,90 (0,60-1,35)	0,254	1,28 (0,84-1,94)
Trastorno renal	0,098	0,73 (0,50-1,06)	0,460	0,87 (0,59-1,27)
Artritis	0,378	0,88 (0,59-1,32)	0,938	1,02 (0,68-1,53)
Trastorno hemático	0,460	0,86 (0,57-1,29)	0,532	1,14 (0,75-1,74)
Eritema vespertilio	0,313	1,20 (0,84-1,70)	0,616	0,91 (0,64-1,31)
Serositis	0,056	1,45 (0,99-2,13)	0,073	1,43 (0,97-2,10)
Fotosensibilidad	0,975	1,01 (0,57-1,80)	0,363	1,31 (0,73-2,34)
Trastorno neurológico	0,269	1,22 (0,86-1,74)	0,068	1,41 (0,97-2,03)
Eritema discoide	0,526	0,89 (0,62-1,27)	0,053	1,44 (0,99-2,09)
Úlceras bucales	0,500	0,71 (0,27-1,91)	0,318	1,70 (0,59-4,91)

* Valor P para la asociación del SNP indicado con los 11 criterios clínicos de ACR

[^] Relación de probabilidad e intervalo de confianza del 95 % para cada fenotipo

Figura 11

$$Z^* = \frac{\sqrt{\frac{\sigma_{Z1}^2}{\lambda_1}}}{\sqrt{\frac{\sigma_{Z1}^2}{\lambda_1} + \frac{\sigma_{Z2}^2}{\lambda_2} + \frac{\sigma_{Z3}^2}{\lambda_3}}} Z_1 + \frac{\sqrt{\frac{\sigma_{Z2}^2}{\lambda_2}}}{\sqrt{\frac{\sigma_{Z1}^2}{\lambda_1} + \frac{\sigma_{Z2}^2}{\lambda_2} + \frac{\sigma_{Z3}^2}{\lambda_3}}} Z_2 + \frac{\sqrt{\frac{\sigma_{Z3}^2}{\lambda_3}}}{\sqrt{\frac{\sigma_{Z1}^2}{\lambda_1} + \frac{\sigma_{Z2}^2}{\lambda_2} + \frac{\sigma_{Z3}^2}{\lambda_3}}} Z_3$$

$$\sigma_Z^2 = \left(\frac{1}{n_{\text{casos}}} + \frac{1}{n_{\text{controles}}} \right) * p(1-p)$$

Figura 12A

Todos los siguientes SNP tienen EIG $P \leq 2.0E-05$													Coordenadas
SNP n.º	SNP_ID	EIG p	P en principal	P en mujeres	Cromosoma	Coordenadas	Citobanda	cM	MAF_CE U	N.º de región (SNP antes de la región)	Coordenadas	SNP después de la región	
1	rs1338078	9E-06	N/A	#N/A	1	111,93 6,717	1p13.2	132,7	0,492	1 (rs2800880)	111,932,147	rs2076590	112,065,205
2	rs10489265	3E-06	5E-04	7E-03	1	171,50 2,688	1q25.1	175,3	0,208	2 (rs3861950)	171,422,915	rs10798267	171,533,067
3	rs10497184	2E-05	8E-03	#N/A	2	157,99 2,794	2q24.1	166,9	0,038	3 (rs16841719)	157,986,114	rs10804386	158,075,558
4	rs10497185	4E-06	4E-03	#N/A	2	157,99 9,785	2q24.1	166,9	0,067	3 (rs16841719)	157,986,114	rs10804386	158,075,558
5	rs10165908	3E-06	N/A	#N/A	2	158,02 3,875	2q24.1	166,9	0,125	3 (rs16841719)	157,986,114	rs10804386	158,075,558
6	rs6738654	8E-06	N/A	#N/A	2	158,04 1,569	2q24.1	166,9	0,067	3 (rs16841719)	157,986,114	rs10804386	158,075,558
7	rs1869554	4E-06	N/A	#N/A	2	158,06 2,960	2q24.1	166,9	0,058	3 (rs16841719)	157,986,114	rs10804386	158,075,558
8	rs4550624	9E-06	7E-03	#N/A	2	158,07 2,052	2q24.1	166,9	0,061	3 (rs16841719)	157,986,114	rs10804386	158,075,558
9	rs3821236	7E-06	8E-05	9E-05	2	191,61 1,003	2q32.2	191,7	0,242	4 (rs925847)	191,605,785	rs6752770	191,681,808
10	rs10168266	3E-06	1E-04	2E-04	2	191,64 4,049	2q32.2	191,7	0,225	4 (rs925847)	191,605,785	rs6752770	191,681,808
11	rs7574865	1E-06	2E-06	2E-06	2	191,67 2,878	2q32.2	191,8	0,208	4 (rs925847)	191,605,785	rs6752770	191,681,808
12	rs978766	2E-05	5E-03	1E-02	6	72,118 181	6q13	85,7	0,392	5 (rs9455395)	72,022,392	rs9283839	72,149,401
13	rs11774146	1E-05	N/A	#N/A	8	26,536 278	8p21.2	45,8	0,059	6 (rs7011131)	26,526,139	rs327225	26,539,107
14	rs17111504	2E-05	N/A	#N/A	10	97,988 574	10q24.1	116,2	0	7 (rs1007764)	97,974,301	rs7905645	98,003,004

Figura 12B

15	rs8007536	2,E-06	1,E-03	#N/A	14	52,449,592	14q22.1	50,0	0,117	8 (rs2025631)	52,353,315	rs2210319	52,509,576
16	rs6572871	1,E-06	1,E-04	4,E-03	14	52,453,660	14q22.1	50,0	0,117	8 (rs2025631)	52,353,315	rs2210319	52,509,576
SNP n.º	SNP ID	Elig P	P en principal	P en mujeres	Cromosoma	Coordenadas	Citobanda	cM	MAF_CE U	N. de región (SNP antes de la región)	Coordenadas	SNP después de la región	Coordenadas
17	rs9937837	1,E-05	7,E-04	6,E-03	16	31,206,440	16p11.2	56,8	0,246	9 (rs889548)	31,045,213	rs8052139	31,294,363
18	rs1368779	2,E-06	2,E-05	4,E-05	18	36,315,699	18q12.3	60,8	0,085	10 (rs16973345)	36,269,528	rs9944655	36,431,363
19	rs1554945	2,E-06	1,E-05	6,E-05	18	36,318,755	18q12.3	60,8	0,083	10 (rs16973345)	36,269,528	rs9944655	36,431,363
20	rs4816519	1,E-05	4,E-04	8,E-04	21	36,416,848	21q22.12	42,6	0,119	11 (rs735142)	36,382,034	rs2252991	36,478,878

Alternancia de sin negrita sin sombrear/con negrita sombreado para la región LD

Figura 12C

SNP n.º	SNP_ID	Tamaño de la región (kb)	Causa de la elección de los genes/región	N.º de región	Primer gen en la región	Descripción	IRIS
1	rs1338078	133	Expresión inmunoespecífica	1	RAP1A	RAP1A, miembro de la familia de oncogenes RAS	3,61
2	rs10489265	110	El receptor es inmunoespecífico	2	TNFSF4	Superfamilia del factor de necrosis tumoral (ligando)	N/A
3	rs10497184	89	Expresión inmunoespecífica	3	PSCDBP	Homología con pleckstrina	11,88
4	rs10497185	89	"	3	PSCDBP	Homología con pleckstrina	11,88
5	rs10165908	89	"	3	PSCDBP	Homología con pleckstrina	11,88
6	rs6738654	89	"	3	PSCDBP	Homología con pleckstrina	11,88
7	rs1869554	89	"	3	PSCDBP	Homología con pleckstrina	11,88
8	rs4550624	89	"	3	PSCDBP	Homología con pleckstrina	11,88
9	rs3821236	76	Expresión inmunoespecífica	4	STAT4	Transductor de señales y activador de la transcripción 4	5,5
10	rs10168266	76	"	4	STAT4	Transductor de señales y activador de la transcripción 4	5,5
11	rs7574865	76	"	4	STAT4	Transductor de señales y activador de la transcripción 4	5,5
12	rs978766	127	Expresión inmunoespecífica	5	OGFR1	Receptor del factor de crecimiento opiode de tipo 1	7,61
13	rs11774146	13	AKA, CRMP2; inmunoespresado implicado en enfermedades neuroinflamatorias	6	DPYSL2	Dihidropirimidinasa de tipo 2	N/A
14	rs17111504	29	Expresión inmunoespecífica	7	BLNK	Conector de linfocitos B	4,43
15	rs8007536	156	Tiene homología con pleckstrina (inmunoespecífico); máximo nivel en el subconjunto de nefritis	8	PLEKHC1	Contiene dominio de unión a pleckstrina	N/A
16	rs6572871	156	"	8	PLEKHC1	Contiene dominio de unión a pleckstrina	N/A
17	rs9937837	249	Expresión inmunoespecífica	9	MYST1	Histona acetiltransferasa MYST	N/A
18	rs1368779	162	Máximo nivel en principal y el subconjunto de nefritis	10	.	.	N/A

Figura 12D

SNP n.º	SNP_ID	Tamaño de la región (kb)	Causa de la elección de los genes/región	N.º de región	Primer gen en la región	Descripción	IRIS
19	rs1554945	162		10			N/A
20	rs4816519	97	Expresión inmunoespecífica	11	CBR3	Carbonil reductasa 3	3,54

Alternancia de sin negrita sin sombrear/con negrita sombreado para la región LD.
Los genes dentro de regiones particulares pueden identificarse por métodos rutinarios basados en la información de las bases de datos del genoma humano. Por ejemplo, para la región 9, se determinó que los siguientes genes residen en la misma región con respecto a SNP n.º 18 - en referencia a MYST1 como gen 1, se determinó que el gen 2-7 en esta región, en orden, es PRSS8, PRSS36, FUS (IRIS = 3,46); PYCARD; ITGAM (IRIS = 5,91); e ITGAX (IRIS = 9,84). Asimismo, para la región 11, en referencia a CBR3 como gen 1, se determinó que el gen 2 era DOPEY2.

Figura 13A

Todos los siguientes SNT tienen EIG $P \leq 2.0E-05$													
SNP n.º	SNP_ID	EIG P	P en principal	Cromosoma	Coordenadas	Citobanda	cM	MAF CE U	N.º de región (SNP antes de la región)	Coordenadas	SNP después de la región	Coordenadas	Tamaño de la región (kb)
1	rs7593100	6,E-06	3,E-04	2	183,402,744	2q32.1	188,2	0,042	1 (rs10931041)	183,393,807	rs288324	183,409,833	16
2	rs6848139	9,E-06	2,E-05	4	123,614,491	4q27	124,0	0,092	2 (rs3108396)	123,241,010	rs2137497	123,777,704	53 7
3	rs10488631	5,E-08	4,E-07	7	128,381,419	7q32.1	130,1	0,167	3 (rs960633)	128,347,997	rs17340646	128,509,750	16 2
4	rs12531711	2,E-07	1,E-06	7	128,404,702	7q32.1	130,1	0,167	3 (rs960633)	128,347,997	rs17340646	128,509,750	16 2
5	rs12537284	2,E-05	8,E-05	7	128,505,142	7q32.1	130,2	0,161	3 (rs960633)	128,347,997	rs17340646	128,509,750	16 2
6	rs1961707	1,E-05	3,E-05	12	117,865,107	12q24.23	138,2	0,317	4 (rs10851051)	117,820,021	rs11069197	117,873,739	54
7	rs4238058	1,E-05	3,E-05	12	117,869,643	12q24.23	138,2	0,317	4 (rs10851051)	117,820,021	rs11069197	117,873,739	54
8	rs12964454	5,E-06	1,E-05	18	70,389,710	18q22.3	108,4	0,364	5 (rs12326826)	70,387,463	rs12957330	70,402,360	15
9	rs5943524	2,E-05	1,E-04	X	28,281,558	Xp21.3	43,5	0	6 (desconocido)				

Alternancia de sin sombreado/con negrita sombreado para la región LD

Figura 13B

SNP n.º	SNP_ID	Causa de la elección de los genes (bloque)	N.º de región	1er gen en el bloque	Descripción	IRIS	2º gen en el bloque	Descripción	IRIS	3er gen en el bloque	Descripción	IRIS
1	rs7593100	FRZB asociado con osteoartritis	1	FRZB	Proteína relacionada con frizzled	N/A	N/A		N/A			N/A
2	rs6848139	En LD con IL21; los receptores de IL2 e IL21 son inmunospecíficos	2	Tnfr		N/A	IL2	Interleucina 2	N/A	IL21	Interleucina 21	N/A
3	rs10488631	En LD con IRF5 (inmunospecífico; asociación previa)	3	IRF5	Factor 5 regulador de interferón	6,9 5	TNPO 3	Trans portina 3	N/A	LOC28 6016		N/A
4	rs12531711		3	IRF5	Factor 5 regulador de interferón	6,9 5	TNPO 3	Trans portina 3	N/A	LOC28 6016		N/A
5	rs12537284		3	IRF5	Factor 5 regulador de interferón	6,9 5	TNPO 3	Trans portina 3	N/A	LOC28 6016		N/A
6	rs1961707	Vinculación previa; SNP en región altamente conservada	4			N/A			N/A			N/A
7	rs4238058		4			N/A			N/A			N/A
8	rs12964454	Asociado con nefropatía diabética	5	CNDP1	Carosina dipeptidasa 1 (familia de metalopeptidasa M20)	N/A			N/A			N/A
9	rs5943524	Crom. X; alelo infrecuente; SNP en región de alto potencial regulador	6			N/A			N/A			N/A

Alternancia de sin negrita sin sombrear/con negrita sombreado para la región LD

Figura 14A

Todos los siguientes SNP tienen EIG $P \leq 7.0E-05$														
SNP n.º	SNP_ID	EIG P	P en mujeres	P en nefritis	Cromo- soma	Coordenadas	Citobanda	cM	MAF _{CE} U	n.º de región (SNP antes del bloque)	Coordenadas	SNP después del bloque	Tamaño del bloque (kb)	
1	rs1079826 9	2.E- 05	2.E-04	5.E-03	1	171,576,336	1q25.1	175, 4	0.375	1 (rs2096147)	171,568,92 9	rs491621 9	171,639,80 6	71
2	rs2422341	4.E- 05	3.E-04	7.E-03	1	171,590,919	1q25.1	175, 4	0.367	1 (rs2096147)	171,568,92 9	rs491621 9	171,639,80 6	71
3	rs4916334	3.E- 05	3.E-04	6.E-03	1	171,600,452	1q25.1	175, 4	0.361	1 (rs2096147)	171,568,92 9	rs491621 9	171,639,80 6	71
4	rs2894600	3.E- 05	4.E-05	4.E-05	2	226,955,846	2q36.3	229, 3	0.192	2 (rs10195178)	226,907,50 7	rs239634 0	226,961,61 7	54
5	rs3734227	6.E- 05	6.E-04	1.E-03	6	144,552,619	6q24.2	146, 8	0.342	3 (rs7452080)	144,515,35 9	rs690372 6	144,562,18 7	47
6	rs1077200 1	7.E- 05	6.E-04	7.E-03	12	9,702,918	12p13.31	24.2	0.317	4 (rs2241005)	9,639,839	rs156001 1	9,714,219	74
7	rs1053605	3.E- 05	6.E-03	3.E-04	16	54,077,108	16q12.2	70.7	0.075	5 (rs243866)	54,069,038	rs243834	54,094,188	25
8	rs4815763	7.E- 06	2.E-05	2.E-04	20	4,994,930	20p13	15.1	0.417	6 (rs404678)	4,972,237	rs242308 9	5,064,199	92

Figura 14B

S N P n.º	SNP_ID	Causa de la elección de los genes (bloque)	n.º de región	1 ^{er} gen en el bloque	Descripción	IRIS	2 ^o gen en el bloque	IRIS	3 ^{er} gen en el bloque	IRIS
1	rs10798269	Vinculación previa	1			N/A				
2	rs2422341		1			N/A				
3	rs4916334		1			N/A				
4	rs2894600	Vinculación previa	2			N/A				
5	rs3734227	Expresión inmunoespecífica	3	STX11	Sintaxina 11	5,4				
6	rs10772001	En LD con CLEC2D (LLT1, inmunoespecífico)	4	KLRB1	Subfamilia B del receptor de tipo lectina de linfocitos citolíticos	10,68	CLR	#N/A	CLEC2D	#N/A
7	rs10533605	Vinculación previa; SNP codificante (sinónimo)	5	MMP2	Metalo peptidasa 2 (gelatinasa A)	N/A				
8	rs4815763	En LD con PCNA (inmunoespecífico)	6	C20orf30	Fase de lectura abierta 30 del cromosoma 20	N/A	PCNA	5,82	CDS2	#N/A

Alternancia de sin negrita sin sombrear/con negrita sombreado para la región LD

Figura 15

<i>SNP de la región LD</i>	<i>Cromosoma</i>	<i>Coordenada A</i>	<i>Coordenada B</i>	<i>SNP A</i>	<i>SNPB</i>
1	1	111,932,147	112,065,205	rs2800880	rs2076590
2	1	171,422,915	171,533,067	rs3861950	rs10798267
3	2	157,986,114	158,075,558	rs16841719	rs10804386
4	2	191,605,785	191,681,808	rs925847	rs6752770
5	6	72,022,392	72,149,401	rs9455395	rs9283839
6	8	26,526,139	26,539,107	rs7011131	rs327225
7	10	97,974,301	98,003,004	rs1007764	rs7905645
8	14	52,353,315	52,509,576	rs2025631	rs2210319
9	16	31,045,213	31,294,363	rs889548	rs8052139
10	18	36,269,528	36,431,363	rs16973345	rs9944655
11	21	36,382,034	36,478,878	rs735142	rs2252991

Subconjunto de nefritis lúpica

Figura 16

<i>SNP de la región LD</i>	<i>Cromosoma</i>	<i>Coordenada A</i>	<i>Coordenada B</i>	<i>SNP A</i>	<i>SNP B</i>
1	2	183,393,807	183,409,833	rs10931041	rs288324
2	4	123,241,010	123,777,704	rs3108396	rs2137497
3	7	128,347,997	128,509,750	rs960633	rs17340646
4	12	117,820,021	117,873,739	rs10851051	rs11069197
5	18	70,387,463	70,402,360	rs12326826	rs12957330

Subconjunto de mujeres

Figura 17

SNP de la región LD	Cromosoma	Coordenada A	Coordenada B	SNP A	SNP B
1	1	171,568,929	171,639,806	rs2096147	rs4916219
2	2	226,907,507	226,961,617	rs10195178	rs2396340
3	6	144,515,359	144,562,187	rs7452080	rs6903726
4	12	9,639,839	9,714,219	rs2241005	rs1560011
5	16	54,069,038	54,094,188	rs243866	rs243834
6	20	4,972,237	5,064,199	rs404678	rs2423089

Grupo principal