



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 356 167**

51 Int. Cl.:
C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **05744466 .3**

96 Fecha de presentación : **03.05.2005**

97 Número de publicación de la solicitud: **1751304**

97 Fecha de publicación de la solicitud: **14.02.2007**

54 Título: **Método de diagnóstico y terapia para enfermedades asociadas con un haplotipo inflamatorio IL-1.**

30 Prioridad: **03.05.2004 US 567727 P**

45 Fecha de publicación de la mención BOPI:
05.04.2011

45 Fecha de la publicación del folleto de la patente:
05.04.2011

73 Titular/es: **INTERLEUKIN GENETICS, Inc.**
135 Beaver Street
Waltham, Massachusetts 02452, US

72 Inventor/es: **Duff, Gordon, W.;**
Kornman, Kenneth;
Wilkins, Leon;
Chen, Hong, Min y
Rogus, John

74 Agente: **Elzaburu Márquez, Alberto**

ES 2 356 167 T3

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

DESCRIPCIÓN

Método de diagnóstico y terapia para enfermedades asociadas con un haplotipo inflamatorio IL-1.

5 1. Fundamento de la invención

Genética de la agrupación de genes IL-1

La agrupación de genes IL-1 está en el brazo largo del cromosoma 2 (2q13) y contiene al menos los genes IL-1 α (IL-1A), IL-1 β (IL-1B), y el antagonista del receptor de IL-1 (IL-1RN), dentro de una región de 430 Kb (Nicklin, *et al.*, (1994) *Genomics*, 19: 382-4). Las moléculas agonistas, IL-1 α y IL-1 β , tienen una actividad pro-inflamatoria potente y están en el principio de muchas cascadas inflamatorias. Sus acciones, frecuentemente vía la inducción de otras citoquinas, tales como IL-6 e IL-8, conducen a la activación y el reclutamiento de leucocitos en el tejido dañado, la producción local de agentes vasoactivos, respuesta de fiebre en el cerebro y respuesta de la fase hepática aguda. Todas las tres moléculas IL-1 se unen a los receptores de IL-1 de tipo I y tipo II, pero solamente el receptor de tipo I transduce una señal al interior de la célula. En contraste, el receptor de tipo II es liberado de la membrana celular y actúa como receptor señuelo. El antagonista del receptor y el receptor de tipo II, por tanto, son ambos anti-inflamatorios en sus acciones.

La producción inapropiada de IL-1 desempeña un papel central en la patología de muchas enfermedades autoinmunes e inflamatorias, incluyendo artritis reumatoide; trastorno intestinal inflamatorio, psoriasis y similares. Además, hay diferencias inter-individuales estables en las tasas de producción de IL-1, y algo de esta variación puede ser explicado por diferencias genéticas en los *loci* de los genes IL-1. Por tanto, los genes IL-1 son candidatos razonables para determinar parte de la sensibilidad genética a las enfermedades inflamatorias, la mayoría de las cuales tienen una etiología multifactorial con un componente poligénico.

Determinados alelos de la agrupación de genes IL-1 se sabe que están asociados con estados morbosos particulares. Por ejemplo, para el alelo 2 de IL-1RN con repeticiones en tándem de número variable (abreviadamente VNTR por la expresión inglesa *Variable Number of Tandem Repeats*) se ha demostrado que está asociado con osteoporosis (Patente de EE.UU. N° 5.698.399), nefropatía en diabetes mellitus (Blakemore, *et al.*, (1996) *Hum. Genet.* 97(3): 369-74), alopecia areata (Cork, *et al.*, (1995) *J. Invest. Dermatol.* 104(5 Supp.): 15S-16S; Cork *et al.*, (1996) *Dermatol. Clin.* 14: 671-8), enfermedad de Graves (Blakemore, *et al.*, (1995) *J. Clin. Endocrinol.* 80(1): 111-5), lupus eritematoso sistémico (Blakemore, *et al.*, (1994) *Arthritis Rheum.* 37: 1380-85), esclerosis del liquen (Clay, *et al.*, (1994) *Hum. Genet.* 94: 407-10), y colitis ulcerosa (Mansfield, *et al.*, (1994) *Gastroenterol.* 106(3): 637-42).

Además, se ha encontrado que el alelo 2 de IL-1A del marcador -889 y el alelo 2 de IL-1B (TagI) del marcador +3954 están asociados con la periodontopatía (Patente de EE.UU. N° 5.686.246; Kornman y diGiovine (1998) *Ann. Periodont.* 3: 327-38; Hart y Kornman (1997) *Periodontol.* 2000 14: 202-15; Newman (1997) *Compend. Contin. Educ. Dent.* 18: 881-4; Kornman *et al.*, (1997) *J. Clin. Periodontol.* 24: 72-77). También se ha encontrado que el alelo 2 de IL-1A del marcador -889 está asociado con la artritis crónica juvenil, particularmente la iridociclitis crónica (McDowell, *et al.*, (1995) *Arthritis Rheum.* 38: 221-28). También se ha encontrado que el alelo 2 de IL-1B (TagI) del marcador +3954 de IL-1B está asociado con la psoriasis y la diabetes dependiente de insulina en pacientes con DR3/4 (diGiovine, *et al.*, (1995) *Cytokine* 7: 606; Pociot, *et al.*, (1992) *Eur. J. Clin. Invest.* 22: 396-402). Adicionalmente, se ha encontrado que el alelo 1 de IL-1RN (VNTR) está asociado con la retinopatía diabética (véanse las solicitudes USSN 09/037472, y PCT/GB97/02790). Además, se ha encontrado que el alelo 2 de IL-1RN (VNTR) está asociado con la colitis ulcerante en la población caucásica de Norte América y Europa (Mansfield, *J. et al.*, (1994) *Gastroenterology* 106: 637-42). Es muy interesante saber que esta asociación es particularmente fuerte con poblaciones de judíos ashkenazis étnicamente relacionados (solicitud PCT WO97/25445).

50 *Cribado de genotipos*

Los métodos tradicionales para el cribado de enfermedades heredables han dependido de la identificación de productos génicos anormales (por ejemplo, la anemia de células falciformes) o de un fenotipo anormal (por ejemplo, el retraso mental). Estos métodos son de una utilidad limitada para las enfermedades heredables que comienzan tarde y para las que no existen fenotipos fácilmente identificables, tales como, por ejemplo, las enfermedades vasculares. Con el desarrollo de la metodología del cribado genético sencillo y barato ahora es posible identificar polimorfismos que indican una propensión a desarrollar la enfermedad, incluso cuando la enfermedad es de origen poligénico. El número de enfermedades que pueden ser cribadas por métodos de la biología molecular continúa creciendo con la mayor comprensión de las bases genéticas de los trastornos multifactoriales.

El cribado genético (también denominado genotipificación o cribado molecular), puede ser definido ampliamente como un análisis para determinar si un paciente tiene mutaciones (alelos o polimorfismos) que causan un estado morbooso o están "ligadas" a la mutación que causa el estado morbooso. El ligamiento se refiere al fenómeno de que las secuencias de DNA que están próximas entre sí en el genoma tienden a ser heredadas conjuntamente. Dos secuencias pueden estar ligadas debido a alguna ventaja selectiva de la herencia conjunta. Más típicamente, sin embargo, dos secuencias polimorfas se heredan conjuntamente porque debido a la infrecuencia relativa con la que los eventos de recombinación meiótica ocurren dentro de la región entre los dos polimorfismos. Los alelos polimórficos heredados

conjuntamente se dice que se encuentran en desequilibrio del ligamiento uno con el otro debido a que, en una población humana dada, tienden ambos a estar presentes conjuntamente o de otro modo no lo están en absoluto en cualquier miembro particular de la población. Realmente, cuando se encuentran polimorfismos múltiples en una región cromosómica dada están en desequilibrio del ligamiento uno con el otro, definiendo un “haplotipo” genético cuasi estable.

5 En contraste, los eventos de recombinación que ocurren entre dos *loci* polimórficos hacen que lleguen a estar separados en distintos cromosomas homólogos. Si la recombinación meiótica entre dos polimorfismos ligados físicamente ocurre con suficiente frecuencia, los dos polimorfismos aparecerán segregados independientemente y se dice de ellos que está en equilibrio de ligamiento.

10 Aunque la frecuencia de la recombinación meiótica entre dos marcadores es generalmente proporcional a la distancia física entre ellos en el cromosoma, la presencia de “puntos calientes”, así como regiones de recombinación cromosómica reprimida puede dar como resultado discrepancias entre la distancia física y la de recombinación entre los dos marcadores. Por tanto, en determinadas regiones cromosómicas, múltiples *locis* polimórficos que se extienden en un amplio dominio cromosómico pueden encontrarse en desequilibrio de ligamiento uno con respecto al otro y,

15 por ello definir un haplotipo genético de amplia extensión. Además, cuando se encuentra una mutación causante de una enfermedad dentro de o en unión con este haplotipo, se pueden usar uno o más alelos polimórficos del haplotipo como indicador de diagnóstico o pronóstico de la probabilidad de desarrollar la enfermedad. Esa asociación entre polimorfismos de otro modo benignos y un polimorfismo causante de la enfermedad ocurre si la mutación causante de la enfermedad surgió en un pasado reciente, de modo que no ha transcurrido un tiempo suficiente para conseguir el equilibrio a través de eventos de recombinación. Por lo tanto, la identificación de un haplotipo humano que abarque o esté dentro de un cambio mutacional causante de la enfermedad, sirve como medida predictiva de la probabilidad de un individuo de tener heredada una mutación causante de la enfermedad. Importantemente, dichos métodos de pronóstico o diagnóstico se pueden utilizar sin necesidad de la identificación y aislamiento de la lesión causante de la enfermedad. Esto es significativo debido a que puede ser difícil y laboriosa la determinación precisa del defecto molecular implicado en un proceso morboso, especialmente en el caso de enfermedades multifactoriales tales como los trastornos inflamatorios.

Realmente, la correlación estadística entre un trastorno inflamatorio y un de polimorfismo de IL-1 no indica necesariamente que el polimorfismo cause directamente el trastorno. En su lugar el polimorfismo correlacionado puede ser una variante alélica benigna que está ligada a (es decir, en desequilibrio del ligamiento con) una mutación causante del trastorno que ha ocurrido en el pasado evolutivo humano reciente, de modo que no ha transcurrido suficiente tiempo para que se consiga el equilibrio a través de eventos de recombinación en el segmento cromosómico intercalado. Por tanto, para los fines de análisis para diagnóstico y pronóstico de una enfermedad particular, la detección de un alelo polimórfico asociado con la enfermedad puede ser utilizado sin la consideración de si el polimorfismo está directamente implicado en la etiología de la enfermedad. Por otra parte, cuando un *locus* polimorfo benigno está en un desequilibrio del ligamiento con un *locus* polimorfo aparentemente causante de la enfermedad, todavía son probables otros *loci* polimorfos que están en desequilibrio del ligamiento estén también en desequilibrio del ligamiento con el *locus* polimorfo causante de la enfermedad. Por tanto, estos otros *loci* polimorfos también serán un pronóstico o diagnóstico de la probabilidad de tener heredado el *locus* polimorfo causante de la enfermedad. Realmente, un haplotipo humano de amplia extensión (que describe el modelo típico de la herencia conjunta de alelos de un conjunto de marcadores polimorfos ligados) puede servir como diana para fines de diagnóstico una vez que se haya establecido una asociación entre una enfermedad o estado particular y un haplotipo humano correspondiente. Por tanto, la determinación de la probabilidad de un individuo de desarrollar una enfermedad o estado particular se puede realizar caracterizando uno o más alelos polimorfos asociados a la enfermedad (o incluso uno o más haplotipo asociados a la enfermedad) si

45 determinar o caracterizar necesariamente la variación genética causante.

2.- Sumario de la invención

50 La presente invención proporciona un método para predecir en un sujeto humano la sensibilidad a padecer la periodontopatía, que comprende la etapa de identificar en una muestra de DNA genómico de un sujeto un modelo alélico que comprende uno de los siguientes conjuntos de alelos IL-1B:

- 55 (1) (-511) alelo 1, (-1468) alelo 1 y (-3737) alelo 2;
- (2) (-511) alelo 2, (-1468) alelo 2 y (-3737) alelo 1;
- (3) (-511) alelo 1, (-1468) alelo 1 y (-3737) alelo 1; o
- 60 (4) (-511) alelo 2, (-1468) alelo 1 y (-3737) alelo 1,

en donde la presencia de uno cualquiera de los modelos 1, 3 o 4 indica que dicho sujeto ha aumentado la sensibilidad a la periodontopatía, y en donde la presencia del modelo alélico 2 indica que el sujeto tiene una sensibilidad disminuida a la periodontopatía.

65

El sujeto es homocigótico para cada uno de los alelos. La transcripción de IL-1B alterado da como resultado la producción de IL-1B alterada. El sujeto es de etnicidad asiática, tal como japonesa, china, taiwanesa o vietnamita.

ES 2 356 167 T3

Un alelo que comprende un haplotipo inflamatorio de IL-1 puede ser detectado por cualquiera de una variedad de técnicas que incluyen: 1) realizar una reacción de hibridación entre una muestra de ácido nucleico y una sonda que es capaz de hibridarse con el alelo; 2) secuenciar al menos una porción del alelo; o 3) determinar la movilidad electroforética del alelo o sus fragmentos (por ejemplo, fragmentos generados por la digestión con endonucleasas). El alelo puede ser sometido opcionalmente a una etapa de amplificación antes de la realización de la etapa de detección. Los métodos de amplificación incluyen por ejemplo la reacción en cadena de la polimerasa (PCR), la reacción en cadena de la ligasa (LCR), la amplificación con desplazamiento de cadenas (abreviadamente en lo sucesivo SDA por la expresión inglesa *Strand Displacement Amplification*), clonación y variaciones de los anteriores (por ejemplo RT-PCR y amplificación específica de alelos). Los oligonucleótidos necesarios para la amplificación se seleccionan, por ejemplo, dentro de los *loci* del gen IL-1, bien los que flanquean el marcador de interés (como se requiere para la amplificación por PCR) o bien directamente los que están solapados con el marcador (como en la hibridación de oligonucleótidos específica de alelos, abreviadamente en lo sucesivo ASO, por la expresión *Allele Specific Oligonucleotide*). Por ejemplo, la muestra se hibrida con un conjunto de cebadores, que se hibridan en 5' y 3' en una secuencia con sentido o antisentido para el alelo asociado al haplotipo inflamatorio de IL-1 y se somete a amplificación por PCR.

Un alelo que contiene un haplotipo inflamatorio de IL-1 también se detecta indirectamente, por ejemplo analizando el producto proteínico codificado por el DNA. Por ejemplo, cuando el marcador en cuestión da como resultado la traducción de una proteína mutante, la proteína se detecta por cualquiera de una variedad de métodos de detección de proteínas. Dichos métodos incluyen inmunodetección y ensayos bioquímicos, tales como fraccionamiento por tamaños, en donde la proteína experimenta un cambio en el peso molecular aparente, a través de truncamiento, prolongación, plegamiento alterado o modificaciones posteriores a la traducción alteradas.

Se describen kits para realizar los análisis antes descritos. El kit puede incluir medios de recogida de muestras de ácidos nucleicos y medios para determinar si un sujeto lleva al menos un alelo que comprende un haplotipo inflamatorio de IL-1. Opcionalmente, el kit contiene una muestra de control positivo o negativo o un patrón y/o un dispositivo algorítmico para evaluar los resultados y reactivos y componentes adicionales que incluyen, por ejemplo reactivos de amplificación del DNA, DNA-polimerasa, reactivos de amplificación de ácidos nucleicos, enzimas de restricción, tampones, un dispositivo de muestreo de ácidos nucleicos, dispositivo de purificación de DNA, desoxinucleótidos, oligonucleótidos (por ejemplo sondas y cebadores) etc.

El control puede ser un control positivo o negativo. Además, la muestra de control puede contener los productos positivos (o negativos) de la técnica de detección de alelos empleada. Por ejemplo, cuando la técnica de detección de alelos es amplificación por PCR, seguida por fraccionamiento por tamaños, la muestra de control puede comprender fragmentos de DNA fragmentos del tamaño apropiado. Análogamente, cuando la técnica de detección de alelos implica la detección de una proteína mutante, la muestra de control puede comprender una muestra de la proteína mutada. Sin embargo, se prefiere que la muestra de control contenga el material que se va analizar. Por ejemplo, los controles pueden ser una muestra del DNA genómico o una porción clonada de la agrupación de genes IL-1. Preferiblemente, la muestra de control es una muestra altamente purificada de DNA genómico en donde la muestra a analizar es el DNA genómico.

Los oligonucleótidos presentes en dicho kit se usan para amplificación de la región de interés o para la hibridación directa de oligonucleótidos específicos de alelos (ASO) a los marcadores en cuestión. Por tanto, los oligonucleótidos flanquean el marcador de interés (como es requerido para la amplificación por PCR) o están solapados directamente con el marcador (como en la hibridación ASO).

También se proporcionan ácidos nucleicos aislados de una longitud menor de 250, 200, 150, 100, 50, 25 o menos nucleótidos que contienen la secuencia de ácidos nucleicos de SEQ ID NO; 36-39. Opcionalmente, el ácido nucleico se fija a un soporte sólido o semi-sólido o está presente como una matriz (chip) de oligonucleótidos.

La información obtenida usando los análisis y kits descritos en la presente memoria (sola o junto con la información sobre otro defecto genético o factor ambiental, que contribuya a la enfermedad o estado que está asociado al haplotipo inflamatorio de IL-1) es útil para determinar si un sujeto no sintomático tiene o es probable que desarrolle la enfermedad o estado particular. Además, la información puede permitir una propuesta más personalizada para impedir el comienzo o progreso de la enfermedad o estado. Por ejemplo, esta información puede facilitar al profesional clínico prescribir más eficazmente una terapia que estará dirigida a la base molecular de la enfermedad o estado.

Se describen métodos para tratar o prevenir el desarrollo de una enfermedad o estado que está asociado con un haplotipo inflamatorio de IL-1 en un sujeto por administración al sujeto de un agente terapéutico adecuado de la invención. En incluso otro aspecto, la invención proporciona ensayos *in vitro* o *in vivo* para cribar los compuestos de ensayo para identificar agentes terapéuticos para tratar o prevenir el desarrollo de una enfermedad o estado que está asociado con un haplotipo inflamatorio de IL-1. El ensayo comprende poner en contacto una célula transfectada con una mutación causante, que está unida operativamente con un promotor apropiado, con un compuesto de ensayo y determinar el nivel de expresión de una proteína en la célula en presencia y en ausencia del compuesto de ensayo. La mutación causante da como resultado una producción disminuida del antagonista del receptor de IL-1, y una producción aumentada del antagonista del receptor de IL-1 en presencia del compuesto de ensayo indica que el compuesto es un agonista de la actividad antagonista del receptor de IL-1. Alternativamente, la mutación causante da como resultado una producción aumentada de IL-1 α o IL-1 β , y una producción disminuida de IL-1 α o IL-1 β en presencia del compuesto de ensayo indica que el compuesto es un antagonista de la actividad de IL-1 α o IL-1 β .

A no ser que se definan de otro modo todos los términos científicos y técnicos usados en la presente memoria tienen los mismos significados que son usualmente entendidos por los expertos ordinarios en la técnica a la que pertenece la presente invención. Aunque en la práctica o análisis de la presente invención se pueden usar materiales y métodos similares o equivalentes a los descritos en la presente memoria, se describen a continuación materiales y métodos adecuados. Todas las publicaciones, solicitudes de patentes, patentes, y otras referencias mencionadas en la presente memoria se incorporan como referencia en su totalidad. En caso de conflicto, prevalecerá la presente memoria, que incluye las definiciones. Además, los materiales, métodos, y ejemplos son solo ilustrativos y no se pretende que sean limitativos.

Otras realizaciones y ventajas se exponen en la siguiente descripción y reivindicaciones.

3.- Breve descripción de las figuras de los dibujos

La Fig. 1 es una representación esquemática de la agrupación de genes IL-1 que incluye unos cuantos marcadores polimórficos.

La Fig. 2 es una gráfica que representa la correlación entre los valores del desequilibrio de ligamiento y la distancia física como se describe en la presente memoria.

La Fig. 3 muestra la del secuencia ácido nucleico para IL-1A (GEN X03833; SEQ ID NO. 1).

La Fig. 4 muestra la secuencia del ácido nucleico secuencia para la IL-1B (GEN X04500; SEQ ID NO. 2).

La Fig. 5 muestra la secuencia del ácido nucleico para IL-1RN secretada (GEN X64532; SEQ ID NO. 3).

La Fig. 6 muestra construcciones con polimorfismo de un solo nucleótido (abreviadamente en lo sucesivo SNP por la expresión inglesa *Single Nucleotide Polymorphism*) dentro de la región del promotor de IL-1B. La construcción parental en donde todos los SNP son el alelo 1 están indicadas por pGL3(IL1)BS para los SNP 1-15, pGL(3-IL1)BS(3)K(M) para los SNP 2-15, y pGL3-(IL1)BL(3)KM para los SNP 2-17 (los paréntesis indican partes de los identificadores de las construcciones que no aparecen en el gráfico de la Fig. 7). Debajo de cada serie parental están las construcciones en donde los SNP individuales han sido convertidos en el alelo 2 por mutagénesis dirigida a un sitio.

Las Fig. 7 A-B muestran la localización de los SNP y las construcciones de promotor. En la Fig. 7A, las posiciones de los SNP se indican dentro de la región del promotor del gen IL-1B en la línea de la parte superior. Las líneas de la parte inferior representan las longitudes de la región del promotor usadas en las diferentes construcciones del indicador (luciferasa), cuyos resultados de actividad transcripcional se muestran en la Fig. 7B. Están indicados los sitios KpnI y BamHI. En las series de IL-1BS, los SNP están indicados por líneas verticales.

La Fig. 8 muestra construcciones de plásmidos de los SNP dentro de la región del promotor de IL-1B en la cual los SNP 14 y 15 se han mantenido mayoritariamente constantes en el alelo 2. La construcción parental, en donde todos los SNP excepto el 14 y el 15 son el alelo 1 está identificada como pGL-BS.2. Debajo están las construcciones en donde el SNP individual ha sido convertido en el alelo 2 mediante mutagénesis dirigida al sitio. Los SNP del alelo 2 SNP aparecen resaltados.

La Fig. 9A-B muestra la localización de los SNP y las construcciones del promotor. En la Fig. 9A, las posiciones de los SNP están indicadas dentro de la región del promotor del gen IL-1B en la línea de la parte superior. Las líneas inferiores representan las longitudes de la región del promotor usadas en las diferentes construcciones del indicador (luciferasa) cuyos resultados de la actividad transcripcional relativa se muestran en la Fig. 9B. Están indicados los sitios KpnI y BamHI. En la serie de IL-1BS, los SNP están indicados por líneas verticales.

La Fig. 10 es un diagrama que muestra el análisis de transfección del SNP 4 en diferentes modelos alélicos de los SNP 14 y 15. La actividad del promotor de los alelos construidos denominados A-D como se miden por la actividad de la luciferasa de luciérnaga se muestra a diferentes dosis de LPS (lipopolisacárido) (0, 1, 10 y 100 ng/ml) y 20 ng/ml de miristato-acetato de forbol (abreviadamente PMA por la expresión inglesa *Phorbol 12-myristate 13-acetate*). La actividad de luciferasa de luciérnaga se normalizó a la actividad de *Renilla* para controlar la eficacia de la transfección. Cada construcción se transfeció por triplicado y la transfección para cada construcción se repitió al menos tres veces.

La Fig. 11 es un diagrama que muestra el análisis de transfección del SNP 10. Las actividades del promotor de las construcciones A (rombo) y B (círculo) como se miden por la actividad de luciferasa de luciérnaga se muestran a diferentes dosis de LPS (0, 1, 10 y 100 ng/ml) y 20 ng/ml de PMA. La actividad de luciferasa de luciérnaga se normalizó a la actividad de *Renilla* para controlar la eficacia de la transfección. Cada construcción se transfeció por triplicado y la transfección para cada construcción se repitió al menos tres veces.

La Fig. 12 es un diagrama que demuestra el análisis de transfección de los SNP 14 y 15. La actividad del promotor de las construcciones con los modelos de alelos representados como A-D como se mide por la actividad de luciferasa de luciérnaga se muestra a diferentes dosis de LPS (0, 1, 10 y 100 ng/ml) y 20 ng/ml de PMA. La actividad de luciferasa de luciérnaga se normalizó a la actividad de *Renilla* para controlar la eficacia de la transfección. Cada construcción se transfeció por triplicado y la transfección para cada construcción se repitió al menos tres veces.

ES 2 356 167 T3

La Fig. 13 es un diagrama que muestra el análisis de transfección del haplotipo del promotor que comprende los SNP 10, 14 y 15. Las combinaciones alélicas se muestran para los plásmidos A-C. Las actividades del promotor de las construcciones A (rombo), B (cuadrado) y C (triángulo) como se miden por la actividad de luciferasa de luciérnaga se muestran a diferentes dosis de LPS (0, 1, 10 y 100 ng/ml) y 20 ng/ml de PMA. La actividad de luciferasa de luciérnaga se normalizó a la actividad de *Renilla* para controlar la eficacia de la transfección. Cada construcción se transfeció por triplicado y la transfección para cada construcción se repitió al menos tres veces.

La Fig. 14 es una imagen fotográfica de los resultados de un ensayo de cambio de movilidad electroforética del SNP 4. Se muestran las secuencias de ambos alelos 1 y 2 oligómeros bicatenarios etiquetados del SNP 4 usadas para el análisis del cambio de la movilidad (abreviadamente en lo sucesivo EMSA por la expresión inglesa *Electrophoretic Mobility Shift Assay*). Se usaron las proteínas nucleares aisladas de las células THP-1 tratadas a las 0, 2, 4, 6 y 24 horas con LPS (100 ng/ml) y PMA (20 ng/ml). Las localizaciones de la sonda libre y los complejos retardados se indican por flechas.

La Fig. 15 es una imagen fotográfica de los resultados de un ensayo de cambio de movilidad electroforética del SNP 4, usando sondas no etiquetadas para competir con la unión específica. Se muestran las secuencias de ambos alelos 1 y 2 oligómeros bicatenarios marcados del SNP 4 usadas para el ensayo EMSA. Para determinar la especificidad de unión se usaron los oligómeros bicatenarios no etiquetados para los alelos 1 y 2 en el ensayo EMSA para competir con las sondas etiquetadas.

La Fig. 16 una imagen fotográfica de los resultados del análisis de super-cambio de los anticuerpos NF-kB para el SNP 4. Los anticuerpos NF-kB para ambas subunidades p50 y p65 se usaron en el ensayo EMSA para determinar que los complejos DNA-proteínas formados fueron debidos a las proteínas NF-kB.

La Fig. 17 una imagen fotográfica de los resultados de un ensayo de cambio de movilidad electroforética del SNP 10. Se muestran las secuencias de ambos alelos 1 y 2 oligómeros bicatenarios etiquetados de SNP 10 usadas para el ensayo EMSA. Se usaron las proteínas nucleares aisladas de las células THP-1 tratadas a las 0, 2, 4, 6 y 24 horas con LPS (100 ng/ml) y PMA (20 ng/ml). Las localizaciones de la sonda libre y de los complejos retardados están indicadas por flechas.

La Fig. 18 es una imagen fotográfica de los resultados de un ensayo de cambio de movilidad electroforética del SNP 10 que muestra la especificidad de la formación del complejo para el SNP 10. Se muestran las secuencias de los alelos 1 y 2 de de los oligómeros bicatenarios etiquetados usadas para el ensayo EMSA. Para determinar la especificidad de unión los alelos 1 y 2 de los oligómeros bicatenarios no etiquetados se usaron en el ensayo EMSA para competir por la unión de las sondas etiquetadas.

La Fig. 19 es una imagen fotográfica de los resultados de un ensayo de cambio de movilidad electroforética del SNP 14. Se muestran las secuencias de los alelos 1 y 2 de de los oligómeros bicatenarios marcados del SNP 14 usadas para el ensayo EMSA. Se usaron las proteínas nucleares aisladas de las células THP-1 tratadas a las 0, 2, 4, 6 y 24 horas con LPS (100 ng/ml) y PMA (20 ng/ml). Las localizaciones de la sonda libre y de los complejos retardados están indicadas por flechas.

La Fig. 20 una imagen fotográfica de los resultados de un análisis del cambio de la movilidad electroforética resultados del SNP 15. Se muestran las secuencias de ambos alelos 1 y 2 de de los oligómeros bicatenarios etiquetados del SNP 15 usados para el ensayo EMSA. Se usaron las proteínas nucleares aisladas de las células THP-1 tratadas a las 0, 2, 4, 6 y 24 horas con LPS (100 ng/ml) y PMA (20 ng/ml). Las localizaciones de la sonda libre y de los complejos retardados están indicadas por flechas.

La Fig. 21 es una imagen fotográfica de los resultados de un ensayo de cambio de movilidad electroforética del SNP 15 que muestra la especificidad de la formación del complejo para el SNP 15. Se muestran las secuencias de los alelos 1 y 2 de de los oligómeros bicatenarios etiquetados de SNP 15 usados para el ensayo EMSA. Para determinar la especificidad de unión los alelos 1 y 2 de los oligómeros bicatenarios no etiquetados se usaron en el ensayo EMSA para competir por la unión de las sondas marcadas.

La Fig. 22 es un diagrama que muestra la gravedad de la periodontitis en sujetos por medida de la profundidad de las cavidades en las encías de los sujetos.

La Fig. 23 es un diagrama que muestra la frecuencia de los haplotipos IL-1B en sujetos japoneses y de raza blanca.

La Fig. 24 es un diagrama que muestra las cantidades relativas de IL-1 β en fluidos gingivales de individuos de raza blanca que tienen haplotipos de IL-1B que tienen haplotipos IL-1B específicos.

La Fig. 25 es un diagrama que demuestra el riesgo reducido de periodontitis in en sujetos japoneses que tienen un genotipo IL-1B específico.

La Fig. 26 es un diagrama que demuestra la asociación de haplotipos IL-1B y la gravedad de periodontitis en sujetos japoneses.

4. Descripción detallada de la invención

4.1.- Definiciones

5 Por conveniencia, se facilitan a continuación el significado de ciertos términos y frases empleados en esta parte descriptiva de la memoria, los ejemplos y las reivindicaciones anexas.

10 El término “alelo” se refiere a las diferentes variantes de secuencias encontradas en diferentes regiones polimórficas. Por ejemplo, IL-1RN (VNTR) tiene al menos cinco alelos diferentes. Las variantes de secuencias pueden ser cambios de una sola base o de múltiples bases, incluyendo sin limitación inserciones, deleciones o sustituciones, o puede ser un número variable de repeticiones de secuencias.

15 La expresión “modelo alélico” se refiere a la identidad de un alelo o alelos en una o más regiones polimórficas. Por ejemplo, un modelo alélico puede consistir en un solo alelo en un sitio polimórfico, como el alelo 1 para IL-1RN (VNTR), que es un modelo alélico que tiene al menos una copia del alelo 1 de IL-1RN en las VNTR de los *loci* del gen IL-1RN. Alternativamente, un modelo alélico puede consistir en un estado homocigótico o heterocigótico en un solo sitio polimórfico. Por ejemplo, el alelo 2,2 de IL1-RN (VNTR) es un modelo alélico en el que hay dos copias del segundo alelo en el marcador de las VNTR de IL-1RN que corresponde al estado del alelo 2 de IL-RN (VNTR) homocigótico. Alternativamente, un modelo alélico puede consistir en la identidad de alelos en más de un sitio polimórfico.

20 El término “anticuerpo” tal como se usa en la presente memoria se refiere a un agente de unión que incluye un anticuerpo completo o uno de sus fragmentos de unión que es específicamente reactivo con un polipéptido IL-1. Los anticuerpos pueden ser fragmentados usando técnicas convencionales y los fragmentos pueden ser cribados para su utilidad del modo descrito antes para los anticuerpos completos. Por ejemplo, los fragmentos F(ab)₂ pueden ser generados tratando un anticuerpo con pepsina. El fragmento F(ab)₂ resultante puede ser tratado para reducir los puentes disulfuro para producir los fragmentos Fab. Los anticuerpos de la presente invención incluyen además moléculas químicas y humanizadas, monocatenarias y bi-específicas que tienen afinidad para un polipéptido IL-1B conferida por al menos una región determinante de la complementariedad (abreviadamente en lo sucesivo CDR por la expresión inglesa *Complementarity Determining Region*) del anticuerpo.

25 “Actividad biológica” o “bioactividad” o “actividad” o “función biológica”, que se usan intercambiamente para los fines de la presente memoria significa una función efectora o antigénica que es realizada directamente o indirectamente por un polipéptido IL-1 (tanto en su conformación natural como desnaturalizada), o por cualquiera de sus sub-secuencias. Las actividades biológicas incluyen unión a un péptido diana, por ejemplo, un receptor de IL-1. Las actividades biológicas también incluyen la transcripción de un gen IL-1, tal como IL-1B. Una bioactividad de IL-1 puede ser modulada afectando directamente a un polipéptido IL-1. Alternativamente, una bioactividad de IL-1 puede ser modulada, modulando el nivel de un polipéptido IL-1, tal como modulando la expresión de un gen IL-1.

30 Como se usa en la presente memoria la expresión “fragmento bioactivo de un polipéptido IL-1” se refiere a un fragmento del polipéptido IL-1 de longitud completa, en donde el fragmento mimetiza o antagoniza específicamente la actividad del polipéptido IL-1 de tipo natural. El fragmento bioactivo es preferiblemente un fragmento capaz de interactuar con un receptor de interleuquina.

35 La expresión “una actividad aberrante”, tal como se aplica a una actividad de un polipéptido tal como IL-1, se refiere a una actividad que difiere de la actividad del polipéptido del tipo silvestre o natural del polipéptido en un sujeto sano. Una actividad de un polipéptido puede ser aberrante porque es más fuerte que la actividad de su correspondiente natural. Alternativamente, una actividad puede ser aberrante porque sea más débil o esté ausente con respecto a la actividad de su correspondiente natural. Una actividad aberrante también puede ser un cambio de actividad. Por ejemplo un polipéptido aberrante puede interactuar con un péptido diana diferente. Una célula puede tener una actividad aberrante de IL-1 debido a la sobre-expresión o sub-expresión de un gen del *locus* IL-1 que codifica un polipéptido del *locus* IL-1.

40 “Células”, “células hospedantes” o “células recombinantes” son términos usados intercambiamente en la presente memoria para referirse no solamente a la célula sujeto particular, sino a la progenie o progenie potencial de dicha célula. Debido a ciertas modificaciones puede ocurrir, en generaciones subsiguientes debido a influencias mutacionales o ambientales, que dicha progenie pueda no ser de hecho idéntica, a la célula parental, pero todavía está incluida dentro del alcance de término tal como se usa en la presente memoria.

45 Una “quimera”, “mosaico”, “mamífero quimérico” y similares, se refiere a un mamífero transgénico no humano con una construcción desactivada (*knock-out*) o activada (*knock-in*) en al menos algunas de sus células que contienen el genoma.

50 El término “animal” tal como se usa en la presente memoria nunca se refiere a un ser humano.

55 Los términos “control” o “muestra de control” se refieren a cualquier muestra apropiada para la técnica de detección apropiada empleada. La muestra de control puede contener los productos de la técnica de detección de alelos empleada o la materia que se ha de analizar. Además, los controles pueden ser positivos o negativos. A modo de ejemplo, cuando

la técnica de detección de alelos es la amplificación por PCR, seguida por fraccionamiento por tamaños, la muestra de control puede comprender fragmentos de DNA de un tamaño apropiado. Análogamente, cuando la técnica de detección de alelos implique la detección de una proteína mutada, la muestra de control puede comprender una muestra de una proteína mutante. Sin embargo, se prefiere que la muestra de control comprenda el material que se ha de analizar.

5 Por ejemplo, los controles pueden ser una muestra de DNA genómico o una porción clonada del agrupamiento de genes de IL-1. Sin embargo, cuando la muestra que se ha de analizar es DNA genómico, la muestra de control es preferiblemente una muestra de control altamente purificada de DNA genómico.

10 La frase “enfermedades y estados asociados con los polimorfismos de IL-1” se refiere a una variedad de enfermedades o estados, cuya sensibilidad a los cuales puede estar indicada en un sujeto basado en la identificación de uno o más alelos dentro del complejo IL-1. Ejemplos incluyen: enfermedad inflamatoria o degenerativa, incluyendo: respuesta inflamatoria sistémica (abreviadamente SIRS por la expresión inglesa *Systemic Inflammatory Response Syndrome*); enfermedad de Alzheimer (y estados y síntomas asociados que incluyen: neuroinflamación crónica, activación glial; aumento de las microglías; formación de placas neuríticas; y respuesta a terapia); esclerosis lateral amiotrófica (abreviadamente ALS por la expresión inglesa *Amyotrophic Lateral Sclerosis*), artritis (y estados y síntomas asociados incluyendo: artritis aguda, artritis inducida por antígenos, artritis asociada con tiroiditis linfocítica crónica, artritis inducida por colágeno, artritis crónica juvenil; artritis reumatoide juvenil, osteoartritis, prognosis y artritis inducida por estreptococos), asma (y estados y síntomas asociados, incluyendo: asma bronquial; enfermedad obstructiva crónica de las vías respiratorias; enfermedad pulmonar obstructiva crónica (EPOC), asma juvenil y asma ocupacional); enfermedades cardiovasculares (y estados y síntomas asociados, incluyendo: aterosclerosis; miocarditis autoinmunitaria, hipoxia cardíaca crónica, insuficiencia cardíaca congestiva, arteriopatía coronaria, miocardiopatía y disfunción de células cardíacas, incluyendo: activación de las células musculares lisas aórticas; apoptosis de las células cardíacas; e inmunomodulación de la función de las células cardíacas; diabetes y estados y síntomas asociados, incluyendo diabetes autoinmune, diabetes dependiente de insulina (Tipo 1), periodontitis diabética, retinopatía diabética, y nefropatía diabética); inflamaciones gastrointestinales (y estados y síntomas relacionados, incluyendo la enfermedad celiaca, osteopenia asociada, colitis crónica, enfermedad de Crohn, enfermedad inflamatoria intestinal y colitis ulcerosa); úlceras gástricas; inflamaciones hepáticas, cálculos biliares de colesterol y fibrosis hepática, infección por el VIH (y estados y síntomas relacionados, incluyendo respuestas degenerativas, respuestas neurodegenerativas, y enfermedad de Hodgkin asociada al VIH), síndrome de Kawasaki (y estados y síntomas asociados, incluyendo síndrome de los ganglios linfáticos mucocutáneos, linfadenopatía cervical, lesiones de las arterias coronarias, edema, fiebre, aumento de leucocitos, anemia leve, descamación de la piel, erupción cutánea, enrojecimiento conjuntival, trombocitosis; esclerosis múltiple, nefropatías (y estados y síntomas asociados, incluyendo nefropatía diabética, nefropatía terminal, glomerulonefritis, síndrome de Goodpasture, lesiones por supervivencia de hemodiálisis y reperfusión isquémica renal), enfermedades neurodegenerativas (y estados y síntomas asociados, incluyendo neurodegeneración aguda, inducción de IL-1 en el envejecimiento y enfermedad neurodegenerativa, plasticidad inducida por IL-1 de las neuronas hipotalámicas e hiperreactividad crónica por estrés), oftalmopatías (y estados y síntomas asociados, incluyendo retinopatía diabética, oftalmopatía de Graves, y uveítis, osteoporosis (y estados y síntomas asociados, incluyendo disminución de la masa ósea o frecuencia de fractura alveolar, femoral, radial, vertebral o de la muñeca, disminución de la masa ósea posmenopáusica, tumores, frecuencia de fracturas o tasa de disminución de la masa ósea), otitis media (en adultos o niños), pancreatitis o acinitis pancreática, periodontitis o periodontopatía (y estados y síntomas asociados, incluyendo en adultos, de iniciación temprana y diabética); enfermedades pulmonares, incluyendo neumopatía crónica, sinusitis crónica, enfermedad de las membranas hialinas, hipoxia y enfermedad pulmonar en el síndrome de muerte súbita del latente (abreviadamente SIDS por la expresión inglesa (*Sudden Infant Death Syndrome*); reestenosis; reumatismo incluyendo artritis reumatoide, nódulos de Aschoff reumáticos, enfermedades reumáticas y miocarditis reumática; tiroiditis incluyendo tiroiditis linfocítica crónica; infecciones de las vías urinarias incluyendo prostatitis crónica, síndrome del dolor pélvico crónico y urolitiasis. Trastornos inmunológicos, incluyendo enfermedades autoinmunitarias, tales como alopecia aerata, miocarditis autoinmune, enfermedad de Graves, oftalmopatías de Graves, esclerosis del liquen, esclerosis múltiple, psoriasis, lupus eritematoso sistémico, esclerosis sistémica, enfermedades del tiroides (por ejemplo bocio y estruma linfomatosa (tiroiditis de Hashimoto, bocio linfadenóide), trastornos del sueño y síndrome de fatiga crónica y obesidad (no diabética o asociada con diabetes). Resistencia a las enfermedades infecciosas, tales como leishmaniosis, lepra, enfermedad de Lyme, carditis de Lyme, malaria, malaria cerebral, meningitis, nefritis túbulo-intersticial asociada con malaria, que son causadas por bacterias, virus (por ejemplo citomegalovirus, encefalitis, virus de Epstein-Barr, virus de la inmunodeficiencia humana, virus de la gripe) o protozoos (por ejemplo, *Plasmodium falciparum*, *trypanosomas*). traumatismo, incluyendo traumatismo cerebral (incluyendo ictus e isquemias, encefalitis, encefalopatías, epilepsia, lesión cerebral perinatal, convulsiones febriles, SIDS y hemorragias subaracnoidea), bajo peso al nacer (por ejemplo parálisis cerebral), lesiones pulmonares (lesiones pulmonares hemorrágicas agudas, síndrome de Goodpasture, reperfusión isquémica aguda), disfunción miocárdica, causada por contaminantes laborales y medioambientales (por ejemplo sensibilidad al síndrome del aceite tóxico y silicosis), traumatismo por radiación, y eficacia de las respuestas de cicatrización de heridas (por ejemplo heridas por quemaduras o térmicas, heridas crónicas, heridas quirúrgicas y lesiones de la médula espinal). Sensibilidad a neoplasias, incluyendo metástasis osteolítica asociadas a cáncer de mama, caquexia, cáncer colo-rectal, enfermedades hiper-proliferantes, enfermedad de Hodgkin, leucemias, linfomas, enfermedades metabólicas y tumores, metástasis, mielomas y diversos cánceres (incluyendo de mama, próstata, ovario, colon, pulmón, etc), anorexia y caquexia. Regulación hormonal incluyendo fertilidad/fecundidad, probabilidad de un embarazo, incidencia de parto prematuro, complicaciones prenatales y neonatales incluyendo bajo peso al nacer prematuro, parálisis cerebral, septicemia, hipotiroxinemia, dependencia del oxígeno, deformidad craneal, menopausia temprana. Una respuesta del sujeto a trasplantes (rechazo o aceptación), respuesta en fase aguda (por ejemplo, respuesta febril), respuesta inflamatoria general, respuesta a la dificultad respiratoria aguda, respuesta inflamatoria sistémica aguda, cicatrización de heridas, adherencia, respuesta inmuno-inflamatoria, respuesta neuroendocrina, desarrollo y re-

ES 2 356 167 T3

sistencia a la fiebre, respuesta en fase aguda, respuesta al estrés, sensibilidad a enfermedades, estrés por movimiento repetitivo, epicondilitis, y tratamiento y respuesta del dolor.

5 Las frases “interrupción del gen” e “interrupción dianizada” o cualquier frase similar se refieren a la interrupción específica de un sitio de una secuencia de DNA de modo que se impida la expresión de ese gen en la célula en comparación con la copia de tipo natural del gen. La interrupción puede ser causada por deleciones, inserciones o modificaciones en el gen o cualquiera de sus combinaciones.

10 El término “haplotipo” como se usa en la presente memoria está destinado a referirse a un conjunto de alelos que se heredan juntos como un grupo (están en desequilibrio de ligamiento) a niveles estadísticamente significativos ($p_{\text{corr}} < 0,05$). Como se usa en la presente memoria, la frase “un haplotipo de IL-1” se refiere a un haplotipo en los *loci* de IL-1. Un haplotipo de IL-1 inflamatorio o pro-inflamatorio se refiere a un haplotipo que es indicativo de actividades agonistas aumentadas y/o antagonistas disminuidas. Un haplotipo inflamatorio o pro-inflamatorio de IL-1 se refiere también a un haplotipo que es indicativo de una expresión alterada de los genes IL-1, tales como una expresión
15 aumentada de IL-1B.

Los términos “agrupación de genes IL-1” y “*loci* de IL-1” como se usan en la presente memoria incluyen la totalidad del ácido nucleico en o cerca de la región 2q13 del cromosoma 2, incluyendo al menos los genes IL-1A, IL-1B y IL-1RN y cualesquiera otras secuencias unidas. (Nicklin *et al.*, *Genomics* 19: 382-84, 1994). Los términos “IL-1A”, “IL-1B”, y “IL-1RN” como se usan en la presente memoria se refieren a los genes que codifican los antagonistas de los receptores de IL-1A, IL-1B, y IL-1, respectivamente. Los números de acceso a los genes (en el banco de datos GenBank) para IL-1A, IL-1B, y IL-1RN son X03833, X04500, y X64532, respectivamente.

20 “Mutación funcional de IL-1” se refiere a una mutación dentro de la agrupación de genes IL-1 que da como resultado un fenotipo alterado (es decir, afecta a la función de un gen o proteína de IL-1). Ejemplos incluyen: el alelo 2 de IL-1A(+4845), el alelo 2 de IL-1B (+3954), el alelo 2 de IL-1B (+6912), el alelo 1 de IL-1B (+3737), el alelo 1 de IL-1B (-1468); y el alelo 2 de IL-1RN (+2018).

30 “Alelo Y de IL-1X (Z)” se refiere a una forma alélica particular, denominada Y, que está presente en un sitio polimórfico del *locus* IL-1 en el gen X, en donde X es IL-1A, B o RN y situado en o cerca del nucleótido Z, en donde el nucleótido Z está numerado con respecto al sitio del comienzo transcripcional principal, que es el nucleótido +1, del gen X particular de IL-1. Como se usa además en la presente memoria, la expresión “alelo (Z) de IL-1X” se refiere a todos los alelos de un sitio polimórfico en IL-1 en el gen X posicionado en o cerca del nucleótido Z. Por ejemplo, la expresión “alelo IL-1RN (+2018)” se refiere a formas alternativas del gen IL-1RN en el marcador +2018. “Alelo 1 de IL-1RN (+2018)” se refiere a una forma del gen IL-1RN que contiene una citosina (C) en la posición +2018 de la cadena con sentido. Clay *et al.*, *Hum. Genet.* 97:723-26, 1996. “Alelo 2 de IL-1RN (+2018)” se refiere a una forma del gen IL-1RN que contiene una timina en la posición (T) +2018 de de la cadena positiva (+). Cuando un sujeto tiene dos alelos IL-1RN idénticos, se dice que el sujeto es homocigótico, o que tiene un estado homocigótico. Cuando un sujeto tiene dos alelos IL-1RN diferentes, se dice que el sujeto es heterocigótico, o que tiene el estado heterocigótico. La expresión “alelo 2,2 de IL-1RN (+2018)” se refiere al estado del alelo 2 de IL-1 RN (+2018) homocigótico. Inversamente, la expresión “alelo 1,1 de IL-1RN (+2018)” se refiere al estado del alelo 1 de IL-1RN (+2018) homocigótico. La expresión “alelo 1,2 de IL-1RN (+2018)” se refiere a los estados 1 y 2 del alelo heterocigótico.

45 “Relacionado con IL-1” como se usa en la presente memoria se emplea para incluir todos los genes relacionados con los genes del *locus* IL-1 humano en el cromosoma 2 humano (2q 12-14). Estos incluyen los genes IL-1 de la agrupación de genes IL-1 humanos localizada en el cromosoma 2 (2q 13-14) que incluye: el gen IL-1A que codifica interleuquina-1 α , el gen IL-1B gene que codifica interleuquina-1 β , y el gen IL-1RN (o IL-1ra) que codifica el antagonista del receptor de interleuquina-1. Además, estos genes IL-1 relacionados incluyen los genes de los receptores de IL-1 de tipo I y tipo II localizados en el cromosoma 2 (2q 12) y sus homólogos de ratón localizados en el cromosoma 1 del ratón en la posición 19,5 cM. La interleuquina-1 α , la interleuquina-1 β y la interleuquina-1RN están relacionadas tanto más cuanto que todas se unen a los receptores de tipo I de IL-1, sin embargo sólo la interleuquina-1 α y la interleuquina-1 β son ligandos agonistas que activan los receptores de tipo I de IL-1, mientras que la interleuquina-1RN es un ligando antagonista que se presenta de modo natural. Cuando se usa el término “IL-1” con referencia un producto o polipéptido génico, pretende significar que se refiere a todos los productos génicos codificados por el *locus* de la interleuquina-1 en el cromosoma 2 (2q 12-14) humano y sus homólogos correspondientes de otras especies o sus variantes funcionales. El término IL-1 incluye por tanto los polipéptidos secretados que promueven una respuesta inflamatoria, tal como IL-1 α y IL-1 β , así como un polipéptido secretado que antagoniza las respuestas inflamatorias, tales como el antagonista del receptor de IL-1 y el receptor de tipo II de IL-1 (señuelo).

60 Un “receptor de IL-1” o “IL-1R” se refiere a receptores proteínicos unidos a la membrana de diversas células capaces de unirse a, y/o transducir, una señal de un ligando codificado por el *locus* IL-1. El término se aplica a cualquiera de las proteínas que son capaces de unirse a moléculas de interleuquina-1 (IL-1) y, en su configuración natural como proteínas de membrana en el plasma de mamífero, desempeñan presumiblemente una función en transducir la señal proporcionada por IL-1 a una célula. Como se usa en la presente memoria, el término incluye análogos de las proteínas naturales con unión a IL-1 o actividad transductora de señales. Ejemplos incluyen los receptores de IL-1 humanos y de muridos descritos en la Patente de EE.UU. N° 4.968.607. La expresión “ácido nucleico de IL-1” se refiere a un ácido nucleico que codifica una proteína IL-1.

ES 2 356 167 T3

Un “polipéptido IL-1” y “proteína IL-1” se emplea para abarcar polipéptidos que comprenden la secuencia de aminoácidos codificada por las secuencias del DNA genómico de IL-1 mostradas en la Figuras 1, 2, y 3, o sus fragmentos, y sus homólogos e incluyen polipéptidos agonistas y antagonistas.

5 “Riesgo aumentado” se refiere a una frecuencia estadísticamente superior de que se produzca la enfermedad o estado en un individuo que lleva un alelo polimórfico particular en comparación con la frecuencia de que se produzca la enfermedad o estado en un miembro de una población que no lleva el alelo polimórfico particular.

10 El término “interactuar” tal como se usa en la presente memoria se emplea para significar relaciones o asociaciones detectable (por ejemplo, interacciones bioquímicas) entre moléculas naturales, tales como interacciones entre proteína-proteína, proteína-ácido nucleico, ácido nucleico-ácido nucleico y proteína-molécula pequeña o ácido nucleico-molécula pequeña.

15 El término “aislado” como se usa en la presente memoria con respecto a ácidos nucleicos, tales como DNA o RNA, se refiere a moléculas separadas de otros DNA o RNA, respectivamente, que están presentes en la fuente natural de la macromolécula. Por ejemplo, un ácido nucleico aislado que codifica uno de los polipéptidos IL-1 objeto incluye no más de 10 kilobases (kb) de la secuencia del ácido nucleico que flanquea natural e inmediatamente el gen IL-1 en el DNA genómico, más preferiblemente no más de 5 kb de dichas secuencias flanqueantes que se presentan de modo natural, y más preferiblemente menos de 1,5 kb de dichas secuencias flanqueante que se presentan de modo natural. El término aislado como se usa en la presente memoria se refiere también a un ácido nucleico o péptido que está sustancialmente exento de material celular, material viral, o medio de cultivo cuando es producido por técnicas de DNA recombinante, o precursores químicos u otros productos químicos cuando se sintetiza químicamente. Además, un “ácido nucleico aislado” se emplea para incluir fragmentos de un ácido nucleico que no se presentan de modo natural como fragmentos y no se encontrarían en estado natural. El término “aislado” se usa también en la presente memoria para referirse a polipéptidos que están aislados de otras proteínas celulares y se emplea para abarcar tanto polipéptidos purificados como recombinantes.

25 Un animal transgénico activado (“*knock-in*”) se refiere a un animal no humano que ha tenido un gen modificado introducido en su genoma y el gen modificado puede ser de origen exógeno o endógeno.

30 Un animal transgénico inactivado (“*knock-out*”) se refiere a un animal no humano en el cual hay una supresión parcial o completa de la expresión de un gen endógeno (por ejemplo, basado en la delección de al menos una porción del gen, reemplazamiento de al menos una porción del gen con una segunda secuencia, introducción de codones de parada, mutación de bases que codifican aminoácidos críticos, o la eliminación de un empalme de intrones, etc.).

35 Una “construcción desactivada (*knock-out*)” se refiere a una secuencia de ácido nucleico que se puede usar para disminuir o suprimir la expresión de una proteína codificada por secuencias de DNA endógeno en una célula. En un ejemplo sencillo, la construcción desactivada está constituida por un gen, tal como el gen IL-1RN, con una delección en una porción crítica del mismo, de modo que no se puede expresar a partir del mismo la proteína activa. Alternativamente, a un gen natural se le pueden añadir cierto número de codones de terminación para provocar la terminación temprana de la proteína o puede ser desactivado un empalme de intrones. En una típica construcción desactivada, alguna porción del gen está reemplazada con un marcador seleccionable (tal como el gen *neo*) de modo que el gen puede ser representado como sigue: IL-1RN 5’/*neo*/ IL-1RN 3’, donde IL-1RN5’ y IL-1RN 3’, se refieren a secuencias genómicas o de cDNA que están, respectivamente en una posición situada hacia el extremo 5 y respectivamente en una posición situada hacia el extremo 3’, respectivamente con respecto a una porción del gen IL-1RN y donde *neo* se refiere a un gen de resistencia a neomicina. En otra construcción desactivada, un segundo marcador seleccionable está añadido en una posición flanqueante de modo que el gen puede ser representado como: IL-1RN/*neo*/IL-1RN/TK, donde TK es un gen de la timidina-quinasa que puede ser añadido a la secuencia IL-1RN5’ o IL-1RN3’ de la construcción precedente y que además puede ser seleccionado contra (es decir, es un marcador seleccionable negativo) en medios apropiados. Esta construcción de dos marcadores permite la selección de eventos de recombinación homólogos, que elimina el marcador con TK flanqueante, a partir de eventos de recombinación no homólogos que típicamente retienen las secuencias TK. La delección y/o el reemplazamiento de genes pueden ser de exones, intrones, especialmente empalmes de intrones, y/o las regiones reguladoras tales como los promotores.

55 “Desequilibrio de ligamiento” se refiere a la herencia conjunta de dos alelos a frecuencias mayores que la que se podría esperar a partir de las frecuencias separadas que presentan cada alelo en una población de control dada. La frecuencia esperada de presentación de dos alelos que se heredan conjuntamente es la frecuencia del primer alelo multiplicada por la frecuencia del segundo alelo. Los alelos que se presentan conjuntamente a las frecuencias esperadas se dice que están en “desequilibrio de ligamiento”. La causa del desequilibrio de ligamiento frecuentemente no está clara. Puede ser debida a la selección para ciertas combinaciones de alelos o a la mezcla reciente de poblaciones genéticamente heterogéneas. Además, en el caso de marcadores que estén vinculados muy fuertemente a un gen de una enfermedad, se espera una asociación de un alelo (o grupo de alelos unidos) con el gen de la enfermedad si la mutación de la enfermedad ocurrió en pasado reciente, de modo que no transcurrido suficiente tiempo para que sea conseguido el equilibrio a través de los eventos de recombinación en la región cromosómica específica. Cuando se hace referencia a modelos alélicos que están constituidos por más de un alelo, un primer modelo alélico está en desequilibrio de ligamiento con un segundo modelo alélico si todos los alelos que comprende el primer modelo alélico están en desequilibrio de ligamiento con al menos uno de los alelos del segundo modelo alélico. Un ejemplo de desequilibrio de ligamiento es el que ocurre entre los alelos en los sitios polimórficos IL-1RN (+2018) y IL-1RN (VNTR). Los dos

ES 2 356 167 T3

alelos en IL-1RN (+2018) están 100% en desequilibrio de ligamiento con los dos alelos más frecuentes de IL-1RN (VNTR), que son el alelo 1 y el alelo 2.

5 El término “marcador” se refiere a una secuencia del genoma que se sabe que varía entre individuos. Por ejemplo, el gen IL-1RN tiene un marcador que consiste en un número variable de repeticiones en tándem (VNTR).

10 Un “gen mutado” o “mutación” o “mutación funcional” se refiere a una forma alélica de un gen, que es capaz de alterar el fenotipo de un sujeto que tiene el gen mutado respecto a un sujeto que no tiene el gen mutado. El fenotipo alterado causado por una mutación puede ser corregido o compensado para ciertos agentes. Si un sujeto debe ser homocigótico para esta mutación que tiene un fenotipo alterado, se dice que la mutación es recesiva. Si una copia del gen mutado es suficiente para alterar el fenotipo del sujeto, se dice que la mutación es dominante. Si un sujeto tiene una copia del gen mutado y tiene un fenotipo que es intermedio entre el de un sujeto homocigótico y el de un sujeto heterocigótico (para ese gen), se dice que la mutación es co-dominante.

15 Un “animal no humano” de la invención incluye mamíferos tales como roedores, primates no humanos, ovejas, perros, vacas, cabras, etc., anfibios, tales como los miembros del género *Xenopus*, y aves transgénicas (por ejemplo, pollos, pájaros, etc.). La expresión “animal quimérico” se usa en la presente memoria para referirse a animales no humanos en los cuales se encuentra el gen recombinante, o en los cuales se expresa el gen recombinante en algunas pero no en todas las células del animal. La expresión “animal quimérico específico de tejido” indica que está presente uno de los genes IL-1 y/o está expresado o interrumpido en algunos tejidos pero no en otros. El término “mamífero” se refiere a cualquier miembro de la clase mamífero, excepto seres humanos.

20 Como se usa en la presente memoria, el término “ácido nucleico” se refiere a polinucleótidos u oligonucleótidos, tales como ácido desoxirribonucleico (DNA), y, cuando sea apropiado, ácido ribonucleico (RNA). El término debe entenderse también para incluir, como equivalentes análogos RNA o DNA hechos de análogos de nucleótidos (por ejemplo, ácidos nucleicos de péptidos) y como aplicables a la realización que describe polinucleótidos, monocatenarios (con sentido o antisentido) y bicatenarios.

30 Los términos “periodontitis” y “periodontopatía” se refieren a un trastorno dental que es el resultado del progreso de la gingivitis, que implica inflamación e infección del tejido, incluyendo ligamentos y huesos, que sostienen la dentadura. La periodontitis es frecuentemente diagnosticada por examen que muestra la encía de color rojo púrpura hinchada y blanda. Los depósitos de placa y sarro pueden ser visibles en la base de los dientes, con cavidades alargadas en las encías. Las encías son indoloras usualmente o suavemente sensibles, salvo que esté presente un absceso dental. Se pueden perder los dientes y reabsorber las encías. La gravedad de la periodontitis se cuantifica midiendo la profundidad de las cavidades (PC, también denominada “profundidad de sondeo”) en las cavidades de las encías de un sujeto.

40 El término “polimorfismo” se refiere a la coexistencia de más de una forma de un gen o una de sus porciones (por ejemplo, variante alélica). Una porción de un gen del cual hay al menos dos formas diferentes, es decir, dos secuencias de nucleótidos diferentes, se denomina una “región polimórfica de un gen”. Una secuencia genética específica en una región polimórfica de un gen es un alelo. Una región polimórfica puede ser de un solo nucleótido, cuya identidad difiere en alelos diferentes. Una región polimórfica también puede tener una longitud de varios nucleótidos.

45 La expresión “propensión a enfermedad,” también “predisposición” o “sensibilidad” a enfermedad o cualquier frase similar, significa que ciertos alelos descubiertos por la presente invención están asociados con, o son predictivos de, una incidencia del sujeto de desarrollar una enfermedad particular (por ejemplo, una enfermedad vascular). Los alelos están por tanto sobre-expresados en frecuencia en individuos con enfermedad en comparación con los individuos sanos. Por tanto, estos alelos se pueden usar para predecir la enfermedad incluso en individuos pre-sintomáticos o en un estado previo a la enfermedad.

50 “Molécula pequeña” como se usa en la presente memoria, está destinada a referirse a una composición, que tiene un peso molecular de menos de aproximadamente 5 kD y más preferiblemente aproximadamente 4 kD. Moléculas pequeñas pueden ser ácidos nucleicos, péptidos, péptido-miméticos, carbohidratos, lípidos u otras moléculas orgánicas o inorgánicas.

55 Como se usa en la presente memoria, la expresión “se hibrida específicamente” o “detecta específicamente” se refiere a la capacidad de una molécula de ácido nucleico de hibridarse a al menos aproximadamente 6 nucleótidos consecutivos de un ácido nucleico de una muestra.

60 “Secuencia reguladora transcripcional” es un término genérico usado a través de la memoria para referirse secuencias de DNA, tales como señales de iniciación, potenciadores, y promotores, que inducen o controlan la transcripción de secuencias codificadoras de proteínas con las que están unidos operativamente.

65 Como se usa en la presente memoria, el término “transgén” significa una secuencia de ácido nucleico (que codifica, por ejemplo, uno de los polipéptidos IL-1 o un de transcrito antisentido para el mismo) que ha sido introducido en una célula. Un transgén podría ser parcial o enteramente heterólogo, es decir, extraño, al animal o célula transgénico en la cual está introducido, o, es homólogo para un gen endógeno del animal o célula transgénico en el cual está introducido, pero que está diseñado para ser insertado, o está insertado, en el genoma del animal de tal modo que

altera el genoma de la célula en la cual está insertado (por ejemplo, está insertado en una localización que difiere del gen natural o su inserción da como resultado una inactivación o *knockout*). Un transgén puede estar presente también en una célula en forma de un episoma. Un transgén puede incluir una o más secuencias reguladoras de la transcripción y cualquier otro ácido nucleico, tal como intrones, que pueden ser necesarios para la expresión óptima de un ácido nucleico seleccionado.

Un “animal transgénico” se refiere a un mamífero preferiblemente no humano, un ave o un anfibio, en el cual una o más de las células del animal contienen ácido nucleico heterólogo introducido por medio de intervención humana, tal como métodos transgénicos muy conocidos por los expertos en la técnica. El ácido nucleico se introduce en la célula, directa o indirectamente mediante la introducción en un precursor de la célula, por medio de manipulación genética deliberada, tal como por micro-inyección o por infección con un virus recombinante. La expresión manipulación genética no incluye cría cruzada clásica, o fertilización *in vitro*, sino que más bien está dirigida a la introducción de una molécula de DNA recombinante. Esta molécula puede estar integrada con un cromosoma, o puede ser DNA que se replica extra-cromosómicamente. En los animales transgénicos típicos descritos en la presente memoria, el transgén hace que las células expresen una forma recombinante de uno de los polipéptidos de IL-1, por ejemplo formas agonísticas o antagonísticas. Sin embargo, también están considerados los animales transgénicos en los cuales el gen recombinante es silencioso, como por ejemplo, la construcción dependiente de recombinasa FLP o CRE descrita más adelante. Además, “animal transgénico” también incluye los animales recombinantes no humanos en los cuales la interrupción de genes de uno o más genes es causada por intervención humana, incluyendo tanto recombinación como técnicas antisentido. El término está destinado a incluir todas las generaciones de la progenie. Por tanto, están incluidos el animal fundador y todos los F1, F2, F3, y así sucesivamente, de su progenie.

El término “tratamiento” como se usa en la presente memoria se pretende que abarque el curado, así como la mejora de al menos un síntoma de un estado o enfermedad.

El término “vector” se refiere a una molécula de ácido nucleico, que es capaz de transportar otro ácido nucleico al cual ha estado unido. Un tipo de vector preferido es un episoma, es decir, un ácido nucleico capaz de replicación extra-cromosómica. Los vectores preferidos son los capaces de replicación autónoma y/o expresión de ácidos nucleicos a los cuales están unidos. Los vectores capaces de dirigir la expresión de genes a los que están unidos operativamente se denominan en la presente memoria “vectores de expresión”. En general, los vectores de expresión de utilidad en técnicas de DNA recombinante están frecuentemente en la forma de “plásmidos” que se refieren generalmente a bucles de DNA bicatenario circular lo cuales, en su forma de vectores no se encuentran en el cromosoma. En la presente memoria, “plásmido” y “vector” se usan intercambiamente puesto que el plásmido es la forma más comúnmente usada de vector. Sin embargo, la invención está destinada a incluir dichas otras formas de vectores de expresión que sirven como equivalentes funcionales y que han llegado a ser conocidos en la técnica con posterioridad a esto.

El término “alelo de tipo natural” se refiere a un alelo de un gen que, cuando está presente en dos copias en un sujeto da como resultado un fenotipo de tipo natural. Puede haber diferentes de alelos de tipo natural de un gen específico, puesto que ciertos cambios de nucleótidos en un gen no pueden afectar al fenotipo de un sujeto que tiene dos copias del gen con los cambios de nucleótidos.

4.2 Medicina predictiva

4.2.1. Haplotipos inflamatorios de IL-1 y su asociación con ciertas enfermedades o estados

La presente invención se basa al menos en parte, en la identificación de ciertos modelos de haplotipos inflamatorios y la asociación (hasta una extensión estadísticamente significativa) de estos modelos con el desarrollo de ciertas enfermedades o estados. Por tanto, la detección de los alelos que comprenden un haplotipo, solo o junto con otros medios en un sujeto puede indicar que el sujeto tiene o está predispuesto al desarrollo de una enfermedad o estado. Sin embargo, debido a que estos alelos están en desequilibrio de ligamiento con otros alelos, la detección de dichos otros alelos unidos puede indicar también que el sujeto tiene o está predispuesto al desarrollo de una enfermedad o estado particular. Por ejemplo, el haplotipo 44112332 comprende el siguiente fenotipo:

55	alelo 4 del marcador 222/223 de IL-1A
	alelo 4 del marcador gz5/gz6 de IL-1A
	alelo 1 del marcador -889 de IL-1A
60	alelo 1 del marcador +3954 de IL-1B
	alelo 2 del marcador -511 de IL-1B
	alelo 3 del marcador gaat.p33330
	alelo 3 del marcador Y31
	alelo 2 de +2018 de IL-1RN
65	alelo 1 de +4845 de IL-1A
	alelo 2 del marcador VNTR de IL-1RN

ES 2 356 167 T3

Otros tres polimorfismos en un exón alternativo de IL-1RN (exón lic, que produce una forma intracelular del producto génico) están también en desequilibrio de ligamiento con el alelo 2 de IL-1RN (VNTR) (Clay *et al.*, (1996) *Hum. Genet.* 97:723-26). Estos incluyen: el exón lic (1812) de IL-1RN (GenBank: X77090 en 1812); el polimorfismo del exón lic (1868) de IL-1RN (GenBank: X77090 en 1868); y el polimorfismo del exón lic (1887) de IL-1RN (GenBank: X77090 en 1887). Además, todavía otro polimorfismo en el promotor para la forma intracelular empalmada alternativamente del gen, el polimorfismo Pic (1731) (GenBank: X77090 en 1731), está también en desequilibrio de ligamiento con el alelo 2 del *locus* polimórfico (VNTR) de IL-1RN. Para cada uno de estos *loci*, se ha determinado la variante de secuencia del alelo 2 en desequilibrio de ligamiento con el alelo 2 del *locus* (VNTR) de IL-1RN (Clay *et al.*, (1996) *Hum. Genet.* 97:723-26).

El haplotipo 33221461 haplotipo comprende el siguiente genotipo:

alelo 3 del marcador 222/223 de IL-1A
alelo 3 del marcador gz5/gz6 de IL-1A
alelo 2 del marcador -889 de IL-1A
alelo 2 del marcador +3954 de IL-1B
alelo 1 del marcador -511 de IL-1B
alelo 4 del marcador gaat.p33330
alelo 6 del marcador Y31
alelo 1 de +2018 de IL-1RN
alelo 2 de +4845 de IL-1A
alelo 1 del marcador VNTR de IL-1RN

Los individuos con el haplotipo 44112332 son típicamente sobre-productores tanto de proteínas IL-1 α como IL-1 β , por estimulación. En contraste, los individuos con el haplotipo 33221461 son típicamente sub-productores de IL-1 α . Cada haplotipo da como resultado una respuesta pro-inflamatoria neta. Cada alelo dentro de un haplotipo puede tener un efecto, así como un efecto de genotipo compuesto. Además, pueden estar asociadas enfermedades particulares con ambos modelos de haplotipos.

La Tabla 1 siguiente expone un número de marcadores de genotipo y diversas enfermedades y estados a los cuales se ha encontrado que están asociados estos marcadores en una extensión estadísticamente significativa.

TABLA 1

Asociación de marcadores de genes del haplotipo IL-1 con ciertas enfermedades

GENOTIPO	IL-1A (-889)	IL-1A (+4845)	IL-1B (-511)	IL-1B (+3954)	IL-1RN (+2018)
ENFERMEDAD					
Periodontopatía	(*2)	*2		*2	
Estenosis de las arterias coronarias			*2		*2
Osteoporosis					*2
Diabetes dependiente de insulina				*2	
Retinopatía diabética					*1
Enfermedades renales en la fase final					(+)
Nefropatía diabética					*2
Fibrosis hepática (alcohólicos japoneses)					(+)
Alopecia areata					*2
Enfermedad de Graves					*2
Oftalmopatía de Graves					(-)
Enfermedad extratiroidal					(+)
Lupus eritematoso sistémico					*2
Esclerosis del liquen					*2
Artritis					(+)
Artritis crónica juvenil	*2				
Artritis reumatoide					(+)
Diabetes dependiente de insulina				*2	*2 VNTR
Colitis ulcerante					*2
Asma			*2	*2	
Esclerosis múltiple				(*2)	*2VNTR
Menopausia de iniciación temprana					*2

ES 2 356 167 T3

También se describen en la presente memoria alelos y haplotipos de IL-1B adicionales que modulan la expresión y los estados inflamatorios de IL-1 β . En la presente memoria se describen los SNP en las posiciones -3737 de IL-1B y -1468 de IL-1B. Los individuos con un haplotipo que incluye el alelo 1 de IL-1B (-511), el alelo 1 de IL-1B (-1468) y el alelo 2 de IL-1B (-3737) son típicamente sobre-productores de ambas proteínas IL-1 β , por estimulación. Los individuos con un haplotipo que incluye el alelo 1 de IL-1B (-511), el alelo 1 de IL-1B (-1468) y el alelo 1 de IL-1B (-3737) son también típicamente sobre-productores de ambas proteínas IL-1 β . En contraste, los individuos con un haplotipo que incluye el alelo 2 de IL-1B (-511), el alelo 2 de IL-1B (-1468) y el alelo 1 de IL-1B (-3737) son típicamente sub-productores de IL-1. Cada alelo dentro de un haplotipo puede tener un efecto, así como un efecto de genotipo compuesto. Además, pueden estar asociadas enfermedades particulares con ambos modelos de haplotipos. Debe advertirse que el SNP en la posición -1468 de IL-1B también ha sido etiquetado IL-1B (-1464) y IL-1B (-1473). (Véase, Lee, *et al.*, *J. Gastroenterol.* 2004, 39:429-433).

Además de los modelos alélicos descritos anteriormente, como se describe en la presente memoria, un experto en la técnica puede identificar fácilmente otros alelos (incluyendo polimorfismos y mutaciones) que estén en desequilibrio de ligamiento con un alelo asociado con una enfermedad o trastorno. Por ejemplo, puede recogerse una muestra de un ácido nucleico de un primer grupo de sujetos sin un trastorno particular, así como DNA de un segundo grupo de sujetos con el trastorno. Las muestras de ácido nucleico pueden ser comparadas para identificar los alelos que están sobre-representados en el segundo grupo comparado con el primer grupo, en donde dichos alelos están supuestamente asociados con un trastorno, que es causado por, o al que contribuye, una regulación inapropiada de la interleuquina 1. Alternativamente, pueden ser identificados los alelos que están en desequilibrio de ligamiento con un alelo que está asociado con el trastorno, por ejemplo, genotipificando una población grande y realizando un análisis estadístico para determinar que alelos aparecen más usualmente juntos que lo esperado. Preferiblemente el grupo se elige para estar constituido por individuos relacionados. Genéticamente los individuos relacionados incluyen individuos de la misma raza, el mismo grupo étnico, o incluso la misma familia. A medida que aumenta el grado de relación genética entre un grupo de control y un grupo de ensayo, aumenta el valor predictivo de los alelos polimórficos que están incluso más distantes ligados a un alelo causante de enfermedad. Esto es porque ha transcurrido menos tiempo evolutivo para permitir polimorfismos que estén ligados a lo largo de un cromosoma en una población fundadora para redistribuirse a través de eventos de cruzamientos genéticos. Por tanto se pueden desarrollar análisis de genotipificación para el diagnóstico específico de raza, específico de etnia e incluso específico de familia para la detección de alelos de enfermedad que surgieron en incluso tiempos más recientes de la evolución humana, por ejemplo, después de la divergencia de las principales razas humanas, después de la separación de las poblaciones humanas en distintos grupos étnicos, e incluso dentro de la historia reciente de una línea familiar particular.

El desequilibrio de ligamiento entre dos marcadores polimórficos o entre un marcador polimórfico y una mutación causante de una enfermedad es un estado metaestable. Estando ausente la presión selectiva o la presentación reiterada ligada esporádica de los eventos mutacionales subyacentes, los polimorfismos llegarán a estar finalmente disociados por los eventos de recombinación cromosómica y alcanzarán por tanto el equilibrio de ligamiento a través del curso de la evolución humana. Por tanto, la probabilidad de encontrar un alelo polimórfico en desequilibrio de ligamiento con una enfermedad o estado puede aumentar con cambios en al menos dos factores: disminución de la distancia física entre el marcador polimórfico y la mutación causante de la enfermedad, y disminución del número de generaciones meióticas disponibles para la disociación del par ligado. La consideración del último factor sugiere que, cuanto más estrechamente estén relacionados dos individuos más probablemente compartirán un cromosoma o región cromosómica parental común que contiene los polimorfismos ligados y menos probablemente que este par ligado habrá llegado a desligarse a través de los eventos de cruzamiento meiótico que ocurren en cada generación. Como resultado, cuánto más estrechamente relacionados estén dos individuos, más probablemente es que sean heredados conjuntamente los polimorfismos ampliamente espaciados. Por consiguiente, para individuos relacionados por raza, etnicidad o familia comunes, la fiabilidad de los *loci* polimórficos incluso los más distantes espaciados puede ser fiable como un indicador de la herencia conjunta de una mutación causante de una enfermedad ligada.

Se pueden diseñar sondas apropiadas para hibridarse a un gen específico del *locus* IL-1, tal como IL-1A, IL-1B o IL-1RN o un gen relacionado. Estas secuencias de DNA genómicas se muestran en las Figuras 3, 4 y 5, respectivamente, y corresponden además a la SEQ ID NOS. 1, 2 y 3, respectivamente. Alternativamente, estas sondas pueden incorporar otras regiones de *locus* genómicos relevantes, incluyendo secuencias intergénicas. Realmente la región IL-1 del cromosoma 2 humano abarca unas 400.000 pares de bases y, suponiendo una media de un polimorfismo de un solo nucleótido cada 1.000 pares de bases, incluye unos 400 loci de SNP solamente. Incluso otros polimorfismos disponibles para uso con la presente invención son obtenibles desde diversas fuentes públicas. Por ejemplo, la base de datos del genoma humano reúne los SNP intragénicos, es buscable por secuencias y usualmente contiene aproximadamente 2.700 entradas (<http://hgbase.interactiva.de>). También está disponible una base de datos de polimorfismo humano mantenida por el Massachusetts Institute of Technology (MIT SNP database (<http://www.genome.wi.mit.edu/SNP/human/index.html>)). A partir de dichas fuentes se pueden encontrar los SNP, así como otros polimorfismos humanos.

Por ejemplo, el examen de la región de IL-1 del genoma humano en una cualquiera de estas bases de datos revela que los genes de *locus* IL-1 están flanqueados por un marcador polimórfico próximo al centrómero denominado marcador microsatélite AFM220ze3 en 127,4 cM (centiMorgans) (véase N° de acceso en GenBank Z17008) y un marcador polimórfico distal denominado marcador de anclaje microsatélite AFM087xa1 en 127,9 cM (véase N° de acceso en GenBank Z16545). Estos *loci* polimórficos humanos son ambos polimorfismos microsatélites de repetición del dinucleótido CA, y, como tales, muestran un alto grado de heterocigosidad en poblaciones humanas. Por ejemplo, un alelo de AFM220ze3 genera un producto de amplificación por PCR de 211 pb con un cebador en 5' de la

ES 2 356 167 T3

secuencia TGTACCTAAGCCACCCTTTAGAGC (SEQ ID No. 4) y un cebador en 3' de la secuencia TGGCCTC CAGAAACCTCCAA (SEQ ID No. 5). Además, un alelo de AFM087xa1 genera un producto de amplificación por PCR de 177 pb con un cebador en 5' de la secuencia GCTGATATTCTGGTGGGAAA (SEQ ID No. 6) y un cebador en 3' de la secuencia GGCAAGAGCAAACTCTGTC (SEQ ID No. 7).

5 El promotor IL-1B humano es de interés en identificar polimorfismos asociados con la producción alterada de IL-1 β . Se usan dos conjuntos de cebadores para amplificar el promotor IL-1B que contienen extremos 3' diferentes. Los cebadores IL-1BF1 5'cccACGCGTGAGTGAAAGGAATCCCGTTAGAAGT (SEQ ID NO: 33) y IL-1BR2 10 5'cccACGCGTGCCTGTTGTGCCTTGTGCCTCGAAG (SEQ ID NO: 34) se usan para generar el fragmento del promotor IL-1B, IL-1BS, (+32 a -5326) (S = corto; sin el primer exón). IL-1BR1 5'cccACGCGTGGCTGCTTCAGA CACCTGTGTA (SEQ ID NO: 35) se usa para generar el fragmento del promotor IL-1BL (+471 a -5326) (L= largo; que contiene el primer exón y el SNP 17 (+45)).

15 Los cebadores equivalentes que corresponden a secuencias únicas que presentan 5' y 3' para estos polimorfismos de repetición del dinucleótido CA en el cromosoma 2 humano serán evidentes para los expertos en la técnica. Los cebadores equivalentes razonables incluyen los que se hibridan dentro de aproximadamente 1 kb del cebador diseñado, y que además están en cualquier parte desde aproximadamente 17 pb a aproximadamente 27 pb de longitud. Una directriz general para diseñar cebadores para la amplificación de secuencias genómicas cromosómicas humanas 20 únicas es que posean una temperatura de fusión de al menos aproximadamente 50°C, en donde se puede estimar una temperatura de fusión aproximada usando la fórmula $T_{\text{fusión}} = [2x(\text{número de A o T}) + 4x(\text{número de G o C})]$.

Numerosos *loci* polimórficos están presentes entre estos dos polimorfismos de repetición del dinucleótido CA y proporcionan dianas adicionales para la determinación de un alelo de pronóstico en una familia u otro grupo de individuos genéticamente relacionados. Por ejemplo, el sitio web del National Center for Biotechnology Information 25 (www.ncbi.nlm.nih.gov/genemap/) contiene un lista de cierto número de marcadores de polimorfismo en la región del *locus* IL-1 y proporciona una guía para diseñar cebadores apropiados para la amplificación y análisis de esos marcadores.

30 Por consiguiente, los segmentos nucleotídicos de la invención pueden usarse por su capacidad para formar selectivamente moléculas dúplex con tramos complementarios de cromosoma 2 humano, q12-13 o cDNA de esa región o proporcionar cebadores para la amplificación del DNA o el cDNA de esta región. El diseño de sondas apropiadas para este fin requiere la consideración de cierto número de factores. Por ejemplo, los fragmentos que tienen una longitud de entre 10, 15 o 18 nucleótidos hasta aproximadamente 20, o a aproximadamente 30 nucleótidos, encontrarán una utilidad particular. Las secuencias más largas, por ejemplo, 40, 50, 80, 90, 100, incluso hasta la longitud completa, son 35 incluso más preferidas para ciertas realizaciones. Las longitudes de oligonucleótidos de al menos aproximadamente 18 a 20 nucleótidos son bien aceptadas por los expertos en la técnica como suficiente para permitir una hibridación suficientemente específica para ser útil como sonda molecular. Además dependiendo de la aplicación prevista, se podrán emplear condiciones variables de hibridación para conseguir grados variables de selectividad de sonda para la secuencia diana. Para aplicaciones que requieren alta selectividad, se deseará típicamente emplear condiciones relativamente 40 estrictas para formar los híbridos. Por ejemplo, condiciones de sal relativamente bajas y/o alta temperatura, tales como las proporcionadas por NaCl 0,02 M – 0,15 M a temperaturas de aproximadamente 50°C a aproximadamente 70°C. Dichas condiciones selectivas pueden tolerar poco, si existe, desapareamiento entre la sonda y la cadena molde o diana.

45 En un sujeto se pueden detectar o monitorizar otros alelos u otros indicios de un trastorno junto con la detección de los alelos descritos antes, por ejemplo, identificar espesores de paredes de vasos sanguíneos (por ejemplo, como se mide por ultrasonido), o si el sujeto es fumador, bebedor, tiene sobrepeso o si está sometido a estrés o a ejercicios.

50 4.2.2 Detección de alelos

Hay disponibles muchos métodos para detectar alelos específicos en *loci* polimórficos humanos. El método preferido para detectar un alelo polimórfico específico dependerá, en parte, de la naturaleza molecular del polimorfismo. Por ejemplo, las diversas formas alélicas del *locus* polimórfico pueden diferir en un solo par de bases del DNA. Dichos 55 polimorfismos de un solo nucleótido (o SNP) son los principales contribuyentes a la variación genética, que comprenden un 80% de todos los polimorfismos conocidos, y su densidad en el genoma humano se estima que es por término medio de 1 por cada 1.000 pares de bases. Los SNP son muy frecuentemente bialélicos que se presentan en solamente dos formas diferentes (aunque son teóricamente posibles hasta cuatro formas diferentes de un SNP, correspondientes a las cuatro bases de nucleótidos diferentes que existen en el DNA). No obstante, los SNP son mutacionalmente más 60 estables que otros polimorfismos, lo que les hace adecuados para estudios de asociación en los cuales el desequilibrio de ligamiento entre marcadores y una variante desconocida se usa para cartografiar mutaciones causantes de una enfermedad. Además, debido a que los SNP tienen típicamente solo dos alelos, pueden ser sometidos a genotipificación mediante un sencillo ensayo más/menos en lugar de una medida prolongada, lo que les hace más susceptibles de automatización.

65 Una variedad de métodos está disponible para detectar la presencia de polimorfismo de un solo nucleótido particular en un individuo. Los avances en este campo han proporcionado precisión, facilidad y genotipificación de los SNP barato y a gran escala. Muy recientemente, por ejemplo, se han descrito nuevas técnicas que incluyen hibridación

dinámica específica de alelos (abreviadamente DASH por la expresión inglesa *Dynamic Allele-Specific Hybridization*), electroforesis en gel diagonal en matrices de microplacas, abreviadamente MADGE por la expresión inglesa *Microplate Array Diagonal Gel Electrophoresis*), piro-secuenciación, ligamiento específico de oligonucleótidos, el sistema TaqMan, así como diversas tecnologías de "chips" de DNA, tal como los chips Affymetrix SNP. Estos métodos requieren la amplificación de la región genética diana, típicamente por PCR. Incluso otros métodos recientemente desarrollados, basados en la generación de pequeñas moléculas de señal por escisión invasiva seguido por espectrometría de masas o sondas candados inmovilizadas y amplificación por círculo rodante, podrían finalmente eliminar la necesidad de la PCR. Se recogen resumidamente más adelante varios de los métodos conocidos en la técnica para detectar polimorfismos de un solo nucleótidos específicos. El método de la presente invención se entiende que incluye todos los métodos disponibles.

Se han desarrollado varios métodos para facilitar tal análisis de los polimorfismos de un solo nucleótido. En una realización, el polimorfismo de una sola base se puede detectar usando un nucleótido resistente a exonucleasas especializadas, como se ha descrito, por ejemplo, en Mundy, C. R. (Patente de EE.UU. N° 4.656.127). De acuerdo con el método, se permite que un cebador complementario a la secuencia alélica inmediatamente 3' respecto al sitio polimórfico se hibride a una molécula diana obtenida de un animal o ser humano particular. Si el sitio polimórfico en la molécula diana contiene un nucleótido que es complementario al derivado de nucleótido resistente a la exonucleasa particular presente, entonces el derivado se incorporará sobre el extremo del cebador hibridado. Dicha incorporación convierte al cebador en resistente a la exonucleasa, y por ello permite su detección. Puesto que se conoce la identidad del derivado resistente a la exonucleasa de la muestra, un hallazgo de que el cebador ha llegado a ser resistente a las exonucleasas revela que el nucleótido presente en el sitio polimórfico de la molécula diana era complementario al del derivado de nucleótido usado en la reacción. Este método tiene la ventaja de que no requiere la determinación de grandes cantidades de datos de secuencias.

Se describe un método basado en una disolución para determinar la identidad del nucleótido de un sitio polimórfico. Cohen, D. *et al.* (Patente francesa 2.650.840; solicitud de patente PCT N° WO91/02087). Como en el método de Mundy de la patente de EE.UU. N° 4.656.127, se emplea un cebador que es complementario a las secuencias alélicas inmediatamente 3' a un sitio polimórfico. El método determina la identidad del nucleótido de ese sitio usando derivados de didesoxinucleótidos marcados, que, si son complementarios al nucleótido del sitio polimórfico se incorporarán en (extremo) del cebador.

Se describe un método alternativo, conocido como Genetic Bit Analysis o GBA™ Goelet, P. *et al.*, (solicitud de patente PCT N° 92/15712). El método de Goelet, P. *et al.*, usa mezclas de terminadores marcados y un cebador que es complementario a la secuencia 3' para un sitio polimórfico. El terminador marcado que se incorpora se determina por tanto por el, y es complementario al, nucleótido presente en el sitio polimórfico de la molécula diana que se evalúa. En contraste con el método de Cohen *et al.*, (Patente francesa 2.650.840; solicitud PCT N° WO91/02087) el método de Goelet, P. *et al.*, es preferiblemente un ensayo en fase heterogénea, en el que el cebador o la molécula diana está inmovilizado en un fase sólida.

Recientemente, se han descrito varios métodos de incorporación de nucleótidos guiados por cebador para analizar sitios polimórficos en DNA (Komher, J. S. *et al.*, *Nucl. Acids. Res.* 17:7779-7784 (1989); Sokolov, B. P., *Nucl. Acids Res.* 18:3671 (1990); Syvanen, A. -C., *et al.*, *Genomics* 8:684-692 (1990); Kuppaswamy, M. N. *et al.*, *Proc. Natl. Acad. Sci. (U.S.A.)* 88:1143-1147 (1991); Prezant, T. R. *et al.*, *Hum. Mutat.* 1:159-164 (1992); Ugozzoli, L. *et al.*, *GATA* 9:107-112 (1992); Nyren, P. *et al.*, *Anal. Biochem.* 208:171-175 (1993)). Estos métodos difieren del GBA™ en que todos se basan en la incorporación de desoxinucleótidos marcados para discriminar entre bases en un sitio polimórfico. En dicho formato, puesto que la señal es proporcional al número de desoxinucleótidos incorporados, los polimorfismos que pueden estar presentes en experimentos del mismo nucleótido pueden dar como resultado señales que son proporcionales a la longitud del experimento (Syvanen, A. -C., *et al.*, *Amer. J. Hum. Genet.* 52:46-59 (1993)).

Para mutaciones que producen terminación prematura de la traducción de proteínas, el ensayo de truncamiento de proteínas (abreviadamente en lo sucesivo PTT por la expresión inglesa *Protein Truncation Test*) ofrece una aproximación eficaz al diagnóstico (Roest, *et al.*, (1993) *Hum. Mol. Genet.* 2:1719-21; van der Luijt, *et al.*, (1994) *Genomics* 20:1-4). Por el PTT, el RNA se aísla inicialmente a partir de tejido disponible y se somete a transcripción inversa, y el segmento de interés se amplifica por PCR. Los productos de la transcripción inversa y la PCR se usan luego como molde para la amplificación por PCR anidada con un cebador que contiene un promotor de RNA-polimerasa y una secuencia para iniciar la traducción en eucariotas. Después de la amplificación de la región de interés, los restos únicos incorporados en el cebador permiten la transcripción y traducción secuenciales *in vitro* de los productos de la PCR. Por electroforesis en gel de dodecil-sulfato sódico-poliacrilamida de los productos de la traducción, la aparición de polipéptidos truncados señala la presencia de una mutación que causa una terminación prematura de la traducción. En una variación de esta técnica, el DNA (en oposición al RNA) se usa como molde para la PCR cuando la región diana de interés se deriva de un solo exón.

Se puede utilizar cualquier tipo de célula o tejido para obtener muestras de ácido nucleico para uso en los diagnósticos descritos en la presente memoria. En una realización preferida, la muestra de DNA se obtiene a partir de un fluido corporal, por ejemplo, sangre, obtenida por técnicas conocidas (por ejemplo, punción en vena) o saliva. Alternativamente, el ensayo del ácido nucleico se puede realizar en muestras secas (por ejemplo, cabello o piel). Cuando se usa RNA o proteína, las células o tejidos que pueden ser utilizados deben expresar un gen IL-1.

Los métodos de diagnóstico se pueden realizar también directamente *in situ* en secciones de tejidos (fijadas y/o congeladas) de tejido de pacientes obtenidos de biopsias o resecciones, tales que no es necesaria una purificación del ácido nucleico. Los reactivos de ácido nucleico se pueden usar como sondas y/o cebadores para tales métodos *in situ* (véase, por ejemplo, Nuovo, G.J., 1992, *PCR in situ hybridization: protocols and applications*, Rayen Press, NY).

Además de los métodos que ponen el énfasis principalmente en la detección de una secuencia de ácido nucleico, también se pueden determinar perfiles en dichos esquemas de detección. Los perfiles de las huellas pueden ser generados, por ejemplo, utilizando un método de presentación diferencial, análisis Northern y /o RT-PCR.

Un método de detección preferido es una hibridación específica de alelos que usa sondas que solapan una región de al menos un alelo de un haplotipo pro-inflamatorio de IL-1 y que tienen aproximadamente 5, 10, 20, 25 o 30 nucleótidos alrededor de la mutación o región polimórfica. En una realización preferida de la invención, se unen a un soporte en fase sólida varias sondas capaces de hibridarse específicamente a otras variantes alélicas implicadas en una reestenosis, por ejemplo, un “*chip*” (que puede contener hasta aproximadamente 250.000 oligonucleótidos). Los oligonucleótidos pueden estar unidos a un soporte sólido mediante una variedad de procesos, incluyendo litografía. El análisis de la detección de la mutación que usa estos *chips* que comprenden oligonucleótidos, también denominado “matrices de sondas de DNA” está descrito por ejemplo por Cronin *et al.*, (1996) *Human Mutation* 7:244. En una realización, un *chip* comprende todas las variantes alélicas de al menos una región polimórfica. El soporte en fase sólida se pone luego en contacto con un ácido nucleico de ensayo y se detecta la hibridación a las sondas específicas. Por consiguiente, se puede identificar en un sencillo experimento de hibridación la identidad de numerosas variantes alélicas de uno o más genes.

Estas técnicas pueden comprender también la etapa de amplificar el ácido nucleico antes del análisis. Los métodos de amplificación son conocidos por los expertos en la técnica e incluyen, pero sin limitación, clonación, reacción en cadena de la polimerasa (PCR), reacción en cadena de la polimerasa de alelos específicos (ASA), reacción en cadena de la ligasa (LCR), reacción en cadena con la polimerasa anidada, replicaciones de secuencias auto-sostenida (Guatelli, J.C. *et al.*, 1990, *Proc. Natl. Acad. Sci. USA*, 87:1874-1878), sistema de amplificación transcripcional (Kwoh, D.Y. *et al.*, 1989, *Proc. Natl. Acad. Sci. USA* 86:1173-1177), y *Q-Beta Replicase* (Lizardi, P.M. *et al.*, 1988, *Bio/Technology* 6:1197).

Los productos de amplificación pueden analizarse por una variedad de modos, incluyendo análisis por tamaños, digestión por restricción seguida de análisis por tamaños, detección de cebadores oligonucleotídicos con una etiqueta específica en los productos de reacción, hibridación de oligonucleótidos específica de alelos (ASO), detección por 5' exonucleasa específica de alelos, secuenciación, hibridación, y similares.

La detección basada en PCR significa que puede incluir amplificación múltiple de una variedad de marcadores simultáneamente. Por ejemplo, es bien conocido en la técnica seleccionar cebadores de PCR para generar productos de PCR que no se solapan en tamaño y pueden ser analizados simultáneamente. Alternativamente, es posible amplificar diferentes marcadores con cebadores que están etiquetados diferentemente y por tanto cada uno puede ser detectado diferencialmente. Naturalmente, la detección basada en hibridación significa permitir la detección diferencial de múltiples productos de PCR en una muestra. En la técnica se conocen otros métodos para permitir análisis múltiples de una pluralidad de marcadores.

En una realización meramente ilustrativa, el método incluye las etapas de: (i) recoger una muestra de células de un paciente, (ii) aislar el ácido nucleico (por ejemplo, genómico, mRNA o ambos) de las células de la muestra, (iii) poner en contacto la muestra de ácido nucleico con uno o más cebadores que se hibridan específicamente en 5' y 3' a al menos un alelo de un haplotipo pro-inflamatorio IL-1 en condiciones tales que ocurre la hibridación y la amplificación del alelo, y (iv) detectar el producto de amplificación. Estos esquemas de detección son especialmente útiles para la detección de moléculas de ácido nucleico si dichas moléculas están presentes en números muy bajos.

En una realización preferida del análisis pertinente, el alelo de un haplotipo pro-inflamatorio de IL-1 se identifica por alteraciones en los modelos de escisión por enzimas de restricción. Por ejemplo, se aíslan el DNA de muestra y de control, se amplifica (opcionalmente), se digiere con una o más endonucleasas de restricción, y se determinan por electroforesis en gel los tamaños de la longitud de los fragmentos.

En incluso otra realización, se pueden usar cualquiera de las reacciones de secuenciación conocidas en la técnica para secuenciar directamente el alelo. Reacciones de secuenciación ilustrativas incluyen las basadas en las técnicas desarrolladas por Maxam y Gilbert ((1977) *Proc. Natl Acad Sci USA*, 74:560) o por Sanger (Sanger *et al.*, (1977) *Proc. Nat. Acad. Sci USA* 74:5463). También se contempla que cualquiera de una variedad de métodos de secuenciación automatizados pueda ser utilizada cuando se realizan los análisis pertinentes (véase, por ejemplo *Biotechniques* (1995) 19:448), que incluyen secuenciar por espectrometría de masas (véase por ejemplo la publicación de la solicitud de patente PCT WO 94/16101; Cohen *et al.*, (1996) *Adv. Chromatogr.*, 36:127-162; y Griffin *et al.*, (1993) *Appl. Biochem. Biotechnol.*, 38:147-159). Será evidente para los expertos en la técnica que, para ciertas realizaciones, se necesita determinar la presencia de solamente uno, dos o tres de las bases de los ácidos nucleicos en la reacción de secuenciación. Por ejemplo, se puede llevar a cabo una huella de A (adenina) o similar, por ejemplo, cuando solamente se detecte un ácido nucleico.

En una realización adicional, se puede usar la protección de los agentes de escisión (tales como una nucleasa, hidroxilamina o tetróxido de osmio y/con piperidina) para detectar bases desapareadas en heterodúplex RNA/DNA o RNA/DNA o DNA/DNA (Myers, *et al.*, (1985) *Science*, 230:1242). En general, la técnica de “escisión de desapareamientos” comienza proporcionando los heterodúplex formados hibridando RNA o DNA (marcados) que contienen el alelo de tipo natural con la muestra. Los dúplex bicatenarios se tratan con un agente que escinde las regiones monocatenarias de los dúplex, tales como las que existirán debido a los desapareamientos de bases entre las cadenas de control y de la muestra. Por ejemplo, los dúplex RNA/DNA se pueden tratar con RNasa y los híbridos DNA/DNA se tratan con nucleasa S1 para digerir enzimáticamente las regiones desapareadas. En otras realizaciones, los dúplex DNA/DNA o RNA/DNA se pueden tratar con hidroxilamina o tetróxido de osmio y con piperidina para digerir las regiones desapareadas. Después de la digestión de las regiones desapareadas, el material resultante se separa luego por tamaño sobre geles de poliacrilamida desnaturizantes para determinar el sitio de la mutación. Véase, por ejemplo, Cotton *et al.*, (1988) *Proc. Natl. Acad. Sci. USA*, 85:4397; y Saleeba *et al.*, (1992) *Métodos Enzymol.* 217:286-295. En una realización preferida, el DNA o RNA de control se puede etiquetar para detección.

En incluso otra realización, la reacción de escisión de desapareamientos emplea una o más proteínas que reconocen pares de bases desapareadas en los DNA bicatenarios (las denominadas “enzimas de reparación de desapareamientos de DNA”). Por ejemplo, la enzima *mutY* de *E. coli* escinde los desapareamientos A en G/A y la timidina DNA-glicosilasa de las células HeLa escinde los desapareamientos T en G/T (Hsu *et al.*, (1994) *Carcinogenesis*, 15:1657-1662). De acuerdo con una realización ilustrativa, una sonda basada en el alelo del haplotipo del locus de IL-1 se hibrida a un cDNA u otro de producto de DNA de una(s) célula(s) de ensayo. El dúplex se trata con la enzima reparadora de desapareamientos del DNA, y los productos escindidos, si los hubiere, pueden ser detectados por protocolos de electroforesis o similares. Véase, por ejemplo, la Patente de EE.UU. N° 5.459.039.

En otras realizaciones, las alteraciones en la movilidad electroforética se usarán para identificar un alelo del locus IL-1. Por ejemplo, el polimorfismo conformacional monocatenario (abreviadamente en lo sucesivo SSCP por la expresión inglesa *Single Strand Conformation Polymorphism*) se puede usar para detectar diferencias en la movilidad electroforética entre los ácidos nucleicos de tipo mutante y de tipo natural (Orita *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:2766, véase también Cotton (1993) *Mutat. Res.* 285:125-144; y Hayashi (1992) *Genet. Anal. Tech. Appl.*, 9:73-79). Los fragmentos de DNA monocatenarios de los alelos de IL-1 de la muestra y del control se desnaturizan y se les deja re-naturalizar. La estructura secundaria de los ácidos nucleicos monocatenarios varía de acuerdo con la secuencia, y la alteración resultante en la movilidad electroforética permite la detección de incluso un cambio de una sola base. Los fragmentos de DNA pueden ser etiquetados o detectados con sondas etiquetadas. La sensibilidad del análisis puede ser mejorada usando RNA (en lugar de DNA), en el cual la estructura secundaria es más sensible a un cambio en la secuencia. En una realización preferida, el método pertinente utiliza análisis de heterodúplex para separar moléculas de heterodúplex bicatenarias basándose en cambios de la movilidad electroforética (Keen *et al.*, (1991) *Trends Genet.*, 7:5).

En todavía otra realización, se analiza el movimiento de los alelos en geles de poliacrilamida que contienen un gradiente de desnaturizante usando electroforesis de gen en gradiente (abreviadamente en lo sucesivo DGGE por la expresión inglesa *Denaturing Gradient Gel Electrophoresis*) (Myers *et al.*, (1985) *Nature*, 313:495). Cuando la DGGE se usa como el método de análisis, se modificará el DNA para asegurar que no se desnaturiza completamente, por ejemplo añadiendo una pinza o abrazadera de GC de aproximadamente 40 pb de DNA rico en GC de alto punto de fusión por PCR. En una realización adicional, se usa un gradiente de temperatura en lugar de un gradiente de agente desnaturizante para identificar diferencias en la movilidad del DNA de la muestra y del control (Rosenbaum y Reissner (1987) *Biophys. Chem.* 265:12753).

Ejemplos de otras técnicas para detectar alelos incluyen, pero sin limitación, hibridación de oligonucleótidos selectivos, amplificación selectiva o extensión de cebadores selectivos. Por ejemplo, pueden prepararse cebadores “oligonucleotídicos” en los cuales la mutación o la diferencia de nucleótidos conocida (por ejemplo, en variantes alélicas) es colocada centralmente y luego hibridada a un DNA diana en condiciones que permiten la hibridación solamente si se encuentra un apareamiento perfecto (Saiki *et al.*, (1986) *Nature*, 324:163; Saiki *et al.*, (1989) *Proc. Natl. Acad. Sci. USA* 86:6230). Dichas técnicas de hibridación de oligonucleótidos específica de alelos se puede usar para analizar una mutación o región polimórfica por reacción cuando los oligonucleótidos se hibridan a un DNA diana amplificado por PCR o un número de diferentes mutaciones o regiones polimórficas cuando los oligonucleótidos se unen a la membrana hibridante y se hibridan con el DNA diana etiquetado.

Alternativamente, se puede usar junto con la presente invención la tecnología de amplificación específica de alelos que depende de la amplificación por PCR selectiva. Los oligonucleótidos usados como cebadores para la amplificación específica pueden llevar la mutación o la región polimórfica de interés en el centro de la molécula (de modo que la amplificación depende de la hibridación diferencial) (Gibbs *et al.*, (1989) *Nucleic Acids Res.*, 17:2437-2448) o en el extremo 3' de un cebador donde, en condiciones apropiadas, puede impedir o reducir la extensión por la polimerasa (Prossner (1993) *Tibtech.*, 11:238. Además puede ser deseable introducir un nuevo sitio de restricción en la región de la mutación para crear una detección basada en la escisión (Gasparini *et al.*, (1992) *Mol. Cell Probes*, 6:1). Se anticipa que en ciertas realizaciones la amplificación se puede realizar también usando Taq ligasa para la amplificación (Barany (1991) *Proc. Natl. Acad. Sci. USA*, 88:189). En tales casos, la ligación ocurrirá solamente si hay un perfecto apareamiento en el extremo 3' de la secuencia en 5' que lo haga posible para detectar la presencia de una mutación conocida en un sitio específico buscando la presencia o ausencia de la amplificación.

En otra realización la identificación de la variante alélica se lleva a cabo usando un ensayo de ligamiento de oligonucleótidos (abreviadamente en lo sucesivo OLA por la expresión inglesa *Oligonucleotide Ligation Assay*), como se describe, por ejemplo, en la patente de EE.UU. N° 4.998.617 y en Landegren, U. *et al.*, ((1988) *Science*, 241:1077-1080). El protocolo OLA usa dos oligonucleótidos que están diseñados para ser capaces de hibridarse a secuencias contiguas de una sola cadena de una diana. Uno de los oligonucleótidos se une a un marcador de separación, por ejemplo, biotinilado y el otro está etiquetado detectablemente. Si la secuencia complementaria precisa se encuentra en una molécula diana, los oligonucleótidos se hibridarán de tal modo que sus extremos estén contiguos y creen un sustrato de ligamiento. El ligamiento permite entonces que el oligonucleótido marcado sea recuperado usando avidina u otro ligando de biotina. Nickerson, D. A. *et al.*, han descrito un análisis de detección de ácidos nucleicos que combina los atributos de la PCR y el OLA (Nickerson, D. A. *et al.*, (1990) *Proc. Natl. Acad. Sci. USA*, 87:8923-27). En este método, la PCR se usa para conseguir la amplificación exponencial del DNA diana, que luego se detecta usando el OLA.

Se han desarrollado varias técnicas basadas en este método OLA y se pueden usar para detectar alelos de un haplotipo del *locus* IL-1. Por ejemplo, la Patente de EE.UU. N° 5.593.826 describe un OLA que usa un oligonucleótido que tiene un grupo 3'-amino y un oligonucleótido 5'-fosforilado para formar un conjugado que tiene un enlace fosforamidato. En otra variación del OLA descrito en Tobe *et al.*, ((1996) *Nucleic Acid Res.*, 24: 3728), el OLA combinado con la PCR permite determinar el tipo de dos alelos en un solo pocillo de microtitulación. Marcando cada uno de los cebadores específicos de alelos con un hapteno único, es decir, digoxigenina y fluoresceína, cada reacción de OLA puede ser detectada usando anticuerpos específicos del hapteno que están etiquetados con diferentes informadores enzimáticos, fosfatasa alcalina o peroxidasa de rábano silvestre. Este sistema permite la detección de los dos alelos usando un formato de alto rendimiento que conduce a la producción de dos colores diferentes.

Otra realización de la invención está dirigida a kits para detectar una predisposición para desarrollar una reestenosis. Este kit puede contener uno o más oligonucleótidos, incluyendo 5' y 3' oligonucleótidos que se hibridan 5' y 3' a al menos un alelo de un haplotipo del *locus* IL-1. Los oligonucleótidos de la amplificación por PCR deben hibridarse entre 25 y 2500 pares de bases separadas, preferiblemente entre aproximadamente 100 y aproximadamente 500 bases separadas, para producir un producto de PCR de tamaño conveniente para análisis posteriores.

Los cebadores particularmente preferidos para uso en el método de diagnóstico de la invención incluyen las SEQ ID NO: 8-23 y 36-39.

El diseño de oligonucleótidos adicionales para la amplificación y detección de alelos polimórficos de IL-1 por el método de la invención se facilita por la disponibilidad de tanto información actualizada de secuencias del cromosoma humano 2q13 - que contienen el *locus* IL-1 humano-, como de información actualizada de polimorfismos humanos disponible para este *locus*. Por ejemplo, la secuencia de DNA para IL-1A, IL-1B y IL-1RN se muestra en la Figuras 1 (N° de acceso en GenBank X03833), 2 (N° de acceso en GenBank X04500) y 3 (N° de acceso en GenBank X64532) respectivamente. Los cebadores adecuados para la detección de un polimorfismo humano en estos genes pueden ser diseñados fácilmente usando información de secuencias y métodos estándares conocidos en la técnica para el diseño y optimización de secuencias de cebadores. El diseño óptimo de dichas secuencias de cebadores se puede conseguir, por ejemplo, mediante el uso de programas de selección de cebadores comercialmente disponibles, tales como Primer 2.1, Primer 3 o GeneFisher (véase también, Nicklin M.H.J., Weith A. Duff G.W., "A Physical Map of the Region Encompassing the Human Interleukin-1 α , interleukin-1 β , and Interleukin-1 Receptor Antagonist Genes" *Genomics*, 19: 382 (1995); Nothwang H.G., *et al.*, "Molecular Cloning of the Interleukin-1 gene Cluster: Construction of an Integrated YAC/PAC Contig and a partial transcriptional Map in the Region of Chromosome 2q13" *Genomics*, 41: 370 (1997); Clark, *et al.*, (1986) *Nucl. Acids. Res.*, 14:7897-7914 [la lista de erratas publicadas aparece en *Nucl. Acids Res.*, 15:868 (1987) y el proyecto de *Genome Database* (GDB) en la URL <http://www.gdb.org>).

Para uso en un kit, los oligonucleótidos pueden ser de cualquier variedad de composiciones naturales y/o sintéticas, tales como oligonucleótidos sintéticos, fragmentos de restricción, cDNA, ácidos nucleicos de péptidos sintéticos (abreviadamente en lo sucesivo PNA por la expresión inglesa *Peptide Nucleic Acids*), y similares. El kit para análisis y el método pueden emplear también oligonucleótidos etiquetados para permitir la fácil identificación en los análisis. Ejemplos de etiquetas que pueden emplearse incluyen radio-etiquetas, enzimas, compuestos fluorescentes, estreptavidina, avidina, biotina, restos magnéticos, restos de unión de metales, restos de antígenos o anticuerpos y similares.

El kit puede, opcionalmente, incluir también medios de muestreo de DNA. Los medios de muestreo de DNA son bien conocidos por los expertos en la técnica y pueden incluir, pero sin limitación, sustratos, tales como papeles de filtro, el AmpliCard™ (Universidad de Sheffield, Sheffield, England S10 2JF; Tarlow, JW, *et al.*, *J. of Invest. Dermatol.*, 103:387-389 (1994)) y similares; reactivos de purificación de DNA, tales como kits Nucleon™, tampones de lisis, soluciones de proteinasa y similares; reactivos de PCR, tales como tampones para reacción de 10 veces, polimerasa termoestable, dNTP y similares; y medios de detección de alelos, tales como la enzima de restricción *HinfI*, oligonucleótidos específicos de alelos, cebadores de oligonucleótidos degenerados para PCR anidada de sangre seca.

4.2.3. Farmacogenómica

El conocimiento de los alelos particulares asociados con una sensibilidad a desarrollar una enfermedad o estado particular, solos o en asociación con la información sobre otros defectos genéticos que contribuyen a la enfermedad o estado particular permite la personalización de la prevención o tratamiento de acuerdo con el perfil genético del

individuo, y esto es lo que constituye el objeto de la “farmacogenómica”. Por tanto, la comparación de un perfil de IL-1 individual con el perfil de la población para un trastorno vascular, permite la selección o diseño de fármacos u otros regímenes terapéuticos que se espera que sean seguros y eficaces para un paciente o población de pacientes particular (es decir, un grupo de pacientes que tienen la misma alteración genética).

Además, la capacidad para señalar como diana las poblaciones que se espera muestren el beneficio clínico más alto, basado en el perfil genético, puede facilitar: 1) el reposicionamiento de fármacos ya comercializados; 2) el rescate de candidatos a fármacos cuyo desarrollo clínico ha sido interrumpido como resultado de limitaciones de seguridad o eficacia, que son específicas de un sub-grupo de pacientes; y 3) un desarrollo acelerado y menos costoso para agentes terapéuticos candidatos y un etiquetado de fármacos más óptimo (por ejemplo, ya que medir el efecto de varias dosis de un agente sobre la mutación causante es útil para optimizar la dosis eficaz).

El tratamiento de un individuo con un agente terapéutico particular puede ser monitorizado determinando la proteína (por ejemplo IL-1 α , IL-1 β o IL-1Ra), el mRNA y/o el nivel de transcripción. Dependiendo del nivel detectado, el régimen terapéutico puede mantenerse o ajustarse luego (aumentando o disminuyendo la dosis). En una realización preferida la eficacia de tratar un sujeto con un agente comprende las etapas de: (i) obtener una muestra para la pre-administración de un sujeto antes de la administración del agente; (ii) detectar el nivel o cantidad de una proteína, mRNA o DNA genómico en la muestra para la pre-administración; (iii) obtener una o más muestras del sujeto para la pos-administración; (iv) detectar el nivel de expresión o actividad de la proteína, mRNA o DNA genómico en la muestra administrada; (v) comparar el nivel de expresión o actividad de la proteína, mRNA o DNA genómico en la muestra de pre-administración con la proteína correspondiente, mRNA o DNA genómico en la muestra de la pos-administración, respectivamente; y (vi) alterar consecuentemente la administración del agente al sujeto.

Las células de un sujeto se pueden obtener también antes y después de la administración de un agente terapéutico para detectar el nivel de expresión de genes distintos del gen IL-1 para verificar que el agente terapéutico no aumenta ni disminuye la expresión de genes que podrían ser perjudiciales. Esto puede hacerse, por ejemplo, usando el método del perfilado transcripcional. Por tanto, el mRNA procedente de células expuestas *in vivo* a un agente terapéutico y mRNA procedente del mismo tipo de células que no fueron expuestas al agente terapéutico podrían ser sometidos a reversión, transcritos e hibridados a un chip que contuviera DNA de numerosos genes para comparar con ello la expresión de genes en células tratadas y no tratadas con el agente terapéutico.

4.3.- Agentes terapéuticos para enfermedades y estados asociados con polimorfismos de IL-1

Los agentes terapéuticos para enfermedades o estados asociados con un polimorfismo o haplotipo de IL-1 se refieren a cualquier agente a régimen terapéutico (incluyendo productos farmacéuticos, productos nutricéuticos y medios quirúrgicos) que impide o pospone el desarrollo de los síntomas o su alivio de la enfermedad o estado particular en el sujeto. El agente terapéutico puede ser un polipéptido, un peptidomimético, ácido nucleico u otra molécula orgánica o inorgánica, preferiblemente una “molécula pequeña” que incluye vitaminas, minerales y otros nutrientes. Preferiblemente el agente terapéutico puede modular al menos una actividad de un polipéptido IL-1, por ejemplo, interacción con un receptor, mimetizando o potenciando (agonizando) o inhibiendo (antagonizando) los efectos de un polipéptido de origen natural. Un agonista puede ser una proteína de tipo natural o uno de sus derivados que tiene al menos una bioactividad de tipo natural, por ejemplo, actividad de unión a receptores. Un agonista puede ser también un compuesto que sobre-regula la expresión de un gen o que aumenta al menos una bioactividad de una proteína. Un agonista puede ser también un compuesto que aumenta la interacción de un polipéptido con otra molécula, por ejemplo, un receptor. Un antagonista puede ser un compuesto que inhibe o disminuye la interacción entre una proteína y otra molécula, por ejemplo, un receptor o un agente que bloquea la transducción de señales o el procesamiento pos-traducción (por ejemplo, el inhibidor de la enzima convertidora de IL-1 (ICE). Por consiguiente, un antagonista preferido es un compuesto que inhibe o disminuye la unión a un receptor y con ello bloquea la activación subsiguiente del receptor. Un antagonista puede ser también un compuesto que sub-regula la expresión de un gen o que reduce la cantidad de una proteína presente. El antagonista puede ser una forma negativa dominante de un polipéptido, por ejemplo, una forma de un polipéptido que sea capaz de interactuar con un péptido diana, por ejemplo, un receptor, pero que no promueve la activación del receptor. El antagonista puede ser también un ácido nucleico que codifica una forma negativa dominante de un polipéptido, un ácido nucleico antisentido o una ribozima capaz de interactuar específicamente con un RNA. Incluso otros antagonistas son moléculas que se unen a un polipéptido e inhiben su acción. Dichas moléculas incluyen péptidos, por ejemplo, formas de péptidos dianas que no tienen actividad biológica, y que inhiben la unión a los receptores. Así, dichos péptidos se unirán al sitio activo de una proteína e impedirán la interacción con los péptidos dianas. Incluso otros antagonistas incluyen anticuerpos que específicamente interactúan con un epítipo de una molécula, de tal modo que la unión interfiere con la función biológica del polipéptido. En otra realización preferida, el antagonista es una molécula pequeña, tal como una molécula capaz de inhibir la interacción entre un polipéptido y un receptor diana. Alternativamente, la molécula pequeña puede funcionar como un antagonista interactuando con sitios distintos del sitio de unión al receptor.

Los moduladores de IL-1 (por ejemplo, antagonista de los receptores de IL-1 α , IL-1 β o IL-1) o una proteína codificada por un gen que está en desequilibrio de ligamiento con un gen IL-1 pueden comprender cualquier tipo de compuesto, incluyendo una proteína, péptido, peptidomimético, molécula pequeña o ácido nucleico. Los agonistas preferidos incluyen ácidos nucleicos (por ejemplo, los que codifican una proteína IL-1 o un gen que está sobre- o sub-regulado por una proteína IL-1), proteínas (por ejemplo proteínas IL-1 o una proteína que está sobre- o sub-regulada) o

una molécula pequeña (por ejemplo, que regula la expresión o unión de una proteína IL-1). Los antagonistas preferidos que pueden ser identificados, por ejemplo, usando los análisis descritos en la presente memoria, incluyen ácidos nucleicos (por ejemplo, DNA monocatenario (antisentido) o DNA o PNA bicatenario (tríplice) y ribozimas), proteína (por ejemplo anticuerpos) y moléculas pequeñas que actúan para suprimir o inhibir la transcripción de IL-1 y/o la actividad de la proteína.

4.3.1. Dosis eficaz

La toxicidad y eficacia terapéutica de dichos compuestos puede ser determinada por métodos farmacéuticos estándares en cultivos celulares o animales de experimentación, por ejemplo, para determinar la DL_{50} (la dosis letal para 50% de la población) y la DE_{50} (la dosis terapéuticamente eficaz en 50% de la población). La relación de dosis entre efectos tóxicos y terapéuticos es el índice terapéutico y puede expresarse como DL_{50}/DE_{50} . Se prefieren los compuestos que exhiben índices terapéuticos grandes. Aunque pueden usarse compuestos que exhiban efectos secundarios tóxicos, debe tenerse cuidado en diseñar un sistema de administración que dirija dichos compuestos al sitio de los tejidos afectados con el fin de minimizar el daño potencial a la células no infectadas y, con ello, reducir los efectos secundarios.

Los datos obtenidos de ensayos en cultivos celulares y estudios en animales se pueden usar en formular un intervalo de dosificaciones para uso en seres humanos. La dosificación de dichos compuestos está preferiblemente dentro de las concentraciones circulantes que incluyen la DE_{50} con poca o ninguna toxicidad. La dosificación puede variar dentro de este intervalo dependiendo de la forma de dosificación empleada y la vía de administración utilizada. Para cualquier compuesto usado en el método de la invención, la dosis terapéuticamente eficaz puede ser estimada inicialmente a partir de ensayos en cultivos celulares. En modelos animales se puede formular una dosis para conseguir un intervalo de concentración circulante en plasma que incluya la CI_{50} (es decir, la concentración del compuesto de ensayo que consigue la inhibición semi-máxima de los síntomas) como se determina en cultivos celulares. Dicha información se puede para determinar más precisamente las dosis útiles en seres humanos. Los niveles en plasma se pueden medir, por ejemplo, por cromatografía de líquidos de alta resolución.

4.3.2. Formulación y uso

Las composiciones para uso de acuerdo con la presente descripción pueden ser formuladas de un modo convencional usando uno o más vehículos o excipientes fisiológicamente aceptables. Por tanto, los compuestos y sus sales y solvatos fisiológicamente aceptables pueden ser formulados para administración, por ejemplo, inyección, inhalación o insuflación (a través de la boca o de la nariz) o administración oral, bucal, parenteral o rectal.

Para dicha terapia los compuestos de la invención pueden ser formulados para una variedad de cargas de administración, incluyendo administración sistémica y tópica o localizada. Las técnicas y formulaciones generalmente se pueden encontrar en *Remington's Pharmaceutical Sciences*, Meade Publishing Co., Easton, PA. Para la administración sistémica se prefiere la inyección, incluyendo la intramuscular, intravenosa, intraperitoneal y subcutánea. Para inyección, los compuestos de la invención se pueden formular en soluciones líquidas, preferiblemente en tampones fisiológicamente compatibles, tales como solución de Hank o solución de Ringer. Además, los compuestos pueden ser formulados en forma sólida y re-disueltos o puestos en suspensión inmediatamente antes de su uso. También están incluidas las formas liofilizadas.

Para la administración oral las composiciones pueden tomar la forma de, por ejemplo, comprimidos o capsulas preparadas por medios convencionales con excipientes farmacéuticamente aceptables, tales como agentes aglutinantes (por ejemplo, almidón de maíz pregelatinizado, polivinilpirrolidona o hidroxipropil-metilcelulosa); cargas (por ejemplo, lactosa, celulosa microcristalina o hidrógeno-fosfato de calcio); lubricantes (por ejemplo, estearato de magnesio, talco o sílice); disgregantes (por ejemplo, almidón de patata o almidón-glicolato sódico); o agentes humectantes (por ejemplo, lauril-sulfato sódico). Los comprimidos pueden estar revestidos por métodos bien conocidos en la técnica. Las preparaciones líquidas para la administración oral puede tomar la forma de, por ejemplo, soluciones, jarabes o suspensiones o pueden ser presentadas como un producto seco para reconstitución con agua u otro vehículo adecuado antes de su uso. Dichas preparaciones líquidas pueden ser preparadas por medios convencionales con aditivos farmacéuticamente aceptables, tales como agentes de puesta en suspensión (por ejemplo, jarabe de sorbitol, derivados de celulosa o grasas hidrogenadas comestibles). Los agentes emulsionantes (por ejemplo, lecitina o goma acacia); vehículos no acuosos, (por ejemplo, aceite de almendras, ésteres oleosos, alcohol etílico o aceites vegetales fraccionados); y conservantes (por ejemplo, p-hidroxibenzoatos de metilo o propilo o ácido sórbico). Las preparaciones pueden contener también sales tampones, agentes aromatizantes, colorantes y edulcorantes según sea apropiado.

Las preparaciones para administración oral pueden ser formuladas adecuadamente para dar una liberación controlada del compuesto activo. Para administración bucal las composiciones pueden tener la forma de comprimidos o píldoras formuladas de la manera usual. Para la administración por inhalación, los compuestos para uso de acuerdo con la presente invención se suministran convenientemente en forma de presentación de spray de aerosol, desde envases presurizados o un nebulizador, con el uso de un gas propulsor adecuado, por ejemplo, diclorodifluorometano, triclorofluorometano, diclorotetrafluoroetano, dióxido de carbono u otro gas adecuado. En el caso de un aerosol presurizado la unidad de dosificación puede ser determinada proporcionando una válvula para suministrar una cantidad medida. Las

cápsulas y cartuchos de, por ejemplo, gelatina para uso en un inhalador o insuflador se pueden formular conteniendo una mezcla de polvos y una base de polvos adecuada, tal como lactosa o almidón.

5 Los compuestos pueden ser formulados para administración parenteral mediante inyección, por ejemplo, inyección de bolo o infusión continua. Las formulaciones para inyección pueden ser presentadas en forma de dosis unitaria, por ejemplo, en ampollas o en envases multidosis, con un conservante añadido. Las composiciones pueden adoptar las formas de suspensiones, soluciones o emulsiones en vehículos oleosos o acuosos, y pueden contener agentes de formulación de suspensiones, tales como agentes de puesta en suspensión, estabilizantes y/o dispersantes. Alternativamente, el ingrediente activo puede estar en forma de polvo para reconstitución con un vehículo adecuado, por ejemplo, agua estéril exenta de pirógenos antes de su uso.

Los compuestos se pueden formular también en composiciones rectales, tales como supositorios o enemas de retención, por ejemplo, que contienen bases de supositorios usuales, tales como manteca de coco u otros glicéridos.

15 Además de las formulaciones descritas previamente, los compuestos se pueden formular también como una preparación de liberación retardada. Dichas formulaciones de acción prolongada pueden ser administradas por implantación (por ejemplo subcutáneamente o intramuscularmente) o por inyección intramuscular. Así, por ejemplo, los compuestos pueden ser formulados con materiales polímeros o hidrófobos (por ejemplo, en forma de una emulsión en un aceite aceptable) o resinas de intercambio iónico o como derivados escasamente solubles, por ejemplo, como una sal escasamente soluble. Otros sistemas de administración adecuados incluyen microesferas que ofrecen la posibilidad de un suministro no invasivo local de fármacos en un periodo prolongado de tiempo. Esta tecnología utiliza microesferas de tamaño pre-capilar que pueden ser inyectadas mediante un catéter coronario en cualquier parte seleccionada por ejemplo el corazón u otros órganos sin causar inflamación o isquemia. El agente terapéutico administrado se libera lentamente desde estas microesferas y es absorbido por las células del tejido circundante (por ejemplo, las células endoteliales).

30 La administración sistémica puede ser también por medios transmucosales o transdérmicos. Para la administración transmucosal o transdérmica se usan en la formulación agentes penetrantes apropiados para la barrera que ha de ser atravesada por permeación. Dichos agentes penetrantes se conocen generalmente en la técnica, e incluyen, por ejemplo, para la administración transmucosal sales biliares y derivados del ácido fusídico. Además se pueden usar detergentes para facilitar la permeación. La administración transmucosal puede ser a través de sprays nasales o usando supositorios. Para la administración tópica los oligómeros de la invención se formulan en ungüentos, bálsamos, geles o cremas como es generalmente conocido en la técnica. Para tratar una lesión o inflamación se puede usar localmente una solución de lavado para acelerar el curado.

35 Las composiciones pueden ser presentadas, si se desea, en un envase o dispositivo dispensador que puede contener una o más formas de dosificación unitaria que contenga el ingrediente activo. El envase o dispositivo dispensador puede estar acompañado por las instrucciones para la administración.

40 4.4 Análisis para identificar agentes terapéuticos

Basándose en la identificación de las mutaciones que causan o contribuyen al desarrollo de una enfermedad o trastorno que está asociado con un polimorfismo o haplotipo de IL-1, la invención se caracteriza además por análisis basados en células o exentos de células para identificar agentes terapéuticos. En una realización, una célula que expresa un receptor de IL-1, o un receptor para una proteína que está codificada por un gen que está en desequilibrio de ligamiento con un gen de IL-1, sobre la superficie exterior de su membrana celular se incuba en presencia de un compuesto de ensayo sólo o en presencia de un compuesto de ensayo y otra proteína y se detecta la interacción entre compuesto de ensayo y el receptor o entre la proteína (preferiblemente una proteína señalada como diana) y el receptor, por ejemplo, usando un microfisiómetro (McConnell *et al.*, (1992) *Science*, 257:1906).

55 Una interacción entre el receptor y el compuesto de ensayo o la proteína se detecta por el microfisiómetro como un cambio en la acidez del medio. Este sistema de análisis proporciona por tanto un medio de identificar antagonistas moleculares que, por ejemplo, funcionen por interferencia con la interacción proteína-receptor, así como agonista molecular que, por ejemplo, funcione activando un receptor.

60 Los análisis celulares o exentos de células se pueden usar también para identificar compuestos que modulan la expresión de un gen IL-1 o un gen en desequilibrio de ligamiento con él, modular la traducción de un mRNA, o que modulan la estabilidad de un mRNA o proteína. Por consiguiente, en una realización, se incuba una célula que es capaz de producir una proteína IL-1 o de otro tipo con un compuesto de ensayo y la cantidad de proteína producida en el medio celular se mide y compara con la producida por una célula que no ha estado en contacto con el compuesto de ensayo. La especificidad del compuesto frente a la proteína puede ser confirmada por diversos análisis de control, por ejemplo, midiendo la expresión de uno o más genes de control. En particular, este análisis se puede usar para determinar la eficacia de compuestos antisentido, ribozimas y triplex.

65 Los análisis exentos de células se pueden usar también para identificar compuestos que sean capaces de interactuar con una proteína, para modificar con ello la actividad de la proteína. Dicho compuesto puede modificar, por ejemplo, la estructura de una proteína afectando a su capacidad de unión a un receptor. En una realización preferida, los análisis

exentos de células para identificar dichos compuestos consisten esencialmente en una mezcla de reacción que contienen una proteína y un compuesto de ensayo o una quimioteca de compuestos de ensayo en presencia o ausencia de una pareja de unión. Un compuesto de ensayo puede ser, por ejemplo, un derivado de una pareja de unión, por ejemplo, un péptido diana biológicamente inactivo o una molécula pequeña.

5 Por consiguiente, un análisis de escrutinio ilustrativo de la presente descripción incluye las etapas de poner en contacto una proteína o uno de sus fragmentos funcionales con un compuesto de ensayo o quimioteca de compuestos de ensayo y detectar la formación de complejos. Para los fines de detección, la molécula puede ser etiquetada con un marcador específico y el compuesto de ensayo o la quimioteca de compuestos de ensayos etiquetados con un
10 marcador diferente. La interacción de un compuesto de ensayo con una proteína o uno de sus fragmentos se puede detectar luego determinando el nivel de las dos etiquetas después de una etapa de incubación y una etapa de lavado. La presencia de los dos marcadores después de la etapa de lavado es indicativa de una interacción.

15 Una interacción entre moléculas puede ser también identificada usando un análisis de interacción biomolecular (abreviadamente en lo sucesivo BIA por la expresión inglesa Biomolecular Interaction Analysis), de Pharmacia Biosensor AB, en tiempo real que detecta la resonancia del plasmón de superficie (SPR = *Surface Plasmon Resonance*) un fenómeno óptico. La detección depende de cambios en la concentración másica de macromoléculas en la interface bioespecífica, y no requiere ningún etiquetado de las entidades que interactúan. En una realización, un quimioteca de compuestos de ensayo puede ser inmovilizada sobre una superficie sensora, por ejemplo, que forma una pared que forma
20 una celda de micro-flujo. Una solución que contiene la proteína o uno de sus fragmentos funcionales se hace fluir luego continuamente sobre la superficie sensora. Un cambio en el ángulo de resonancia como es mostrado en el registro de la señal, indica que ha ocurrido una interacción. Esta técnica está descrita, por ejemplo, en *BIA technology Handbook* por Pharmacia.

25 Otro análisis de escrutinio ilustrativo de la presente descripción incluye las etapas de: (a) formar una mezcla de reacción que incluye: (i) una proteína IL-1 o de otro tipo, (ii) un receptor apropiado, y (iii) un compuesto de ensayo; y (b) detectar una interacción de la proteína y el receptor. Un cambio estadísticamente significativo (potenciación o inhibición) en la interacción de la proteína y el receptor en la presencia del compuesto de ensayo, relativo a la interacción en la ausencia del compuesto de ensayo, indica un antagonismo potencial (inhibidor). Los compuestos
30 de este análisis pueden ser puestos en contacto simultáneamente. Alternativamente, una proteína se puede poner en contacto primeramente con un compuesto de ensayo durante una cantidad apropiada de tiempo, tras lo cual se añade el receptor a la mezcla de reacción. La eficacia del compuesto puede determinarse generando curvas de dosis-respuestas a partir de los datos obtenidos usando diversas concentraciones del compuesto de ensayo. Además, se puede realizar un análisis de control para proporcionar una línea base para la comparación.

35 La formación de un complejo entre una proteína y un receptor puede ser detectada por una variedad de técnicas. La modulación de la formación de complejos puede ser cuantificada usando, por ejemplo, proteínas etiquetadas detectablemente, tales como proteínas o receptores radio-etiquetadas, etiquetadas fluorescentemente o enzimáticamente, por inmunoensayo o por detección cromatográfica.

40 Típicamente, será deseable inmovilizar la proteína o el receptor para facilitar la separación de formas complejas de formas no complejadas de una o de ambas de las proteínas, así como acomodar la automatización del ensayo. La unión de la proteína y el receptor se puede conseguir en cualquier recipiente adecuado para contener los reaccionantes. Ejemplos incluyen placas de microtitulación, tubos de ensayo, y tubos de micro-centrífuga. En una realización se puede proporcionar una proteína de fusión que añade un dominio que permite que la proteína sea unida a una matriz. Por
45 ejemplo, las proteínas de fusión de glutation-S-transferasa pueden ser adsorbidas sobre perlas de glutation-Sefarosa (Sigma Chemical, St. Louis, MO) o placas de microtitulación derivatizadas con glutatión, que luego se combinan con el receptor, por ejemplo, un receptor etiquetado con ³⁵S, y el compuesto de ensayo, y la mezcla se incuba en condiciones que conducen a la formación de complejos, por ejemplo en condiciones fisiológicas en cuanto a sal y pH, aunque puede ser deseable, en condiciones ligeramente más estrictas. Tras la incubación, las perlas se lavan para eliminar la etiqueta no unida, y la matriz inmovilizada y radio-etiquetada se determina directamente (por ejemplo, colocando las perlas en un dispositivo destellante), o en el líquido sobrenadante después de que los complejos sean subsiguientemente disociados. Alternativamente, los complejos pueden ser disociados de la matriz, separados por
50 SDS-PAGE, y el nivel de proteína o de receptor encontrado en la fracción de perlas cuantificado a partir del gel usando técnicas electroforéticas estándares, tales como las descritas en los ejemplos anexos. También están disponibles otras técnicas para inmovilizar proteínas sobre matrices para uso en el análisis pertinente. Por ejemplo, la proteína o el receptor pueden ser inmovilizados utilizando la conjugación de biotina y estreptavidina. Se pueden obtener animales transgénicos también para identificar agonistas y antagonistas o para confirmar la seguridad y eficacia de un agente terapéutico candidato. Los animales transgénicos de la invención pueden incluir animales no humanos que contienen una mutación causante de la reestenosis bajo el control de un promotor endógeno apropiado o bajo el control de un promotor heterólogo.

55 Los animales transgénicos pueden ser también animales que contengan un transgén, tal como un gen informador, bajo el control de un promotor apropiado o uno de sus fragmentos. Estos animales son útiles, por ejemplo, para identificar fármacos que modulan la producción de una proteína IL-1, es decir, modulando la expresión del gen. Los métodos para obtener animales transgénicos no humanos son bien conocidos en la técnica. En realizaciones preferidas, la expresión de la mutación causante de la reestenosis está restringida a subconjuntos específicos de células, tejidos o fases del desarrollo que utilizan, por ejemplo, secuencias de acción cis que controlan la expresión en el modelo

deseado. En la presente invención, dicha expresión en mosaico de una proteína puede ser esencial para muchas formas de análisis de linajes y pueden proporcionar adicionalmente medios para determinar los efectos de, por ejemplo, el nivel de expresión que podría alterar extremadamente el desarrollo en pequeños parches de tejido dentro de un embrión por los demás normal. Para este fin, se pueden usar secuencias reguladoras específicas de tejidos y secuencias reguladoras condicionales para controlar la expresión de la mutación en ciertos modelos espaciales. Además, pueden proporcionarse modelos temporales de expresión, por ejemplo, sistemas de recombinación condicionales o secuencias reguladoras transcripcionales procarióticas. Las técnicas genéticas, que permiten que la expresión de una mutación pueda ser regulada mediante manipulación genética específica del sitio *in vivo*, son conocidas por los expertos en la técnica.

Los animales transgénicos descritos incluyen todos dentro de una pluralidad de sus células un transgén de la mutación causante, que altera el fenotipo de la “célula hospedante”. En una realización ilustrativa, puede usarse el sistema de recombinasa *cre/loxP* del bacteriófago P1 (Lakso *et al.*, (1992) *PNAS*, 89:6232-6236; Orban *et al.*, (1992) *PNAS*, 89:6861-6865) o el sistema de recombinasa FLP de *Saccharomyces cerevisiae* (O’Gorman *et al.*, (1991) *Science*, 251:1351-1355; publicación *PCT* WO 92/15694) para generar sistemas de recombinación genética específicos del sitio *in vivo*. La recombinasa *Cre* cataliza la recombinación específica del sitio de una secuencia diana intercalada localizada entre las secuencias *loxP*. Las secuencias *loxP* son secuencias de repetición de nucleótidos de 34 pares de bases a las cuales se une la recombinasa *Cre* y son requeridas para la recombinación genética mediada por la recombinasa *Cre*. La orientación de las secuencias *loxP* determina si la secuencia diana intercalada es escindida o invertida cuando está presente la recombinasa *Cre* (Abremski *et al.*, (1984) *J. Biol. Chem.* 259:1509-1514); catalizando la escisión de la secuencia diana cuando las secuencias *loxP* están orientadas como repeticiones directas y cataliza la inversión de la secuencia diana cuando las secuencias *loxP* están orientadas como repeticiones inversas.

Por consiguiente, la recombinación genética de la secuencia diana es dependiente de la expresión de la recombinasa *Cre*. La expresión de la recombinasa puede ser regulada por elementos promotores que están sometidos a control regulador, por ejemplo, específico de la fase de desarrollo del tejido, inducible o reprimible por agentes añadidos externamente. Este control regulado dará como resultado recombinación genética de la secuencia diana solamente en células en donde la expresión de la recombinasa esté mediada por el elemento promotor. Por tanto, la activación de la expresión del transgén de la mutación causante puede ser regulada mediante el control de la expresión de la recombinasa.

El uso del sistema recombinasa *cre/loxP* para regular la expresión de un transgén causante requiere la construcción de un animal transgénico que contenga los transgenes que codifican tanto la recombinasa *Cre* como la proteína pertinente. Los animales que contienen tanto la recombinasa *Cre* como el transgén de la mutación causante de la reestenosis pueden ser obtenidos a través de la construcción de animales transgénicos “dobles”. Un método conveniente para obtener dichos animales es aparear dos animales transgénicos que contenga cada uno un transgén.

Se pueden proporcionar transgenes condicionales similares usando secuencias de promotores procarióticos que requieren que sean simultáneamente expresadas proteínas procarióticas para facilitar la expresión del transgén. En la patente de EE.UU. N° 4.833.080 se dan promotores ilustrativos y las proteínas procarióticas transactivantes correspondientes.

Además, la expresión de los transgenes condicionales puede ser inducida por métodos similares a la terapia de genes en donde se suministra al tejido un gen que codifica la proteína transactivante, por ejemplo una recombinasa o una proteína procariota, y se hace que se exprese, tal como en un modo específico del tipo de célula. Mediante este método, el transgén podría permanecer silencioso en un adulto hasta ser “conectado” mediante la introducción del transactivador.

En una realización ilustrativa, los “animales transgénicos no humanos” de la invención se obtienen introduciendo transgenes en la línea germinal del animal no humano. Para introducir los transgenes pueden utilizarse células diana embrionarias en diversas etapas de desarrollo. Se utilizan diferentes métodos dependiendo de la etapa de desarrollo de la célula diana embrionaria. La(s) línea(s) específica(s) de cualquier animal no humano utilizada(s) para realizar esta invención se selecciona(n) para conseguir buena salud general, buen rendimiento de embriones, buena visibilidad pronuclear en el embrión y buena aptitud reproductiva. Además, el haplotipo es un factor significativo. Por ejemplo, cuando se quieren obtener ratones transgénicos, se utilizan con frecuencia razas tales como las líneas C57BL/6 o FVB (Jackson Laboratory, Bar Harbor, ME). Las razas preferidas son aquellas con haplotipos H-2b, H-2d o H-2q, tales como C57BL/6 o DBA/1. La(s) línea(s) utilizada(s) para realizar esta invención puede(n) ser ella(s) misma(s) transgénica(s) y/o puede(n) estar desactivada(s) (es decir, obtenida de animales que tienen uno o más genes parcial o completamente suprimidos).

En una realización, la construcción transgénica se introduce en un embrión de una sola etapa. El cigoto es la mejor diana para micro-inyección. En el ratón, el pronúcleo macho alcanza el tamaño de aproximadamente 20 micrómetros de diámetro que permite la inyección reproducible de 1-2 pl de solución de DNA. El uso de cigotos como diana para la transferencia de genes tiene la ventaja principal de que en la mayoría de los casos el DNA inyectado se incorporará al gen hospedante antes de la primera escisión (Brinster *et al.*, (1985) *PNAS*, 82:4438-4442). En consecuencia, todas las células del animal transgénico llevarán el transgén incorporado. Esto se reflejará también en general en la transmisión eficiente del transgén a la descendencia del animal fundador, puesto que el 50% de las células germinales albergará el transgén.

Normalmente, se incuban embriones fertilizados en medios adecuados hasta que aparecen los pronúcleos. Aproximadamente en ese momento, la secuencia nucleotídica que comprende el transgén se introduce en el pronúcleo hembra o macho como se describe más adelante. En algunas especies tales como ratones, se prefiere el pronúcleo macho. Se prefiere mejor, que el material genético exógeno sea añadido al complemento de DNA macho del cigoto antes de que sea procesado por el núcleo del óvulo o el pronúcleo hembra del cigoto. Se sabe que el núcleo del óvulo o pronúcleo hembra libera moléculas que afectan al complemento de DNA macho, quizás sustituyendo las protaminas del DNA macho por histonas, facilitando con ello la combinación de los complementos de DNA hembra y macho para formar el cigoto diploide. Así, se prefiere que el material genético exógeno sea añadido al complemento macho de DNA o cualquier otro complemento de DNA antes de que sea afectado por el pronúcleo hembra. Por ejemplo, el material genético exógeno se añade al pronúcleo macho temprano, tan pronto como sea posible después de la formación del pronúcleo macho, que es cuando los pronúcleos macho y hembra están bien separados y ambos están situados próximos a la membrana celular. Alternativamente, el material genético exógeno podría añadirse al núcleo del espermatozoide después que haya sido inducido para experimentar descondensación. El espermatozoide que contiene el material genético exógeno puede añadirse entonces al óvulo o el espermatozoide descondensado podría añadirse al óvulo añadiéndose después tan pronto como sea posible las construcciones transgénicas.

La introducción de la secuencia nucleotídica transgénica en el embrión puede realizarse por cualquier medio conocido en la técnica, tal como, por ejemplo, micro-inyección, electroporación o lipofección. Después de la introducción de la secuencia nucleotídica transgénica en el embrión, dicho embrión puede ser incubado *in vitro* durante diversos periodos de tiempo o reimplantado en el hospedante sustituto o ambas técnicas. La incubación *in vitro* hasta la madurez está incluida en el alcance de esta invención. Un método usual es incubar los embriones *in vitro* durante aproximadamente 1-7 días, dependiendo de las especies y a continuación reimplantarlos en el hospedante sustituto.

Para los fines de esta invención, un cigoto es esencialmente la formación de una célula diploide que es capaz de convertirse en un organismo completo. Generalmente, el cigoto comprenderá un huevo que contiene un núcleo formado, natural o artificialmente, por la fusión de dos núcleos haploides procedentes de un gameto o gametos. Por tanto, los núcleos de los gametos deben ser naturalmente compatibles, es decir, que den como resultado un cigoto viable capaz de experimentar diferenciación y convertirse en un organismo funcional. En general, se prefiere un cigoto euploide. Si se obtiene un cigoto aneuploide, entonces el número de cromosomas no debe variar en más de uno respecto al número euploide del organismo del que se originó cada gameto.

Además de consideraciones biológicas similares, las físicas también gobiernan la cantidad (por ejemplo, volumen) de material genético exógeno que se puede añadir al núcleo del cigoto o al material genético que forma una parte del núcleo del cigoto. Si no se retira material genético, entonces la cantidad de material genético exógeno que puede añadirse está limitada por la cantidad que será absorbida sin ser físicamente perjudicial. En general, el volumen de material genético exógeno insertado no será superior a aproximadamente 10 picolitros. Los efectos físicos de adición no deben ser tan grandes que destruyan físicamente la viabilidad del cigoto. El límite biológico del número y variedad de secuencias de DNA variará dependiendo del cigoto particular y de las funciones del material genético exógeno y será fácilmente evidente para los expertos en la técnica, debido a que el material genético, incluyendo el material genético exógeno, del cigoto resultante debe ser biológicamente capaz de iniciar y mantener la diferenciación y el desarrollo del cigoto en un organismo funcional.

El número de copias de las construcciones transgénicas que se añade al cigoto es dependiente de la cantidad total de material genético exógeno añadido y será la cantidad que permita que tenga lugar la transformación genética. Teóricamente, solo se requiere una copia; sin embargo, generalmente, se utilizan numerosas copias, por ejemplo, 1.000-20.000 copias de la construcción transgénica, con el fin de garantizar que la copia sea funcional. En cuanto a la presente invención, será con frecuencia una ventaja tener más de una copia funcional de cada una de las secuencias de DNA exógenas insertadas para mejorar la expresión fenotípica de las secuencias de DNA exógeno.

Puede utilizarse cualquier técnica que permita la adición del material genético exógeno en el material genético nucleico siempre que no destructivo para la célula, la membrana nuclear u otras estructuras celulares o genéticas existentes. El material genético exógeno se inserta preferiblemente en el material genético nucleico por micro-inyección. La micro-inyección de células y estructuras celulares es conocida y usada en la técnica.

La reimplantación se consigue utilizando métodos estándares. Generalmente, se anestesia el hospedante sustituto y se insertan los embriones en el oviducto. El número de embriones implantado en un hospedante particular variará según las especies, pero será en general comparable al número de descendientes que las especies producen de forma natural.

Se puede escrutar la descendencia transgénica del hospedante sustituto para detectar la presencia y/o expresión del transgén por cualquier método adecuado. La selección se realiza con frecuencia por análisis de transferencia Southern o transferencia Northern, utilizando una sonda complementaria con al menos una porción del transgén. Se puede emplear el análisis de transferencia Western utilizando un anticuerpo contra la proteína codificada por el transgén como método alternativo o adicional para escrutar la presencia del producto transgénico. Típicamente, el DNA se prepara a partir del tejido de la cola y se analiza por análisis Southern o por PCR para determinar el transgén. Alternativamente, se analizan los tejidos o células que se cree que expresan el transgén a niveles más altos para determinar la presencia y expresión del transgén utilizando análisis Southern o PCR, aunque pueden usarse para este análisis cualesquiera tejidos o tipos de células.

Métodos alternativos o adicionales para evaluar la presencia del transgén incluyen, sin limitación, valoraciones bioquímicas adecuadas, tales como valoraciones enzimáticas y/o inmunológicas, razas histológicas para detectar un marcador particular o actividades enzimáticas, análisis por citometría de flujo y similares. El análisis de la sangre también puede ser útil para detectar la presencia del producto transgénico en la sangre, así como para evaluar el efecto del transgén sobre los niveles de diversos tipos de células sanguíneas y otros constituyentes de la sangre.

La progenie de los animales transgénicos se puede obtener apareando el animal transgénico con una pareja adecuada, o por fertilización *in vitro* de huevos y/o espermatozoides obtenidos del animal transgénico. Cuando se realiza el apareamiento con una pareja, dicha pareja puede ser o no ser transgénica y/o inactivada; si es transgénica, puede contener el mismo o diferentes transgenes o ambos. Alternativamente, la pareja puede ser una línea parental. Si se usa la fertilización *in vitro*, el embrión fertilizado puede implantarse en un hospedante sustituto o incubarse *in vitro*, o ambas cosas. Utilizando cualquier método, se puede evaluar la progenie para determinar la presencia del transgén aplicando los métodos antes descritos u otros métodos apropiados.

Los animales transgénicos producidos de acuerdo con la presente invención incluirán material genético exógeno. Además, en dichas realizaciones la secuencia se unirá a un elemento de control de la transcripción, por ejemplo, un promotor, que permite preferiblemente la expresión del producto transgénico en un tipo específico de célula.

También se puede utilizar la infección retroviral para introducir el transgén en un animal no humano. El embrión no humano en desarrollo se puede cultivar *in vitro* hasta la etapa del blastocisto. Durante este tiempo, los blastómeros pueden ser dianas para la infección retroviral (Jaenich, R. (1976) *PNAS* 73:1260-1264). La infección eficaz de los blastómeros se obtiene por tratamiento enzimático para eliminar la zona pelúcida (*Manipulating the Mouse Embryo*, Hogan eds. (Cold Spring Harbor Laboratory Press, Cold Spring Harbor, 1986). El sistema vector viral utilizado para introducir el transgén es típicamente un retrovirus de replicación defectuosa que lleva el transgén (Jahner *et al.*, (1985) *PNAS*, 82:6927-6931; Van der Putten *et al.*, (1985) *PNAS*, 82:6148-6152). La transfección se obtiene fácil y eficazmente cultivando los blastómeros sobre una monocapa de células productoras de virus (Van der Putten, *supra*; Stewart *et al.*, (1987) *EMBO J.*, 6:383-388). Alternativamente, la infección se puede realizar en una etapa posterior. El virus o las células productoras de virus se pueden inyectar en el blastocelo (Jahner *et al.*, (1982) *Nature*, 298:623-628). La mayoría de los fundadores serán mosaicos para el transgén puesto que la incorporación sólo ocurre en un subconjunto de células que formaban el animal transgénico no humano. Además, el fundador puede contener diversas inserciones retrovirales del transgén en diferentes posiciones en el genoma que generalmente se segregarán en la descendencia. Por otra parte, también es posible introducir los transgenes en la línea germinal por infección retroviral intrauterina del embrión en la mitad de la gestación (Jahner *et al.*, (1982) *supra*).

Un tercer tipo de célula diana para la introducción de los transgenes es la célula madre embrionaria. Las células madre embrionarias se obtienen de embriones antes de la implantación cultivados *in vitro* y fusionados con embriones (Evans *et al.*, (1981) *Nature*, 292:154-156; Bradley *et al.*, (1984) *Nature*, 309:255-258; Gossler *et al.*, (1986) *PNAS*, 83: 9065-9069; y Robertson *et al.*, (1986) *Nature*, 322:445-448). Los transgenes se pueden introducir eficazmente en las células madre por transfección de del DNA o por traducción mediada por un retrovirus. Dichas células madre embrionarias transformadas se pueden combinar a continuación con blastocistos procedentes de un animal no humano. Las células madre colonizan a continuación el embrión y contribuyen a la línea germinal del animal quimérico resultante. Véase una revisión de Jaenisch, R. (1988) en *Science*, 240:1468-1474.

La presente invención está además ilustrada por los siguientes ejemplos que no deben considerarse de ningún modo como limitativos. El contenido de todas las referencias citadas (incluyendo las referencias bibliográficas, las patentes concedidas y las solicitudes de patentes publicadas que se citan a lo largo de esta solicitud) se incorporan expresamente como referencia. La práctica de la presente invención empleará, salvo indicación contraria, métodos convencionales que están incluidos en la técnica. Dichos métodos están completamente explicados en la bibliografía. Véase, por ejemplo, *Molecular Cloning A Laboratory Manual*, (2nd ed., Sambrook, Fritsch and Maniatis, eds., Cold Spring Harbor Laboratory Press: 1989); *DNA Cloning*, Volumes I and II (D. N. Glover ed., 1985); *Oligonucleotide Synthesis* (M. J. Gait ed., 1984); Patente de EE.UU. N° 4.683.195; Patente de EE.UU. N° 4.683.202; y *Nucleic Acid Hybridization* (B. D. Hames & S. J. Higgins eds., 1984).

5. Ejemplos

Ejemplo 1

Genotipificación

Todos los sujetos humanos fueron donantes de sangre sanos, de raza blanca, no relacionados entre sí de Sheffield (n = 112). Los sujetos fueron tipificados en los *loci* indicados en la Tabla 1.

ES 2 356 167 T3

TABLA 2

Marcadores utilizados en el estudio del haplotipo

Marcador	Gen	Referencia
222/223	IL1A	Todd & Naylor, <i>Nucleic Acids Res.</i> 19: 3756 (1991)
gz5/gz6	IL1A	Zuliani, <i>et al.</i> , <i>Am. J. Hum. Genet.</i> 46: 963-69 (1990)
-889	IL1A	McDowell, <i>et al.</i> , <i>Arth. & Rheum.</i> 38: 221-8 (1995)
+3954	IL1B	di Giovine, <i>et al.</i> , <i>Cytokine</i> 7(6): 606 (1995)
-511	IL1B	di Giovine, <i>Hum. Molec. Genet.</i> 1(6): 450 (1992)
gaat.p33330	entre IL1B e IL1RN	Murray, <i>et al.</i> , <i>Coop. Hum. Link. Center</i> , sin publicar
Y31	entre IL1B e IL1RN	Spurr, <i>et al.</i> , <i>Cytogenet. & Cell Genet.</i> 73: 255-73 (1996)
VNTR	IL1RN	Tarlow, <i>et al.</i> , <i>Hum. Genet.</i> 91: 403-4 (1993)

Las secuencias de los cebadores y las etiquetas fluorescentes utilizadas en la amplificación por PCR de los marcadores se muestran en la Tabla 3.

TABLA 3

Secuencia de los cebadores y etiqueta fluorescente para la genotipificación

Marcador	Etiqueta	Secuencia de los cebadores
222/223	HEX	ATGTATAGAATTCCATTCTG (SEQ ID NO. 8) TAAAATCAAGTGTTGATGTAG (SEQ ID NO. 9)
gz5/gz6	FAM	GGGA7TACAGGCGTGAGCCACCGCG (SEQ ID NO. 10) TTAGTATTGCTGGTAGTATTCATAT (SEQ ID NO. 11)
-889	NINGUNA	TGTTCTACCACCTGAACTAGG (SEQ ID NO. 12) TTACATATGAGCCTTCCATG (SEQ ID NO. 13)
+3954	NINGUNA	CTCAGGTGTCCTCGAAGAAATCAAA (SEQ ID NO. 14) GCTTTMGCTGTGAGTCCCG (SEQ ID NO. 15)
-511	NINGUNA	TGGCATTGATCTGGTTCATC (SEQ ID NO. 16) GTTTAGGAATCTTCCCATT (SEQ ID NO. 17)
gaat.p33330	FAM	GAGGCGTGAGAATCTCAAGA (SEQ ID NO. 18) GTGTCCCTCAAGTGGATCTGG (SEQ ID NO. 19)
Y31	HEX	GGGCAACAGAGCAATGTTTCT (SEQ ID NO. 20) CAGTGTGTCAGTGTACTGTT (SEQ ID NO. 21)
VNTR	NINGUNA	CTCAGCAACACTCCTAT (SEQ ID NO. 22) TCCTGGTCTGCAGGTAA (SEQ ID NO. 23)
-1468	NINGUNA	CCT ATT TCC CTC GTG TCT CAA ATA CT (SEQ ID NO: 36) TCA AAT TAG TAT GTG CCA GGT ATC GT
-3737	NINGUNA	CTA AGC CCT CCT TGT TCT AGA CCA CTC TTG GAT CCA GGT TCT CAG ATT

ES 2 356 167 T3

Las condiciones de reacción eran las incluidas en la Tabla 4.

TABLA 4

Condiciones de reacción

Marcador	Condiciones
222/223	KCl 50 mM, Tris-HCl 10 mM pH 9,0, MgCl ₂ 1,5 mM, dNTP 200 mM, 25 ng de cebadores, 50 ng de molde, 0,004% W-1 (Gibco-BRL) 0,2 u Taq, la PCR se realizó en 30 ciclos de 94°C durante 1", 55°C durante 1", 72°C durante 1"
gz5/gz6	como para el marcador 222/223, excepto que se añadió 1 u de Perfect Match (StrataGene)
-889	como para el marcador 222/223, excepto que la PCR se realizó con 1 ciclo a 96°C durante 1", 40 ciclos de 94°C durante 1", 46°C durante 1", 72°C durante 1" y 1 ciclo de 72°C durante 4", los productos se escindieron con NcoI para su análisis
+3954	como para el marcador 222/223, excepto que la PCR se realizó en 35 ciclos con reasociación a 67,5°C, los productos se escindieron con Taq 1 para su análisis
-511	como para el marcador 222/223, excepto que la PCR se realizó en 1 ciclo a 95°C durante 2", 35 ciclos de 95°C durante 1", 53°C durante 1", 74°C durante 1" y 1 ciclo de 74°C durante 4", los productos se escindieron con Aval y Bsu361 para su análisis
gaat.p33330	como para el marcador 222/223
Y31	como para el marcador 222/223
VNTR	como para el marcador 222/223, excepto con MgCl ₂ 1,7 mM, 1 ciclo a 96°C durante 1", 30 ciclos de 94°C durante 1", 60°C durante 1", 70°C durante 1" y 1 ciclo a 70°C durante 2"

Los productos de la PCR 222/223, gz5/gz6, gaat.p33330 y Y31 se examinaron por electroforesis en gel de agarosa y el resto de los productos de la PCR se mezclaron de acuerdo con la intensidad de tinción con bromuro de etidio. 2 g/l de la mezcla se analizaron en un secuenciador automático ABI 373A y se determinaron los tamaños de los alelos utilizando el programa informático *Genescan and Genotyper*. Los alelos se agruparon globalmente utilizando un sencillo programa de ordenador y numerados en orden de tamaño.

Los productos de la PCR -889 se sometieron a digestión con NcoI y los fragmentos resultantes fueron ordenados por tamaños por PAGE al 8%. El alelo 1 produce fragmentos de 83 y 16 pb. El alelo 2 produce un fragmento de 99 pb.

Los productos de la PCR +3954 se sometieron a digestión con la enzima de restricción Taq I. El alelo 1 produce fragmentos de 97, 85 y 12 pb y el alelo 2 produce fragmentos de 182 y 12 pb.

Los productos de la PCR -511 se sometieron a digestión con Aval y Bsu361 y los fragmentos fueron ordenados por tamaños por PAGE al 8%. El alelo 1 produce fragmentos de 190 y 114 pb cuando se sometió a digestión con Aval y un fragmento de 304 pb cuando se sometió a digestión con Bsu361. El alelo 2 produce un fragmento de 304 pb cuando se sometió a digestión con Aval y fragmentos de 190 y 114 pb cuando se sometió a digestión con Bsu361.

Los productos de la PCR VNTR fueron ordenados por tamaños por electroforesis sobre gel de agarosa al 2% a 90V durante 45 minutos. El alelo 1 tiene 4 repeticiones y el producto de la PCR tiene 412 pb, el alelo 2 tiene 2 repeticiones y el producto de la PCR tiene 240 pb, el alelo 3 tiene 3 repeticiones y el producto de la PCR tiene 326 pb, el alelo 4 tiene 4 repeticiones y el producto de la PCR tiene 498 pb, el alelo 5 tiene 6 repeticiones y el producto de la PCR tiene 584 pb.

Las distancias intergénicas se determinaron por estimación basándose en los tamaños de los insertos de los clones PAC correspondientes a partir de un conjunto de clones contiguos que abarca la agrupación de genes de IL-1 (Nicklin, *et al.*, *Genomics* 19:382-4 (1994)). Las distancias intragénicas se determinaron a partir de la secuencia de nucleótidos correspondientes obtenidos de la base de datos GENBANK.

Ejemplo 2

Método para estimar el desequilibrio de ligamiento

Debido a que cuatro de los marcadores estudiados en la presente invención son multialélicos, se realizó un análisis preliminar para determinar que combinaciones alélicas entre pares de *loci* contribuían al mayor desequilibrio, con el fin de que el desequilibrio no fuera enmascarado cuando se agrupan los alelos en sistemas bialélicos. Se utilizó el programa E.H. de Xie y Ott (*Handbook of Human Genetic Linkage*, 1994, John Hopkins University Press, 188-98), incorporado aquí como referencia, para estimar las frecuencias de haplotipos bajo H_0 (sin ligamiento) y H_1 (permitido el ligamiento alélico). Se encontró que la estrategia de agrupamiento de alelos elaborada tenía algunas ventajas sobre los métodos utilizados corrientemente, en que se detectó el desequilibrio entre casi todas las combinaciones de marcadores en parejas examinadas y existía buena correlación entre el desequilibrio y la distancia física.

Más específicamente, se utilizó el programa E.H. de Xie y Ott para determinar las estimaciones de probabilidad máxima de desequilibrio (D_{ij}) entre cada combinación de alelos en parejas, donde $D_{ij} = h_{ij} - p_i q_j$ son las frecuencias para el alelo i en el *locus* 1 y para el alelo j en el *locus* 2 respectivamente, y h_{ij} es la frecuencia del haplotipo ij . El programa calculó los valores de probabilidad máxima para las frecuencias de los haplotipos (y por consiguiente frecuencias de los alelos) bajo H_0 (sin asociación) y las frecuencias de los haplotipos bajo H_1 (permitida la asociación alélica). Para marcadores con más de dos alelos, la estimación por E. H. para las frecuencias de los alelos está mal relacionada con las frecuencias de los alelos estimadas directamente en la población de muestra y por consiguiente no dieron ninguna confianza a las estimaciones de D_{ij} dadas. Por consiguiente no era necesario agrupar alelos de los marcadores multialélicos en un sistema bialélico. El análisis de los marcadores en un formato bialélico tiene las ventajas adicionales de que la anotación $^{\wedge}D_{ij}$, p_j y q_j se pueden simplificar en $^{\wedge}D$, p y q respectivamente, donde p y q se definen como las frecuencias de los alelos más raros en ambos *loci* (de tal forma que sin pérdida de generalidad $p \propto \alpha 0,5$) y $^{\wedge}D$ es el desequilibrio estimado entre dichos alelos.

En un sistema bialélico, la potencia es también mucho más sencilla de determinar utilizando ecuaciones como las detalladas por Hill (Hill, *Heredity*, 33: 229-39 (1974)). Además, el signo de $^{\wedge}D$ se hace más informativo, de tal forma que $^{\wedge}D > 0$ cuando se asocian los alelos más raros en cada uno de los dos *loci* y $^{\wedge}D < 0$ cuando el alelo raro en un *locus* se asocia con el alelo común en el otro *locus*.

Debido a que el método de agrupamiento de alelos afectaba claramente a la potencia para detectar el desequilibrio (Zouros, *et al.*, *Gent.* 85: 543-50 (1977); Weir, *et al.*, *Gent.* 88: 633-42 (1976)), se realizó un análisis preliminar para garantizar que el agrupamiento no enmascaraba el desequilibrio entre subconjuntos de alelos. En este análisis, se calculó $\delta_{ij} = (O'_{ij} - E_{ij})/E_{ij}$ para cada haplotipo, donde E_{ij} es el número esperado de haplotipos ij suponiendo el equilibrio ($E_{ij} = 2n p_i q_j$, donde n = número de individuos en el estudio) y O'_{ij} es una estimación básica para el recuento de haplotipos observados, determinada como sigue. Todos los genotipos que podrían ser resueltos sin ambigüedad eran un haplotipo contado. Cada heterocigoto doble ($i_1 i_2 / j_1 j_2$) podría ser resuelto en dos posibles conjuntos de haplotipos, $[i_1 j_1, i_2 j_2]$ o $[i_1 j_2, i_2 j_1]$. Utilizando las frecuencias de haplotipos estimadas a partir del recuento de haplotipos sin ambigüedad, se calculó la probabilidad de cada conjunto y se usó como recuento "parcial". De este modo los genotipos ambiguos eran también haplotipos contados y los recuentos totales (ambiguos más no ambiguos) constituían las O'_{ij} utilizadas en δ_{ij} . Una vez estabilizado, se utilizaron la magnitud y signo de los δ_{ij} para determinar que combinación alélica mostraba mayor desviación de la hipótesis nula de no asociación. Esta información se utilizó para agrupar alelos en los *loci* multialélicos en sistemas bialélicos para poder utilizar eficazmente el programa E.H.

Con el fin de comparar el grado de desequilibrio entre diferentes combinaciones de *loci* en parejas, se calculó una medida independiente de la frecuencia del desequilibrio $\sim D$, (la proporción de desequilibrio máximo posible en la dirección dada), donde $\sim D = ^{\wedge}D / |D_{\max}|$ (Thompson, *et al.*, *Am. J. Hum. Gent.* 42: 113-24 (1988)). La relación entre p y q es tal que $p \propto q \propto 0,5$, por consiguiente puede escribirse que $-pq \propto D \propto p(1-q)$, tal que cuando $^{\wedge}D < 0$, $D_{\max} = -pq$ y cuando $^{\wedge}D > 0$, $D_{\max} = p(1-q)$. La salida del programa E.H. incluía probabilidades logarítmicas para los valores máximos del parámetro de probabilidad bajo H_0 y H_1 , y puesto que $-2 \ln(L_0/L_1) \sim X^2_1$, donde L_0 y L_1 son las probabilidades bajo H_0 y H_1 , se podrían entonces determinar los valores de p para cada ensayo.

Las varianzas asintóticas para $^{\wedge}D$, bajo H_0 : $D = 0$ y H_1 , se calcularon utilizando la fórmula definida por Hill (*Heredity*, 33: 229-39 (1974)) para datos genotípicos. Utilizándolas, se pudo calcular la potencia para cada comparación en parejas.

Se identificaron haplotipos comunes que contenían todos 8 *loci* a partir del análisis preliminar de δ_{ij} antes descrito y respaldados por la magnitud y signo de los desequilibrios una vez que los alelos en los *loci* multialélicos fueron agrupados. Para estos *loci*, fue identificado en el haplotipo el alelo en el grupo que contribuía en mayor medida al desequilibrio. Para estimar las frecuencias de haplotipos en la población, se determinaron las tasas de transporte de al menos una copia de los alelos pertinentes en la población. Estos no representan verdaderos haplotipos puesto que la fase es desconocida. Se utilizaron técnicas de simulación de Monte Carlo para analizar la desviación significativa de una distribución nula simulada para estos transportes combinados suponiendo que no hay asociación.

ES 2 356 167 T3

Ejemplo 3

Estimación del desequilibrio del ligamiento en la agrupación de genes de IL-1

5 Se han identificado diversos marcadores bialélicos y multialélicos en los genes de IL-1 y alrededor de ellos. Sin embargo, no han sido identificados hasta ahora el grado de desequilibrio del ligamiento entre los marcadores y la prevalencia de haplotipos multimarcadores en la población general.

10 La Figura 1 muestra las posiciones relativas de los 8 *loci* marcadores utilizados en este estudio. Muestras de DNA de 212 voluntarios sanos no relacionados fueron genotipificadas para cada uno de estos marcadores, y las estimaciones resultantes de frecuencias de alelos se muestran en la Tabla 5.

TABLA 5

Frecuencias estimadas de alelos marcadores

20	222/223	frec.	gz5/gz6	frec.	-889	frec.	+3953	frec.
	1 (126 pb)	0,005	1 (79 pb)	0,003	1(NcoI)	0,714	1(2 TagI)	0,812
	2 (128 pb)	0,018	2 (82 pb)	0,005	2	0,286	2	0,188
	3 (130 pb)	0,378	3 (88 pb)	0,676				
25	4 (132 pb)	0,299	4 (91 pb)	0,316				
	5 (134 pb)	0,016						
	6 (136 pb)	0,208						
	7 (138 pb)	0,055						
	8 (140 pb)	0,003						
30	9 (142 pb)	0,010						
	10 (144 pb)	0,008						
	*total	384		392		398		398
	-511	frec.	gaat p33330	frec.	Y31	frec.	VNTR	frec.
35	1	0,618	1 (189 pb)	0,658	1 (148 pb)	0,092	1	0,744
	2 (Bsu361)	0,382	2 (193 pb)	0,002	2 (158 pb)	0,008	2	0,256
			3 (197 pb)	0,255	3 (160 pb)	0,454		
			4 (201 pb)	0,084	4 (162 pb)	0,062		
40					5 (164 pb)	0,003		
					6 (166 pb)	0,122		
					7 (168 pb)	0,035		
					8 (170 pb)	0,030		
					9 (172 pb)	0,095		
45					10 (174 pb)	0,087		
					11 (176 pb)	0,003		
					12 (178 pb)	0,011		
		398		404		370		398

50 *número de cromosomas analizado

Nota – los nombres de los alelos (y tamaños) están en negrita.

55 Para determinar el desequilibrio del ligamiento entre combinaciones de *loci* en parejas, se utilizó el programa de ordenador de Xie y Ott. Se encontró que este programa era el más eficaz cuando se utilizaba con sistemas bialélicos, por consiguiente se agruparon los alelos en los *loci* multialélicos del modo más apropiado para cada comparación en parejas, de forma que no se enmascarara el desequilibrio entre subconjuntos de alelos.

60 En la Tabla 6, los desequilibrios entre parejas de *loci* se expresan como $\sim D$, la relación entre ΔD y su valor máximo D_{\max} , y se muestran junto con la distancia física aproximada entre los *loci* en kilopares de bases.

65

ES 2 356 167 T3

TABLA 6

Desequilibrio ($\sim D = \wedge D/D_{\max}$) y distancia física entre marcadores

	222/223	gz5/gz6	-889	+3953
222/223	-	+0,872	+0,829	+0,710
gz5/gz6	2,5	-	-0,889	-0,695
-889	7	4,5	-	+0,804
+3953	55	55	50	-
-511	60	60	55	4,5
gaat.p33330	260	260	255	205
Y31	310	310	305	255
VNTR	380	380	375	325
	-511	gaat.p33330	Y31	VNTR
222/223	+0,535	+0,433	+0,364	-0,499
gz5/gz6	+0,540	+0,517	-0,503	+0,286
-889	-0,264	+0,337	+0,318	-0,207
+3954	-0,617	+0,409	-0,475	0,439
-511	-	+0,691	-0,456	+0,448
gaat.p33330	200	-	+0,639	+0,442
Y31	250	50	-	-0,765
VNTR	320	120	70	-

Nota – los valores de desequilibrio se muestran en la parte superior derecha, las distancias físicas en Kb se muestran en la parte inferior izquierda. Las distancias intergénicas se dan con un error de 5 Kb.

La Tabla 7 muestra la potencia para detectar el 50% de D_{\max} para cada combinación de *loci* y los valores de *p* para cada *D* correspondiente.

TABLA 7

*Potencia para detectar el 50% de D_{\max} y valores de *p* de $-2\ln(L_0/L_1)$*

	222/223	gz5/gz6	-889	+3953
222/223	-	-100(+)	-100(+)	98(+)
gz5/gz6	$<1 \times 10^{-10}$	-	87(-)	60(-)
-889	$<1 \times 10^{-10}$	$<3 \times 10^{-8}$	-	-100(+)
+3954	-1×10^{-7}	$*9 \times 10^{-3}$	$<1 \times 10^{-10}$	-
-511	-9×10^{-10}	-4×10^{-10}	$*9,4 \times 10^{-2}$	$*2,6 \times 10^{-2}$
gaat.p33330	-9×10^{-8}	-2×10^{-9}	$*1,7 \times 10^{-2}$	-5×10^{-4}
Y31	-1×10^{-4}	-4×10^{-4}	-6×10^{-4}	-1×10^{-7}
VNTR	-1×10^{-3}	-1×10^{-3}	$*3 \times 10^{-1}$	$*1,2 \times 10^{-1}$
	-511	gaat.p33330	Y31	VNTR
222/223	$\sim 100(+)$	$\sim 100(+)$	$\sim 100(+)$	93(-)
gz5/gz6	$\sim 100(+)$	$\sim 100(+)$	98(-)	$\sim 100(+)$
-889	96(-)	89(+)	$\sim 100(+)$	78(-)
+3954	79(-)	97(+)	$\sim 100(+)$	52(-)
-511	-	$\sim 100(+)$	$\sim 100(-)$	$\sim 100(+)$
gaat.p33330	$<1 \times 10^{-10}$	-	49(+)	$\sim 100(+)$
Y31	$\sim 2 \times 10^{-4}$	$*\sim 7 \times 10^{-3}$	-	89(-)
VNTR	$\sim 8 \times 10^{-6}$	$\sim 1 \times 10^{-9}$	2×10^{-7}	-

Nota – La potencia se muestra en la parte superior derecha con el signo de desequilibrio entre paréntesis; los valores de *p* puntuales se muestran (sin corregir) en el fondo izquierda. Para un nivel de significancia global de $p = 0,05$, el nivel de significancia puntual es 0,0018 para 28 comparaciones.

* No significativo a un umbral de $p = 0,0018$

Se detectó el desequilibrio del ligamiento significativo ($p_{\text{corr}} < 0,05$) entre la mayor parte de las combinaciones de *loci*, sólo con algunas excepciones. Estas incluyen las comparaciones entre el marcador VNTR y los marcadores bialélicos más distantes, +3954 y -889, en los que el desequilibrio se produce en dirección negativa y en consecuencia se reduce la potencia (Tabla 7). La correlación entre el desequilibrio $\sim D$ y la distancia física era $r = -0,752$ ($p < 0,0001$, unilateral) (Figura 2).

ES 2 356 167 T3

Con el fin de comparar diferentes métodos de agrupamiento para los marcadores multialélicos, se calculó $\sim D$ para todas las comparaciones que implicaban 222/223 utilizando dos estrategias de agrupamiento adicionales. La primera consistía en una propuesta de “alelo común frente al resto” y la segunda en un agrupamiento basado en el tamaño del alelo, utilizando la distribución bimodal de la frecuencia del alelo frente al tamaño que se observaba para todos los marcadores multialélicos examinados. Los resultados de este análisis se muestran en la Tabla 8, donde se comparan los valores de $\sim D$ para los tres métodos de agrupamiento.

TABLA 8

Valores de $\sim D$ para los tres métodos agrupamiento de alelos en los loci marcadores multialélicos

	Δ_{ij}	común frente al resto	tamaño del alelo
gz5/gz6	0,87	0,79	0,77
-899	0,83	0,81	0,98
+3954	0,71	*0,74	0,77
-511	0,54	*0,15	0,61
gaat.p33330	0,43	*0,03	0,53
Y31	0,36	*0,12	0,16
VNTR	0,5	0,48	*0,04

Nota – Valores dados para el desequilibrio entre 222/223 y los otros marcadores indicados.
* indica no significativo a un nivel $p = 0,05$, incluso antes de la corrección para ensayos múltiples.

Se puede observar que el desequilibrio no se detecta en varios casos utilizando estas otras estrategias de agrupamiento, principalmente 222/223 con -511 y gaat.p33330 en la propuesta común frente al resto, 222/223 con Y31 tanto en la propuesta común frente al resto como en la del tamaño de los alelos y 222/223 con VNTR en la propuesta del tamaño de los alelos.

El examen de qué alelos de los *loci* multialélicos contribuyeron en mayor grado al desequilibrio, a partir de la determinación de δ_{ij} reveló la existencia de 2 haplotipos que contenían alelos de los 8 *loci*. Estos se confirmaron por examen de las frecuencias de haplotipos y valores del desequilibrio obtenidos después del agrupamiento. El primer haplotipo: alelos 44112332 (expresados en orden cromosómico, véase la Fig. 1) es el más común (transporte de 34/198) y está presente 7 veces más frecuentemente que lo esperado (esperado = 4,5/198) ($p < 0,000001$). El segundo haplotipo: alelos 33221461 (transporte de 2/206) estaba presente 4 veces más frecuentemente que lo esperado (esperado 0,5/206), pero esto no era estadísticamente significativo ($p \sim 0,106$). Sin embargo, el examen de un tamaño de muestra mayor podía ayudar a aumentar el significado estadístico de este hallazgo.

Los datos presentados indican un grado significativo de desequilibrio del ligamiento en un tramo de aproximadamente 400 Kb del cromosoma 2q13. El desequilibrio era grande tanto para los tres marcadores en el gen de IL-1 α , lo que podría esperarse, pero también entre alguno de los marcadores más separados (-899/+3954; $\sim D = +0,804$, distancia física = 50 Kb) (Tabla 6). Sin embargo, $\sim D$ disminuyó considerablemente entre los extremos más alejados del agrupamiento. Dentro del gen de IL-1 β , se obtuvo un valor moderado de $\sim D$ (+3954/-511; $\sim D = -0,617$), aunque este valor no era significativo cuando se corrigió para comparaciones múltiples, reflejando probablemente la reducción en potencia cuando el desequilibrio se produce en la dirección negativa (Thompson, *et al.*, *Am. J. Hum. Gent.* 42: 113-24 (1988)).

En general, existe una buena correlación entre la distancia física y el desequilibrio del ligamiento (Figura 2); $r = -0,752$. La fiabilidad de r propiamente dicha depende parcialmente de la fiabilidad de las estimaciones tanto de la distancia física como de $\sim D$. En distancias cortas, las distancias físicas son exactas puesto que se determinan a partir de secuencias de DNA conocidas, mientras que las estimaciones de distancias largas son menos precisas. La potencia puede ser considerada provisionalmente como un indicador de la fiabilidad de $\sim D$, puesto que si la potencia es baja esto indica que el tamaño de la muestra era demasiado pequeño y con tamaños de muestras pequeños las estimaciones de $\sim D$ pueden no ser fiables.

El éxito de la estrategia elaborada de agrupamiento se indica en la Tabla 8, que presenta varios casos en los que el desequilibrio entre los *loci* particulares es aparentemente bajo o no detectado cuando se emplean otros métodos de agrupación usualmente utilizados. Las desventajas de la estrategia de agrupamiento utilizada en la presente invención son que dicha estrategia es bastante laboriosa puesto que la información utilizada para el agrupamiento se basaba en una estimación aproximada de las frecuencias de haplotipos “observadas” (véase el Ejemplo 2). Para los marcadores más polimórficos la mayor heterocigosis significa que la estimación de δ_{ij} era menos precisa puesto que había una mayor proporción de haplotipos ambiguos. A pesar de este inconveniente, se procuró tener en cuenta tanto el signo como la magnitud de δ_{ij} y las frecuencias de los alelos relacionados.

ES 2 356 167 T3

El método pudo simplificarse, en un estudio suficientemente amplio, sólo considerando los haplotipos no ambiguos cuando se determina el agrupamiento. La determinación de δ_{ij} utiliza todo el conocimiento anterior para el agrupamiento de los marcadores multialélicos y esta puede ser la razón por la que se detectó el desequilibrio entre casi todas las combinaciones de marcadores en parejas.

5

Los dos haplotipos que contienen los 8 marcadores, así como otros haplotipos más cortos, son de particular interés puesto que es probable que las combinaciones particulares de alelos de los genes de IL-1 puedan actuar conjuntamente para determinar un fenotipo inflamatorio global. La comprensión de cuales marcadores están en un desequilibrio de ligamiento fuerte no sólo permite un diseño más racional de los estudios genéticos sino que también puede proporcionar claves para comprender el mecanismo de la enfermedad. Por consiguiente, además de los alelos identificados en la presente memoria, el haplotipo de IL-1 (44112332) puede contener los siguientes alelos:

10

15
alelo 2 del marcador 1731 del gen IL1RN (A en la posición 1731);
alelo 2 del marcador 1812 del gen IL1RN (A en la posición 1812);
alelo 2 del marcador 1868 del gen IL1RN (G en la posición 1868);
alelo 2 del marcador 1887 del gen IL1RN (C en la posición 1887);
alelo 2 del marcador 8006 del gen IL1RN (contiene un sitio HpaI o MspI)
20 alelo 2 del marcador 8061 del gen IL1RN (carece de un sitio MwoI)
alelo 2 del marcador 9589 del gen IL1RN (contiene un sitio SspI)

20

Además, para amplificar estos alelos se pueden utilizar los siguientes cebadores de la PCR:

25

TTACGCAGATAAGAACCAGTTTGG (SEQ ID NO. 24)
TTTCCTGGACGCTTGCTCACCA (SEQ ID NO. 25)
(utilizados para 1731, 1812, 1868 y 1887)

25

TTCTATCTGAGGAACAACCAACTAGTAGC (SEQ ID NO. 26)
CACCAGACTTGACACAGGACAGGCACATC (SEQ ID NO. 27)
(utilizados para 8006)

30

CGACCCTCTGGGAGAAAATCCAGCAAG (SEQ ID NO. 28)
(utilizado con SEQ ID NO. 20 para 8006)

30

ACACAGGAAGGTGCCAAGCA (SEQ ID NO. 29)
35 TGCAGACAGACGGGCAAAGT (SEQ ID NO. 30)
(utilizados para 8006 y 9589)

35

TTGTGGGGACCAGGGGAGAT (SEQ ID NO. 31) y
40 AGCCTGGCACTCTGCTGAAT (SEQ ID NO. 32)
(utilizados para 9589).

40

Ejemplo 4

45

El haplotipo de IL-1 (441123-32) está asociado a la nefropatía diabética

50

Se investigó la presencia de los dos haplotipos descritos en la presente memoria en poblaciones sanas y enfermas para determinar si los haplotipos estaban asociados a una enfermedad inflamatoria. En el ejemplo 3 se compararon 81 pacientes con diabetes mellitus no dependiente de insulina (NIDDM) y con nefropatía con 198 sujetos sanos étnicamente similares y 147 pacientes con NIDDM sin nefropatía. Se realizó la genotipificación como en el ejemplo 1.

55

El haplotipo de IL-1 (44112332) era portado por 24 de 79 de los pacientes con NIDDM y nefropatía y 25 de 141 de los pacientes con NIDDM sin nefropatía. Sin embargo, no se encontró el segundo haplotipo (3322146 1) en los pacientes con nefropatía (0/8 1). El haplotipo de IL-1 (44112332) estaba significativamente sobre-representado en el grupo de pacientes en comparación con el grupo de control sano (24/79 frente a 34/198; $p = 0,015$) y en el grupo con NIDDM sin nefropatía (24/79 frente a 25/141; $p = 0,03$).

60

Ejemplo 5

Un haplotipo de IL-1 está asociado a una enfermedad inflamatoria

65

Es un ejemplo profético o predictor. Se examinaron otras enfermedades como en el Ejemplo 4. Se encontró que el haplotipo de IL-1 (44112332) está asociado a la arteriopatía coronaria, osteoporosis, nefropatía en diabetes mellitus, alopecia areata, enfermedad de Graves, lupus eritematoso diseminado, esclerosis del liquen y colitis ulcerosa.

ES 2 356 167 T3

Igualmente, el haplotipo de IL-1 (33221461) está asociado a la periodontopatía, artritis crónica juvenil, psoriasis, diabetes dependiente de la insulina y retinopatía diabética.

5 Ejemplo 6

Nuevos marcadores están ligados a un haplotipo de IL-1

10 Es un ejemplo predictor. Se identificaron más marcadores por análisis de las secuencias y enzimas de restricción de la región 2q13-14. Se identificó que estos nuevos marcadores pertenecen a un haplotipo de IL-1 de la manera descrita en los Ejemplos 2 y 3.

Ejemplo 7

15 *El haplotipo de IL-1 (44112332) se utiliza para predecir la sensibilidad a enfermedades*

20 Es un ejemplo predictor. Un paciente con un historial familiar de colitis ulcerosa se genotipificó para determinar la presencia del haplotipo de IL-1 (44112332). La genotipificación se realizó como en el Ejemplo 1 y se detectó que el paciente llevaba uno o más alelos del haplotipo. El paciente se trató por consiguiente con antagonistas de IL-1 para evitar la enfermedad.

25 Un segundo paciente con un historial familiar de arteriopatía coronaria se genotipificó para detectar el agrupamiento de genes IL-1. Se encontró que el paciente llevaba uno o más alelos del haplotipo de IL-1 (44112332) y que era homocigótico para el alelo 2 de VNTR. Por tanto, era 5,4 veces más probable que el paciente desarrollara la arteriopatía coronaria que la población general y se trató exhaustivamente para evitar la enfermedad.

Ejemplo 8

30 *Los haplotipos adicionales son estadísticamente significativos*

35 Es un ejemplo predictor se tipificaron, como en el Ejemplo 1, 400 cromosomas más y se valoró el desequilibrio de ligamiento como en el Ejemplo 2. Se encontró que el haplotipo de IL-1 (33221461) estaba presente aproximadamente 4 veces más frecuentemente que lo que se esperaba ($p < 0,05$).

De manera similar, se determinó que estaban presentes los siguientes marcadores en el haplotipo de IL- 1 (44112332) ($p < 0,05$).

40 Alelo 2 del marcador 1731 del gen IL1RN
alelo 2 del marcador 1812 del gen IL1RN
alelo 2 del marcador 1868 del gen IL1RN
alelo 2 del marcador 1887 del gen IL1RN
45 alelo 2 del marcador 8006 del gen IL1RN
alelo 2 del marcador 8061 del gen IL1RN
alelo 2 del marcador 9589 del gen IL1RN

50 Ejemplo 9

Identificación y caracterización de los polimorfismos de un solo nucleótido en el gen IL-1B humano asociado con la expresión alterada de IL-1 β y estados inflamatorios

55 Se estudiaron el gen IL-1B y otros miembros del agrupamiento de genes IL-1. En particular, se investigó el IL-1B debido a que diversas enfermedades están asociadas a los SNP en su región en dirección 5'. Se identificaron diversos SNP nuevos, como se muestra en las Tablas 9^a-E. Se utilizaron valoraciones con genes indicadores para analizar las combinaciones alélicas de los SNP en la región promotora de IL-1B, algunas de las cuales representan haplotipos frecuentes en la población. Una construcción control tiene el tipo natural (designado alelo 1) en todos los alelos que contienen SNP. La actividad transcripcional diferencial de alelos individuales en las valoraciones indicadoras se estudió más a fondo analizando la unión de proteínas nucleares al DNA alélico en valoraciones del desplazamiento de la electro-movilidad.

65

ES 2 356 167 T3

Cartografiado de los SNP de alta densidad

Se seleccionaron veinticinco voluntarios sanos de diferente origen étnico, auto-descrito. Cada sujeto proporcionó un consentimiento escrito informado y el protocolo fue aprobado por el Comité de ética de la Universidad de Sheffield (U.K.). A partir de muestras de sangre se preparó DNA genómico. El DNA se envió a Genome Therapeutics Corporation (Waltham, MA, USA) para su secuenciación y cartografiado de los SNP de alta densidad. La secuencia genómica públicamente disponible en la región de IL-1 se utilizó para establecer la secuencia de referencia de esta región para identificación de los SNP. Además, se escrutó una genoteca del cromosomas artificiales bacterianos (abreviadamente en lo sucesivo BAC por la expresión inglesa *Bacterial Artificial Chromosomes*) comercialmente disponible para identificar clones de BAC que abarcan la región de IL-1. Se utilizaron BAC positivos para generar un pequeña genoteca de insertos de plásmidos con insertos de 3 a 3,5 kb. Se estudiaron cuatro genes (IL-1A, IL-1B, IL-RN e IL-1F5) y se identificaron los SNP para cada gen en las regiones del exón y 5 kb del extremo 5' del gen. La detección de los SNP se realizó automáticamente utilizando el programa PolyPhred en un instrumento ABI 3700, seguido de inspección visual de los SNP recogidos en las Tablas 9A-E.

TABLA 9A

Descubrimiento de los SNP de IL-1 α

Nº SNP	Con relación a +1	Localización	Tipo	Secuencia de los SNP
1	-3885	Promotor	C/T	GCATGAGCCA\CGGCACCCAGCCACT (SEQ ID NO:)
2	-3764	Promotor	C/T	TGAACTAGAA\CTCAAGAAATTGA (SEQ ID NO:)
3	-3726	Promotor	T/C	CACACTCTCA\TATGAATTCTCCAT (SEQ ID NO:)
4	-3608	Promotor	G/A	CATATTCTGG\AACCTTCAATAAA (SEQ ID NO:)
5	-3115	Promotor	C/T	AAAAGTTATG\TTTTCTCTTCATTCA (SEQ ID NO:)
6	-2756	Promotor	C/T	TCTTTATAAG\CCATCACTTGGTG (SEQ ID NO:)
9	-2269	Promotor	G/A	CGAGAGGTGG\GTGCCCTGAAGCCACC (SEQ ID NO:)
10	-2118	Promotor	C/T	TGTTACAGT\CCCAGAAAAGCGGGC (SEQ ID NO:)
11	-1506	Promotor	C/T	TAAAGAGGAA\CCAAGGTAAGCAGAA (SEQ ID NO:)
12	-1202	Promotor	C/T	ACACAAGCTG\CTTTCCTCCCAGATC (SEQ ID NO:)
13	-889	Promotor	C/T	CCAGGCAACA\CCATTGAAGGCTC (SEQ ID NO:)
14	+145	Promotor	C/T	AAAGCTACAG\CCTCTCCTTTCTTT (SEQ ID NO:)
15	+251	Promotor	T/A	CTGATTCTTT\TTACTGAGGGACG (SEQ ID NO:)
16	+263	Promotor	G/A	ACTGAGGGAC\GGCAGAAGTATTTT (SEQ ID NO:)
17	+1864	Promotor	A/C	CTTGAATCTT\AAATACTTTTGTT (SEQ ID NO:)
18	+1892	Promotor	G/A	CTCACTAGAG\GTCCAGAGACCT (SEQ ID NO:)
19	+2822	Promotor	G/A	AAGAAGAGAC\GGTTGAGTTT (SEQ ID NO:)
24	+4718	Promotor	A/C	TCTGGATTGG\AATATTCCTA (SEQ ID NO:)
25	+4728	Promotor	T/C	ATTCCATAATA\TCCCCTCCAG (SEQ ID NO:)
26	+4845	Intrón 1	G/T	GCCTAGGTCA\GCACCTTTTAG (SEQ ID NO:)
30	+6627	Intrón 1	G/A	GCCCCACCT\GCCCCACCCA (SEQ ID NO:)
31	+8811	Intrón 1	A/C	CTTTTTCTA\ACATCTTGTTCTCTA (SEQ ID NO:)
32	+9460	Intrón 3	G/A	TTTGCCTTCT\GCTTTAAGTT (SEQ ID NO:)
33	+9613	Intrón 3	C/T	AAACTTCT\CGAAGCCGAGC (SEQ ID NO:)
34	+9749	Exón 4	C/G	CTGAGTGTGA\CCAGGCATCCTC (SEQ ID NO:)

ES 2 356 167 T3

TABLA 9B

Descubrimiento de los SNP de IL-1 β

Nº SNP	Con relación a +1	Localización	Tipo	Secuencia de los SNP
1	-5164	Promotor	T/C	GCCTGGGTCC/CAGACTTGACAAA (SEQ ID NO:)
2	-3959	Promotor	T/C	AGAAAAGACA/TAGAGTAGGA (SEQ ID NO:)
3	-3893	Promotor	G/A	TCCAAAGGAA/GGACAAGGTC (SEQ ID NO:)
4	-3737	Promotor	C/T	GGGAGGAGAATGGAATGT/CCCTTGGACTCT (SEQ ID NO:)
5	-3686	Promotor	A/G	GAAGAAGCCC/ATTGGAGATGATG (SEQ ID NO:)
6	-3254	Promotor	T/C	GATAACTGGC/TGCGAAGCCCATGAT (SEQ ID NO:)
7	-2905	Promotor	A/G	GGAAGACAGG/ATCTGATACATAC (SEQ ID NO:)
8	-2394	Promotor	G/A	CCTGTCACTG/GCTTTGATCCTCCTT (SEQ ID NO:)
9	-2378	Promotor	G/A	ATCCTCCTTC/GTTCAGCTTGTAAATC (SEQ ID NO:)
10	-1468	Promotor	G/C	CACTCCCTTG/GATAATGCAGAGCGAG (SEQ ID NO:)
14	-511	Promotor	C/T	AGAGAGCTCC/TGAGGCAGAGAAC (SEQ ID NO:)
15	-31	Promotor	T/C	TTTTGAAAGC/CATAAAAACAGCGAGGGAG (SEQ ID NO:)
17	+45	Exón 1 5' UTR	T/C	GCTCTGGGAT/TCTCTTCAGCCAATCTTCAT (SEQ ID NO:)
19	+746	Intrón 2	C/T	ACTACTTTCC/CATTACAAGTCCCTCCAG (SEQ ID NO:)
24	+3876	Intrón 4	G/A	AAATTTGCC/GCCTCGCCTCACGAG (SEQ ID NO:)
25	+3954	Exón 5	C/T	CTATCTTCTT/CGACACATGGGATAACG (SEQ ID NO:)
30	+5548	Intrón 6	G/A	CCTTCTCCCC/GCCCCCATCCCTAGG (SEQ ID NO:)
35	+6506	Exón 7 3' UTR	C/A	ATAGCCTGGA/CTTTCCTGTTGTCT (SEQ ID NO:)
37	+6911	Exón 7 3' UTR	G/C	CTTTAATTAA/GACTGAAAATATATAAGC (SEQ ID NO:)
38	+7214	Desconocido	C/T	CAGTGACAT/TTGGAACAGGATC (SEQ ID NO:)

ES 2 356 167 T3

TABLA 9C

Descubrimiento de los SNP de IL-Rnic

5

10

15

20

25

30

35

40

45

50

55

60

65

Nº SNP	Con relación a +1	Localización	Tipo	Secuencia de los SNP
14	-4864		G/A	GTTGGGTAC/GTACCCGACGTGCTA (SEQ ID NO:)
15	-4744		T/G	ACTGTTACA/GAGCCAAGATATGGA (SEQ ID NO:)
18	-3662		T/G	TGTTTGCTT/TTCCTCTCTCAGC (SEQ ID NO:)
19	-3417		A/G	AGCTGGGTCT/ATGAGTTGTGGTGGC (SEQ ID NO:)
21	-3230		C/T	GTGTGTGTGT/CTGTGTTGTGTGTG (SEQ ID NO:)
22	-3188		G/A	GAGAGAATGA/GAATATGAGTGGTGG (SEQ ID NO:)
25	-2448		T/C	TATCTTGCTC/TTCCATTCTGATGC (SEQ ID NO:)
26	-2343		T/C	GTCACCATCA/TTGGGGTTGTGGATC (SEQ ID NO:)
29		Promotor	C/T	TGAGCCAAGG/CGGAAGAGAACAGGA (SEQ ID NO:)
30	-1002	Promotor	T/C	GCAGATAGCA/TCAGGTCCATTTTGC (SEQ ID NO:)
31	-501	Promotor	G/C	CTCAGAGAG/GGGCTTCCCTGGCCA (SEQ ID NO:)
32	-500	Promotor	G/A	TCTCAGAGAG/GGCTTCCCTGGCCAC (SEQ ID NO:)
33	-388	Promotor	C/T	TATTTTATTTG/CTAACTTGTTTCTTG (SEQ ID NO:)
34	-252	Promotor	A/C	GCACACATGC/ATGAGCTGGCGGCAG (SEQ ID NO:)
35	-41	Promotor	G/A	GGGAGGGGAG/GCTGGGCTCCTCCTT (SEQ ID NO:)
36	+40	Exón 1 5' UTR	G/A	CCCAGGACT/GCCCGGGTGCTACTT (SEQ ID NO:)
37	+96	Exón 1 5' UTR	G/A	GGAAGACCTC/AGAAGACCTCCTGTC (SEQ ID NO:)
38	+115	Exón 1 5' UTR	G/C	CCTGTCCTAT/GAGGCCCTCCCATG (SEQ ID NO:)
39	+162	Intrón 1	T/C	TCTCATTTTT/TCACCTGAGAAATGA (SEQ ID NO:)
40	+304	Intrón 1	A/G	GGAGGCATCC/ATGGGAGACCATGCA (SEQ ID NO:)
42	+2027	Intrón 1	G/T	GTGCATACTC/TGACTGGAACTGGA (SEQ ID NO:)
43	+2060	Intrón 1	G/C	AAAGGATAGA/GATGGAACCATGTGC (SEQ ID NO:)
44	+2245	Intrón 1	A/C	AGGGTAAATT/ATTTTTAGGATCCAA (SEQ ID NO:)
61	+11732	Exón 3	T/C	AACTAGTTGC/TGGATACTTGCAAGG (SEQ ID NO:)
62	+11787	Intrón 3	T/C	GCCAGGAAAG/CCAATGTATGTGGGC (SEQ ID NO:)
63	+11797	Intrón 3	G/A	CAATGTATGT/GGGCATCACGTCACT (SEQ ID NO:)
64	+11819	Intrón 3	G/C	CACTTTGCCC/GTCTGTCTGCAGCAG (SEQ ID NO:)
65	+11860	Intrón 3	T/G	TGCACAAACCCTAGG/TCAATGTCCTAATC (SEQ ID NO:)
66	+11855	Intrón 3	A/C	AAACCCTAGG/TGCAATGTCCTAATC (SEQ ID NO:)
67	+11890	Intrón 3	T/C	TGTATTCAAG/TTTGAAGCTGGGAGG (SEQ ID NO:)
68	+11892	Intrón 3	G/A	ATTCAAGTTT/GAAGCTGGGAGGGCC (SEQ ID NO:)
76	+13070	Intrón 3	G/A	AAAAATACCC/GGGGTCTCTTATTA (SEQ ID NO:)
77	+13092	Intrón 3	T/C	TTATTGCTGC/TTCCTCTTCTATTA (SEQ ID NO:)
78	+13307	Intrón 3	A/T	AAACAACCA/AATTTTTTCTTATGA (SEQ ID NO:)
79	+13414	Intrón 4	T/A	GCAGAGTGCC/TGGCTTGCCTGGGC (SEQ ID NO:)
80	+13520	Intrón 4	T/G	GCATGGCGGC/TGACTTCCAAAAGGG (SEQ ID NO:)
81	+13575	Intrón 4	G/A	CGCTTATTAT/GACTTCTGCTTGCAT (SEQ ID NO:)
82	+13614	Intrón 4	C/T	AAGCCAGTCA/CTGGCTAAGTCTAG (SEQ ID NO:)
89	+14818	Exón 4	T/C	GCTCAGACAG/CGGCCCCACCACCAG (SEQ ID NO:)

ES 2 356 167 T3

TABLA 9D

Descubrimiento de los SNP de IL-Rnsec

Nº SNP	Con relación a +1	Localización	Tipo	Secuencia de los SNP
61	+2014	Exón 2	T/C	AACTAGTTGC/TGGATACTTGCAAGG (SEQ ID NO:)
62	+2069	Intrón 2	T/C	GCCAGGAAAG/CCAATGTATGTGGGC (SEQ ID NO:)
63	+2080	Intrón 2	G/A	CAATGTATGT/GGGCATCACGTCCT (SEQ ID NO:)
64	+2101	Intrón 2	G/C	CACTTTGCCC/GTCTGTCTGCAGCAG (SEQ ID NO:)
66	+2142	Intrón 2	A/C	AAACCCTAGG/TGCAATGTCCTAATC (SEQ ID NO:)
67	+2180	Intrón 2	T/C	TGTATTC AAG/TTTGAAGCTGGGAGG (SEQ ID NO:)
68	+2183	Intrón 2	G/A	ATTCAAGTTT/GAAGCTGGGAGGGCC (SEQ ID NO:)
76	+3192	Intrón 2	G/A	AAAAATACCC/GGGGTCTCTTCATTA (SEQ ID NO:)
77	+3214	Intrón 2	T/C	TTATTGCTGC/TTCTCTTCTATTAA (SEQ ID NO:)
78	+3428	Intrón 2	A/T	AAACAACCAA/AATTTTTTCTTATGA (SEQ ID NO:)
79	+3535	Intrón 3	T/A	GCAGAGTGCC/TGGCTTGCGCTGGGC (SEQ ID NO:)
80	+3641	Intrón 3	T/G	GCATGGCGGC/TGACTTCCAAAAGGG (SEQ ID NO:)
81	+3696	Intrón 3	G/A	CGCTTATTAT/GACTTCTGCTTGCAT (SEQ ID NO:)
82	+3735	Intrón 3	C/T	AAGCCAGTCA/CGTGGCTAAGTCTAG (SEQ ID NO:)
89	+4939	Exón 4	T/C	GCTCAGACAG/CGGCCCCACCACCAG (SEQ ID NO:)
90	+5221	Exón 4 3' UTR	G/C	GGCGTCACAA/CAACCTGGTCACAGG (SEQ ID NO:)
91	+6047	Exón 4 3' UTR	G/A	CTCCCCACC/AGGCTGGGAGCTCTG (SEQ ID NO:)
93	+6410		T/C	GCAAAAAGA/TATGGGGCAGCACTG (SEQ ID NO:)
94	+6440		A/G	AACAGCCTCT/ACTGAAACAACCCA (SEQ ID NO:)
95	+6179		A/G	AAAGTCCCT/ACTCCTGTGACTTC (SEQ ID NO:)

ES 2 356 167 T3

TABLA 9E

Descubrimiento de los SNP de IL-1F5

Nº SNP	Con relación a +1	Localización	Tipo	Secuencia de los SNP
15	-2346	Promotor	T/C	ATTCAGGGAA/TGGAGGCTTAT (SEQ ID NO:)
16		Promotor	T/C	AGAGAACTTA/TGTTGGTGAGG (SEQ ID NO:)
17	-2063	Promotor	A/G	TTGACAGTAT/ACACAGTCATA (SEQ ID NO:)
18	-1953	Promotor	C/T	CTCACTGCAA/CCTCCACCTCC (SEQ ID NO:)
19			A/T	TGCAACCTCC/ACCTCCCAGGT (SEQ ID NO:)
20	-1834	Promotor	C/T	GGTTTTGCCA/CGTTGGCCAGG (SEQ ID NO:)
21	-1019	Promotor	C/T	GCTATTTTGC/CACATGAAAAA (SEQ ID NO:)
22	-926	Promotor	T/A	GTGGAGGAAA/TAGTGTTCAT (SEQ ID NO:)
23			G/C	GGAGGAAATA/GTGTTCATTG (SEQ ID NO:)
24			G/A	GAAATGGGAA/GGGAAAGAAG (SEQ ID NO:)
25		UTR	T/C	AAAGTTCGCGCTTCCCTGAGC (SEQ ID NO:)
26			C/T	GCCAGGATCC/CTGCTCAGAGCTCCG (SEQ ID NO:)
27		Intrón	T/C	CTGAGGGGGA/TAAGTGAACCTCCCA (SEQ ID NO:)
28		Intrón	T/C	CAGCACCTCA/CTGCTCTTAGAGTTT (SEQ ID NO:)
29		Intrón	G/A	CTGTGCCCT/ACCTCCCTGCTACTC (SEQ ID NO:)
30		Intrón	T/C	TTACCATCAG/TAGTCTGTCCA (SEQ ID NO:)
31		Intrón	A/T	TTCTGGGCTT/ACAACAGAGGA (SEQ ID NO:)
32		Intrón	T/C	GCACTTCCCA/TATGTGTTCTA (SEQ ID NO:)
33		Intrón	G/T	ATTGGAGGCA/GGATGGGGCAA (SEQ ID NO:)
34		Intrón	G/A	TTCTCTTCCA/GAGGAGATGGG (SEQ ID NO:)
35		Intrón	G/A	CCCAGGCATT/GGGGCAGCACAAACC (SEQ ID NO:)
36		Intrón	G/C	TGAAGACACA/GGCTCTGGGGCCACC (SEQ ID NO:)
37		Intrón	G/A	TGAAGCTTCT/ATGCAGAAGTGGGAA (SEQ ID NO:)
38		Intrón	G/A	TGCAGAAGTG/GGAATGTTTTCCAAG (SEQ ID NO:)
39		Intrón	G/C	CACTGCAACA/GTCCCTCTCTA (SEQ ID NO:)
40		Intrón	T/C	CAGGCCGTCT/TACAGCAGTCTGTG (SEQ ID NO:)
41			A/G	GTGGTCCCA/ATCGGTGGCTGGATG (SEQ ID NO:)
42		UTR	G/A	ATTTCTGTGG/GGTGGGGTGGG (SEQ ID NO:)
43		UTR	C/G	TGAGGAGGTG/GTGATGTCAGAAGAA (SEQ ID NO:)
44			G/A	GGTGGTGGCA/GTATAGGTGATTTT (SEQ ID NO:)
45		UTR	T/C	TGTTAATTTA/TCTGTATTTTCTAAT (SEQ ID NO:)
46		UTR	G/C	TTAATCTTGA/GGTAAGCAGAGTAGA (SEQ ID NO:)
47		UTR	T/C	GTAAGCAGAG/TAGACATCATCTCTGA (SEQ ID NO:)
48		UTR	A/C	GGAGGCTGTG/ATGAGTTTGTGTGGC (SEQ ID NO:)
49		UTR	T/C	ACAGAGGAGA/CGCGGGGAAGACTAT (SEQ ID NO:)
50		UTR	C/T	TAATTGGTTA/TGGCAGCTCTAGGAA (SEQ ID NO:)
51		UTR	A/G	AATTGGTTAC/GGCAGCTCTAGGAAA (SEQ ID NO:)
52			T/A	TTACAGCAGC/TCTAGGAACTAATA (SEQ ID NO:)
53			C/A	TACAGCAGCT/CTAGGAACTAATAC (SEQ ID NO:)
54		UTR	G/T	GAAGCTAGCT/TCCATGCTATGAGCA (SEQ ID NO:)
55		UTR	C/T	AAGAGACTTA/TGTGGTAAAAATGAA (SEQ ID NO:)
56		UTR	G/A	TGAGATAATC/AATGTTTGTGTTTT (SEQ ID NO:)
57		UTR	T/C	TTTGTGAGT/CATCTGGTTCTGCAA (SEQ ID NO:)

En el gen IL-1A, la cartografía reveló un total de 25 SNP. Dos SNP estaban en las regiones del exón y dieron como resultado cambios no sinónimos. Uno de ellos en el Exón 5 era conocido y se encontraba en la posición +4845, el otro en el Exón 4 tenía una frecuencia baja (2%) y los intentos posteriores para confirmarlos en otra población no mostraron variación. De los 11 SNP en la región promotora, 3 sólo se encontraron en un sujeto (2%) y los otros 8 estaban también en la base de datos pública del National Center for Biotechnology Information (NCBI dbSNP). Los 12 SNP restantes eran bien intrónicos (n=9) o se encontraban en la región no traducida (abreviadamente en lo sucesivo UTR por la expresión inglesa *UnTranslated Region*) 3' (n=3) del Exón 7. Diez de estos se encontraban también en la base de datos pública y los 2 restantes tenían una frecuencia muy baja (2-4%) en nuestra población.

Para identificar las variantes en el gen de IL-1B (IL-1F2) que más probablemente eran funcionales, los siete exones y las regiones intrónicas adyacentes, 5 kb en dirección 5' del gen y 2 kb en la región 3', que comprendían 12 kb, se secuenciaron en un panel de 25 individuos sanos étnicamente diversos. Se identificaron veinte SNP (Tabla 9B). Es

ES 2 356 167 T3

significativo que en siete exones secuenciados sólo se identificó un SNP en una región codificadora, el SNP sinónimo bien descrito en +3954. Existían tres SNP intrónicos, tres SNP en la UTR 3' del exón 7 y un SNP en la UTR 5' de exón 1 identificado. Los 12 SNP restantes estaban en la región promotora/potenciadora. Dos de los 12 SNP promotores no se encontraban en las bases de datos públicas y seis de los 12 tenían frecuencias de alelos menores $\leq 4\%$. Se confirmó la presencia de 5 SNP previamente indicados en las posiciones -511 (SNP 14), -31 (SNP15), -1468 (SNP10), +3954 (SNP 25) y +3877 (SNP 24) en el gen de IL-1B. Estos SNP se han asociado a estados clínicos. También se identificó la existencia de otros cinco SNP: -3893 (SNP3); -3737 (SNP4), +5548 (SNP30), +6911 (SNP37) y +7214 (SNP 38).

10 *Escrutinio de los SNP funcionales*

No se identificaron SNP no sinónimos en los exones de IL-1B. Fueron identificados SNP que alteraban la actividad transcripcional y/o la unión al factor de transcripción. Ajustando todos los SNP al alelo 1 excepto el SNP individual del foco, se realizó el escrutinio de la actividad del promotor después de la transfección de las construcciones de SNP en células THP-1. El análisis por transfección con construcciones parentales de longitud variable mostró que la delección del fragmento distal entre SNP 1 y SNP 10 potenció la actividad promotora, indicando fuertes represores en esta región. Una mayor actividad promotora de los plásmidos que contenían esta región entre SNP 17 y el comienzo del exón también reveló la existencia de potenciadores en la UTR en 5'. Con el fin de evitar la interferencia de los represores y potenciadores durante la identificación de los SNP funcionales individuales en la región promotora, se analizaron los SNP 2 a 17 en los fondos de plásmidos con o sin la región UTR en 5' en dirección 3' del SNP 15 (posible potenciador). También se analizaron los SNP 10 a 17 en los fondos de plásmidos con y sin fuerte interferencia potencial desde las regiones promotoras en dirección 5'. Debido a que se encontró que el clon del BAC inicial contenía el alelo 2 tanto en el SNP 14 como 15, se escrutaron también los SNP individuales para la actividad transcripcional en este fondo. Los resultados aquí recogidos utilizaron el plásmido promotor más largo que contenía secuencias en dirección 5' incluyendo SNP 2.

Se anotaron los SNP y se agruparon en forma esquemática con relación a su posición en el gen, cDNA, y la proteína como base para desarrollar una clara estrategia para sondear las diferencias específicas de los alelos en la función génica.

Se identificaron cinco SNP funcionales; B4, B7, B10, B14 y B15, demostrando las diferencias alélicas en la transcripción. Los SNP, B4, B10, B14 y B15 se presentan con elevada frecuencia ($>25\%$) en la población utilizada para el descubrimiento de los SNP; mientras que B7 se encontró sólo en el 2% de los sujetos secuenciados.

35 *Generación de las construcciones promotoras del gen IL-1B*

El promotor de IL-1B humano se amplificó por PCR a partir del clon del BAC RP11-67L14 (IL-1B, BAC PAC Resources, Children's Hospital Oakland Research Institute) utilizando DNA-polimerasa turbo *Pfu* turbo (Stratagene, CA). Se utilizaron dos conjuntos de cebadores para amplificar el promotor de IL-1B que contenía diferentes extremos 3'. Se utilizaron los cebadores de IL-1BF1 5'cccACGCGTGAGTGAAAGGAATCCCGTTAGAAAGT (SEQ ID NO: 33) e IL-1BR2 5'cccACGCGTGCCTGTTGTGCCTTGTGCCTCGAAG (SEQ ID NO: 34) para generar el fragmento promotor de IL-1B, IL-1BS (+32 a -5326) (S = corto; sin el primer exón). Se utilizó IL-1BR1 5'cc-cACGCGTGGCTGCTTCAGACACCTGTGTA (SEQ ID NO: 35) para generar el fragmento promotor de IL-1BL (+471 a -5326) (L=largo; que contenía el primer exón y SNP 17 (+45)). Los fragmentos promotores de IL-1B amplificados se clonaron en el vector pCRBlunt II-TOPO utilizando el kit de clonación Zero Blunt TOPO PCR (Invitrogen, CA) y a continuación se subclonó en el vector pGL3-Basic (Promega, WI) en dirección a 5' del gen indicador de la luciferasa para generar construcciones parentales con diferentes longitudes de la región promotora. La secuenciación de los clones del DNA del BAC reveló que los SNP en las posiciones 14 y 15 eran los alelos secundarios. Estos SNP se convirtieron secuencialmente en las secuencias de los alelos principales por mutagénesis dirigida al sitio de modo que las construcciones en serie parentales tanto para el plásmido más largo (que contenía los SNP 2-15 utilizando el sitio KpnI justo en dirección 5' del SNP 2) como para un plásmido corto (que contenía la región promotora fuera del sitio BamHI entre los SNP 9 y 10) eran uniformemente el alelo 1 en los 13 SNP (12 en la región flanqueante en 5' y 1 en el primer exón). La mutagénesis dirigida al sitio del SNP individual a las secuencias del alelo 2 se realizó con el kit de mutagénesis dirigida al sitio QuikChange XL (Stratagene, CA). Se confirmó el éxito de la mutagénesis del promotor de IL-1B por análisis de secuenciación y se volvió a clonar en el vector pGL3-Basic original para evitar posibles mutaciones introducidas por la PCR en el vector. El DNA utilizado en la transfección celular fue preparado por el kit EndoFree Plasmid Maxi (Qiagen, CA) para garantizar la mínima contaminación con endotoxinas. Los genotipos de las construcciones de SNP de IL-1B se recogen en la Tabla 6.

60 *Cultivo celular y Reactivos*

La línea de células monocíticas humanas, THP-1, se obtuvo de American Type Culture Collection (ATCC, VA). Las células THP-1 se desarrollaron en un medio RPMI 1640 (ATCC, VA) complementado con suero de bovino fetal al 10% (Hyclone, UT). Las células se subcultivaron cada dos días y se mantuvieron a una densidad celular entre 2×10^5 a 7×10^5 células/ml. El lipopolisacárido (LPS; de *Escherichia coli* 055:B5) y 12-miristato-13-acetato de forbol (PMA) se obtuvieron de Sigma (MO). Anti-NF- κ B p50 (11-119) y anti-NF- κ B p65(c-20) fueron adquiridos a Santa Cruz Biotechnology, Inc. (CA).

ES 2 356 167 T3

Transfección de las células y valoración de la actividad indicadora

Antes de la transfección, las células THP-1 se contaron, lavaron y volvieron a poner en suspensión en medio RPMI natural a $1,5-2 \times 10^7$ células/ml. Para la transfección, se mezclaron 200 ng de construcciones promotoras de pGL-IL1 y 6,25 ng de pRL-TK (Promega, WI) con 750 μ l de RPMI, 50 μ l de TRIS 1 M a pH 7,4, 100 μ l de DEAE-dextrano (2mg/ml) en solución salina tamponada con fosfato y 100 μ l de células THP-1. Las mezclas se incubaron a 37°C durante 30 minutos. Las células se centrifugaron, lavaron, volvieron a poner en suspensión en 0,5 ml de medio de cultivo y se transfirieron a una placa de cultivo de 24 pocillos. Dieciocho a veinticuatro horas después de la transfección, se estimularon las células con 12-miristato-13-acetato de forbol (PMA, 20 ng/ml) y lipopolisacárido (LPS) durante 6 horas. A continuación se lisaron las células en 100 μ l de tampón de lisis pasivo (Promega, WI). Se analizó el lisado celular (20 μ l) para determinar las actividades tanto de la luciferasa de luciérnaga como la de *Renilla* utilizando el sistema de valoración del indicador dual-luciferasa (Promega, WI). La actividad de la luciferasa se midió con un luminómetro 1450 MicroBeta (Perkin Elmer, MA). La actividad promotora se normalizó a la actividad de la *Renilla* a partir de pRL-TK transfectado conjuntamente. Cada construcción se transfectó y valoró para determinar por triplicado la actividad promotora y se realizó la transfección para cada construcción al menos tres veces.

Valoración del desplazamiento de la electromovilidad (abreviadamente EMSA por la expresión inglesa Electromobility Shift Assay)

Se cultivaron células THP-1 como se ha descrito anteriormente y se recolectaron sin tratamiento o se trataron con 100 ng/ml de LPS y 20 ng/ml de PMA durante varios periodos de tiempo. Los extractos nucleares se prepararon según Andrews N.C. y Faller D.V.³¹. En resumen, las células se volvieron a poner en suspensión en tampón A frío (Hepes-KOH 10 mM, pH 7,9 - MgCl₂ 1,5 mM - KCl 10 mM - DTT 0,5 mM - AEBSF 2 mM) y se incubaron sobre hielo durante 10 minutos. Los sedimentos nucleares se volvieron a poner en suspensión en tampón C frío (Hepes-KOH 20 mM, pH 7,9 - glicerol al 25% - NaCl 420 mM - MgCl₂ 1,5 mM - EDTA 0,2 mM - DTT 0,5 mM - AEBSF 2 mM) durante 30 minutos sobre hielo. Después de centrifugación durante 2 minutos a 4°C, se conservaron los extractos nucleares a -80°C. Se sintetizaron los oligonucleótidos de cada SNP (MWG Biotech, High Point, NC) y se purificaron en gel de poliacrilamida al 10%. Los oligonucleótidos monocatenarios complementarios se re-asociaron y etiquetaron con α -³²P-dGTP por relleno con Klenow (exo-). Para la unión proteína-DNA, se incubaron 7 μ g de extractos nucleares con $5,0 \times 10^4$ cpm de sondas marcadas con ³²P durante 20 minutos a la temperatura ambiente en un volumen de 20 μ l de tampón de unión (Hepes-KOH 12 mM, pH 7,9 - MgCl₂ 1,5 mM - KCl 60 mM - EDTA 0,2 mM - DTT 0,5 mM - AEBSF 0,4 mM - glicerol al 10%) y 2 μ g de Poly (dI-dC) (Sigma, St. Louis, MO). Los complejos proteína-DNA se resolvieron sobre gel de poliacrilamida al 4% y se visualizaron por exposición a películas de rayos X. Para las valoraciones de competición con oligonucleótidos fríos, se incubaron extractos nucleares con una sonda fría durante 10 minutos a la temperatura ambiente antes de la adición de la sonda etiquetada. En las valoraciones del superdesplazamiento de anticuerpos, se añadieron 1-2 μ g de anticuerpo después de la unión de los extractos nucleares a la sonda marcada y se incubó durante 15 minutos a la temperatura ambiente antes de la carga del gel. Las secuencias de oligonucleótidos son las siguientes:

Nombre del oligonucleótido	5' - 3'
B4C	cggagaatggaatgtcccttgactctgcatga (SEQ ID NO:)
B4C-r	cgcgatcatgcagagtcgaaggacattccattctccgagct (SEQ ID NO:)
B4T	cggagaatggaatgttcccttgactctgcatga (SEQ ID NO:)
B4T-r	cgcgatcatgcagagtcgaaggacattccattctccgagct (SEQ ID NO:)
B10G	ctcactcccttgataatgcagagcgagca (SEQ ID NO:)
B10G-r	cgcggtctcgctctgattatccaaggagtgagagct (SEQ ID NO:)
B10C	ctcactcccctgcataatgcagagcgagca (SEQ ID NO:)
B10C-r	cgcggtctcgctctgattatccaaggagtgagagct (SEQ ID NO:)
B14C	ccaattgacagagagctcccaggcagagaacagca (SEQ ID NO:)
B14C-r	cgcggtctgttctctgctcgggagctctctgtcaattggagct (SEQ ID NO:)
B14T	ccaattgacagagagctcctgaggcagagaacagca (SEQ ID NO:)
B14T-r	cgcggtctgttctctgctcaggagctctctgtcaattggagct (SEQ ID NO:)
B15T	ccttctgctttgaaagctataaaaaacagcgaggagca (SEQ ID NO:)
B15T-r	cgcggtctcctcgctgttttatagctttcaaaagcagaaggagct (SEQ ID NO:)
B15C	ccttctgctttgaaagccataaaaaacagcgaggagca (SEQ ID NO:)
B15C-r	cgcggtctcctcgctgttttatagctttcaaaagcagaaggagct (SEQ ID NO:)

SNP B4 (-3737)

SNP B4 no mostró diferencias de transcripción por alelo, sobre los fondos de los demás

ES 2 356 167 T3

SNP incluidos en el alelo 1 (Figura 10).

El análisis EMSA muestra diferencias significativas en la unión de proteínas entre sondas que contienen el alelo 1 y el alelo 2 de SNP B4. (Figuras 15-17) Tanto el alelo 1 como el 2 muestran unión constante de una banda de migración rápida (complejo 3) a las 0, 4, 6 y 24 horas de estimulación con LPS de las células. El alelo 1 muestra una fuerte unión de dos complejos de alta movilidad después de la estimulación con LPS. El complejo 1 consiste en dos bandas próximas entre sí y está presente al comenzar las 4 horas. A las 24 horas, desaparece la banda inferior del complejo 1 y se vuelve evidente el complejo 2 (Figura 15) y la competición con la sonda no marcada demostró la especificidad de la unión (Figura 16). El alelo 2 también se unió al complejo 2 a las 24 horas pero no había pruebas del complejo 1 en cualquiera de estos momentos. Debido a la presencia de un elemento regulador NF- κ B próximo a SNP B4, se utilizaron anticuerpos tanto para subunidades p50 como p65 de NF- κ B en la incubación de la EMSA de las sondas del alelo-1 y el alelo-2. Como se muestra en la Figura 17, pistas 7 y 8, el anticuerpo p50 es capaz de superdesplazar ambas bandas del complejo 1 y el anticuerpo p65 es capaz de super-desplazar sólo la banda superior del complejo 1. Este resultado indica que NF- κ B estimulado por LPS/PMA se une preferiblemente al alelo 1 de SNP B4.

SNP B10 (-1468)

En la construcción de fondo con los demás SNP situados en el alelo 1, A) la presencia del alelo 2 en SNP 1310 produjo una actividad transcripcional ligeramente menor. (Véase la Figura 11).

El análisis EMSA con sondas que contenían el alelo-1 y el alelo-2 de SNP10 demostró claras diferencias en los perfiles de unión de las proteínas (Figura 18-19) con el alelo 2 de SNP B 10 que se unía a un grupo de proteínas de alta movilidad (HMG) (complejo 1) más fuertemente que el alelo 2. La competición con una sonda no marcada mostró que esta fuerte preferencia de unión era específica del alelo (Figura 19).

SNP B14 (-511)

Los resultados del análisis transcripcional para efectos individuales de diversos alelos sólo en SNP 14 se incluyen en la Figura 12. En las construcciones de fondo en las que los demás SNP son el alelo 1 (Fig. 20), SNP B14 se convirtió en el alelo 2 (Fig. 20) y mostró una actividad promotora mayor.

La EMSA con sondas que contenían variantes alélicas de SNP 14 mostró 2 complejos distintos formados con esta sonda pero sin diferencias alélicas en la movilidad de complejos proteínicos. (Figura 20). Sin embargo, la actividad promotora potenciada por el alelo 2 de SNP 14 indica que puede haber diferencias sutiles en las composiciones de las proteínas entre los dos complejos unidos a elementos reguladores que rodean el alelo 1 y 2 del SNP 14 lo que conduce al impacto funcional en la actividad transcripcional.

SNP 15 (-31)

El efecto de variación alélica sólo en SNP B15 en la actividad transcripcional está también incluido en las Figuras 21-22. Con la conversión de SNP B15 en el alelo 2 (Figura 21), se produjo una reducción significativa de la actividad promotora en comparación con la construcción base (Figura 21).

El perfil proteínico en EMSA de SNP B15 es notablemente diferente para las sondas del alelo-1 (T) que para las del alelo-2 (C) (Figura 22). Como se muestra en la Figura 22, se formaron dos bandas intensas con la sonda del alelo 1 mientras que se formaron 3 bandas distintas con el alelo 2. Se realizó una valoración de competición en frío para determinar la especificidad de la unión de las proteínas a estas sondas alélicas (Figura 22) y se indicó que la unión proteínica del complejo 1 es muy específica para el alelo 2, mientras que los complejos II y III tienen mayor afinidad para el alelo 1.

Ejemplo 10

Determinación de la gravedad de la periodontitis en sujetos que tenían genotipos de IL-1B

Una indicación de la gravedad de la periodontitis en un sujeto es la profundidad de las cavidades (abreviadamente PD por la expresión inglesa *pocket depth*). Se clasifica periodontitis grave en un sujeto que tiene una PD > 2 en una o más cavidades en las encías. La PD se midió en sujetos que tenían diversos genotipos de IL-1B como se muestra en la Tabla 10A (por ejemplo, un sujeto con el genotipo BP1 tiene dos alelos del alelo 1 de IL-1B (-511), dos alelos del alelo 1 de IL-1B (-1468) y dos alelos del alelo 2 de IL-1B (-3737). La Tabla 10B muestra la distribución en porcentaje de los haplotipos de IL-1B en sujetos con grados variables de periodontitis, medidos por el número de cavidades en las encías por sujeto (menos de 2 cavidades, entre 2 y 4 cavidades, entre 4 y 10 cavidades, más de 10 cavidades).

ES 2 356 167 T3

TABLA 10A

Genotipos	Alelos de SNP de IL-1B		
	-511	-1468	-3737
BP1	1,1	1,1	2,2
BP2	2,2	2,2	1,1
BP3	1,1	1,1	1,1
BP4	2,2	1,1	1,1

TABLA 10B

Genotipos	% de distribución			
	0-2 cavidades en encías	2 a 4 cavidades en encías	4 a 10 cavidades en encías	>10 cavidades en encías
BP1	0,429	0,089	0,304	0,179
BP2	0,833	0,000	0,167	0,000
BP3	0,375	0,094	0,188	0,344
BP4	0,364	0,091	0,000	0,545

Como se muestra en la Tabla 10B y en la Figura 23, los sujetos que tienen los genotipos BP1, BP3 o BP4 tienen mayor probabilidad de desarrollar periodontitis grave, mientras que los sujetos que tienen el genotipo BP2 tienen menor probabilidad de desarrollar periodontitis grave.

Ejemplo 11

Determinación de la gravedad de la periodontitis en poblaciones étnicas

Anteriormente, en múltiples estudios se asociaron en hombres de raza blanca las variaciones específicas de los genes de IL-1 (IL-1A(+4845), IL-1A(-889) e IL-1B(+3954)) con el aumento de la gravedad de la periodontitis. Estos SNP de IL-1 son de baja prevalencia en las poblaciones asiáticas que han sido estudiadas, que incluyen japoneses, coreanos y chinos. Puesto que la prevalencia de un polimorfismo y su penetración con relación a la expresión de la enfermedad afecta el tamaño de la muestra necesario para identificar la asociación entre el alelo de interés y la expresión de la enfermedad, la mayor parte de los estudios anteriores en poblaciones asiáticas han sido demasiado pequeños para evaluar fielmente si IL-1A(+4845) e IL-1B(+3954) están asociados a la periodontopatía. Este ejemplo detalla las variaciones específicas de los genes de IL-1 que predicen la periodontitis más grave en adultos japoneses. Esto se realizó analizando los haplotipos en las agrupaciones de genes de IL-1 para determinar las variaciones genéticas asociadas a la varianza clínica en la periodontitis, utilizando la cartografía de SNP de elevada densidad de todos los exones y regiones reguladoras de los genes para IL-1 α (IL-1A), IL-1 β (IL-1B) y el gen para el antagonista receptor IL-1 (IL-1RN), que abarca la agrupación de nueve genes de IL-1 en el cromosoma 2q13-14.

Existen varios haplotipos predominantes de IL-1 que se definen por seis SNP. De los 64 haplotipos posibles, 23 haplotipos constituyen más del 98% de los haplotipos en la raza blanca y afroamericana. Los mismos haplotipos constituyen más del 97% de los haplotipos observados en pacientes japoneses. (Véase la Tabla 11A). Los haplotipos que contienen el alelo 2 en IL-1RN(+2018) sólo representan el 3,7% de los haplotipos, lo que produciría un índice esperado de individuos homocigóticos en dicho *locus* de aproximadamente 1/1000 sujetos. El alelo 2 en IL-1A(+4845) y en IL-1B(+3954) era también tan poco frecuente que sólo se esperaría que aproximadamente 1/100 individuos fueran homocigóticos en dicho *locus*. Los haplotipos se estimaron utilizando el programa *haplo.score* versión 1.0 (Schaid *et al*, 2002). Este programa informático aplica una versión del algoritmo EM (Excoffier and Slatkin, 1995) para valorar correctamente la ambigüedad de fase.

Tabla 11A. Frecuencia de haplotipos de IL-1 en sujetos de diversas etnias.									
Gen:	IL-1A		IL-1B		IL-1RN				
	4845	3954	-511	-1468	-3737	2018	Frec. en japoneses	Frec. en raza blanca	Frec. en raza negra
Haplotipo									
IL1C1	1	1	1	1	2	1	0,458	0,339	0,241
IL1C2	2	2	1	1	1	1	0,039	0,145	0,029
IL1C3	1	1	2	2	1	2	0,03	0,137	0,05
IL1C4	1	1	2	2	1	1	0,279	0,097	0,056
IL1C5	1	1	1	1	2	2	0,007	0,071	ND
IL1C6	1	1	2	1	1	1	0,048	0,038	0,331
IL1C7	2	1	1	1	2	1	0,015	0,031	0,013
IL1C8	1	2	1	1	1	1	ND	0,021	ND
IL1C9	2	1	2	2	1	2	ND	0,02	ND
IL1C10	2	2	1	1	1	2	ND	0,019	ND
IL1C11	2	1	2	1	1	1	0,025	0,016	0,039
IL1C12	1	1	1	1	1	1	0,027	0,015	0,061
IL1C13	2	2	2	2	1	1	ND	0,01	ND
IL1C14	2	2	1	1	2	1	ND	0,007	0,019
IL1C15	2	2	2	2	1	2	ND	0,007	ND
IL1C16	1	1	2	1	1	2	ND	0,006	0,009
IL1C17	2	1	2	2	1	1	0,008	0,006	ND
IL1C18	1	1	1	2	1	1	0,021	ND	ND
IL1C19	1	1	2	1	2	1	0,01	ND	ND
IL1C20	2	1	1	1	1	1	0,005	ND	0,049
IL1C21	1	1	1	1	1	2	?	?	0,008
IL1C22	1	2	2	1	1	1	?	?	0,035
IL1C23	2	2	2	1	1	1	?	?	0,047
						Totales	0,972	0,985	0,987

Se ha identificado tres SNP en el promotor de IL-1B que son funcionales y altamente prevalentes en sujetos japoneses. Estos SNP incluyen IL-1B (-511), IL-1B (-1468) e IL-1B (-3737). La Figura 24 muestra la distribución de las frecuencias relativas en los haplotipos de los alelos recogidos en la Tabla 11A en sujetos de raza blanca y japoneses. Los autores de la presente invención han determinado que los haplotipos predominantes, basados en estos 3 SNP, muestran diferencias significativas en los niveles de IL-1 β en el fluido gingival (GCF). La Figura 25 demuestra que los sujetos de raza blanca que tienen el genotipo H1 o H3 presentan mayor inflamación, medida en comparación con los demás haplotipos utilizando una estadística de puntuación. Una estadística de puntuación es una estadística utilizada para evaluar el significado estadístico de estimaciones paramétricas calculadas por métodos de probabilidad máxima. También se denomina algunas veces *estadística de puntuación eficaz*. El ensayo se basa en el comportamiento de la función de probabilidad logarítmica en el punto en el que la estimación paramétrica respectiva es igual a 0,0 (cero); específicamente, se usa la derivada (pendiente) de la función de probabilidad logarítmica evaluada en el valor de hipótesis nula del parámetro (parámetro = 0,0).

En sujetos de etnia japonesa, se ha determinado que la predisposición a la periodontopatía y la gravedad de dicha enfermedad están afectadas por el haplotipo de IL-1.

Un método para medir la periodontopatía utiliza el ensayo generalizado de hemorragia por sondamiento (en inglés *Bleeding on Probing* (BOP)). Como se muestra en la Figura 26, los sujetos japoneses que tienen un genotipo que es homocigótico en el alelo 2 de IL-1B (-511), homocigótico en el alelo 2 de IL-1B (-1468) y homocigótico en el alelo 1 de IL-1B (-373) tienen menor riesgo de periodontitis medido por el ensayo generalizado de BOP. Esta disminución del riesgo se observa en dos aumentos diferentes en porcentaje en las superficies cubiertas de placa.

Otro método para medir la periodontopatía que permite obtener la asociación entre haplotipos específicos de IL-1 y la gravedad de la periodontitis implica la medida de la profundidad del sondeo y el nivel de acoplamiento de la sonda de un sujeto. Estas medidas se realizan a individuos, cuyos haplotipos están identificados, y se determinan el porcentaje medio de los sitios medidos que tengan una profundidad de sondeo mayor que 4 mm o el nivel de acoplamiento de la sonda mayor o igual a 4 mm. Como se muestra en la Figura 27, los sujetos japoneses que tienen un haplotipo H1, H3 o H4 de IL-1 presentan una mayor gravedad de la periodontitis en comparación con sujetos que tienen un haplotipo H2, medido por estos métodos.

REIVINDICACIONES

5 1. Un método para predecir la sensibilidad de un sujeto humano a la periodontopatía, que comprende la etapa de identificar en una muestra de DNA genómico del sujeto un modelo alélico que comprende una de las siguientes series de alelos de IL-1B:

(1) alelo 1 (-511), alelo 1 (-1468) y alelo 2 (-3737);

10 (2) alelo 2 (-511), alelo 2 (-1468) y alelo 1 (-3737);

(3) alelo 1 (-511), alelo 1 (-1468) y alelo 1 (-3737); o

(4) alelo 2 (-511), alelo 1 (-1468) y alelo 1 (-3737),

15 en donde la presencia de uno cualquiera de los modelos alélicos 1, 3 o 4 indica que dicho sujeto tiene mayor sensibilidad a la periodontopatía y en donde la presencia del modelo alélico 2 indica que el sujeto tiene menor sensibilidad a la periodontopatía.

20 2. El método de la reivindicación 1, en donde el sujeto es homocigótico para cada uno de los alelos en el modelo alélico.

25 3. El método de la reivindicación 1 o 2, en donde dicho sujeto es de etnia asiática.

30

35

40

45

50

55

60

65

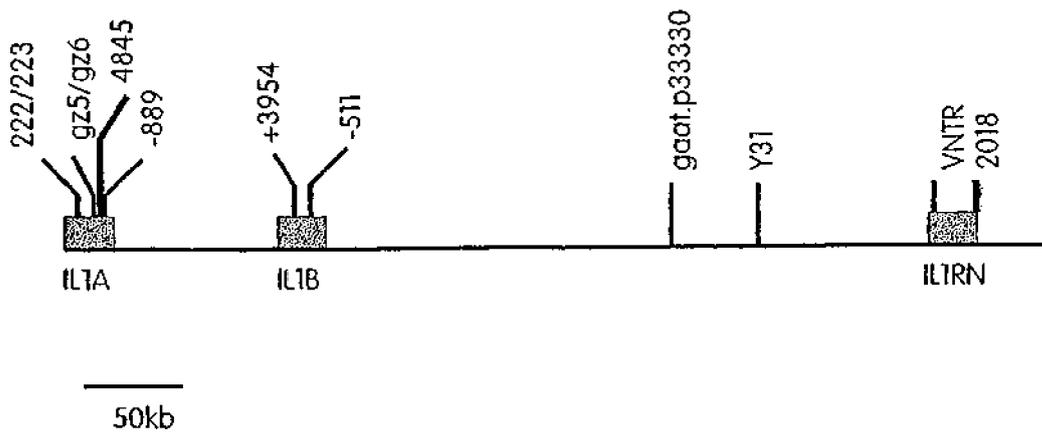


Fig. 1

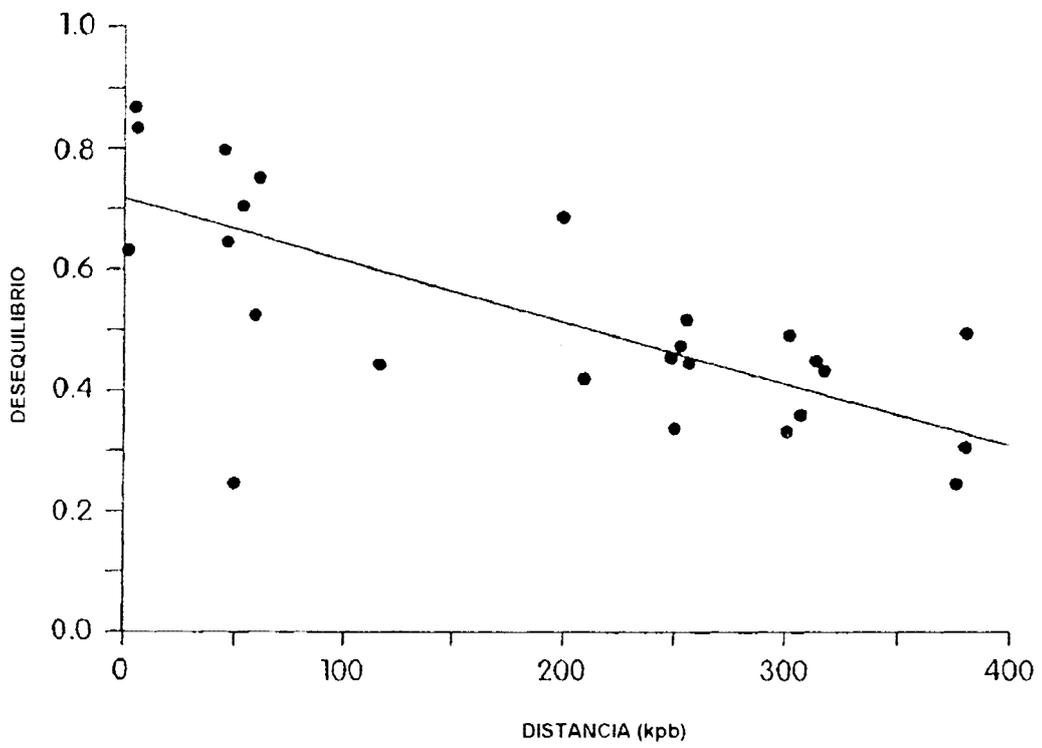


Fig. 2

ES 2 356 167 T3

-1437 AAGCTTCTAC CCTAGTCTGG TGCTACACTT ACATTGCTTA CATCCAAGTG TGGTTATTTT
-1377 TGTGGCTCCT GTTATAACTA TTATAGCACC AGGTCTATGA CCAGGAGAAT TAGACTGGCA
-1317 TTAAATCAGA ATAAGAGATT TTGCACCTGC AATAGACCTT ATGACACCTA ACCAACCCCA
-1257 TTATTTACAA TTAACAGGA ACAGAGGGAA TACTTTATCC AACTCACACA AGCTGTTTTT
-1197 CTCCCAGATC CATGCTTTTT TGCGTTTATT ATTTTTTAGA GATGGGGGCT TCACTATGTT
-1137 GCCCACACTG GACTAAAACCT CTGGGCTCA AGTGATTGTC CTGCCTCAGC CTCCTGAATA
-1077 GCTGGGACTA CAGGGGCATG CCATCACACC TAGTTCATTT CCTCTATTTA AAATATACAT
-1017 GGCTTAAACT CCAACTGGGA ACCCAAACA TTCATTTGCT AAGAGTCTGG TGTTCTACCA
-957 CCTGAACTAG GCTGGCCACA GGAATTATAA AAGCTGAGAA ATTCTTTAAT AATAGTAACC
-897 AGGCAACATC ATTGAAGGCT CATATGTAAA AATCCATGCC TTCCTTTCTC CCAATCTCCA
-837 TTCCCAAACCT TAGCCACTGG TTCTGGCTGA GGCCTTACGC ATACCTCCCG GGGCTTGCAC
-777 ACACCTTCTT CTACAGAAGA CACACCTTGG GCATATCCTA CAGAAGACCA GGCTTCTCTC
-717 TGGTCCITGG TAGAGGGCTA CTTTACTGTA ACAGGGCCAG GGTGGAGAGT TCTCTCCTGA
-657 AGCTCCATCC CCTCTATAGG AAATGTGTTG ACAATATTCA GAAGAGTAAG AGGATCAAGA
-597 CTCTTTGTG CTCAAATACC ACTGTTCTCT TCTTACCCT GCCCTAACCA GGAGCTTGTC
-537 ACCCCAAACT CTGAGGTGAT TTATGCCTTA ATCAAGCAAA CTTCCTCTT CAGAAAAGAT
-477 GGCTCATTTT CCCTCAAAG TGCCAGGAG CTGCCAAGTA TTCTGCCAAT TCACCCTGGA
-417 GCACAATCAA CAAATTCAGC CAGAACAAA CTACAGCTAC TATTAGAAT ATTATTATTA
-357 ATAAATTCCT CTCCAAATCT AGCCCCTTGA CTTCGGATTT CACGATTTCT CCCTCCTCC
-297 TAGAAACTTG ATAAGTTTCC CGCGCTTCCC TTTTCTAAG ACTACATGTT TGTCATCTTA
-237 TAAAGCAAAG GGGTGAATAA ATGAACCAA TCAATAACTT CTGGAATATC TGCAAACAAC
-177 AATAATATCA GCTATGCCAT CTTTCACTAT TTTAGCCAGT ATCGAGTTGA ATGAACATAG
-117 AAAAATACAA AACTGAATTC TTCCCTGTAA ATTCCCCGTT TTGACGACGC ACTTGTAGCC
-57 ACGTAGCCAC GCCTACTTAA GACAATTACA AAAGGCGAAG AAGACTGACT CAGGCTTAAG
4 CTGCCAGCCA GAGAGGGAGT CATTTTCATTG GCGTTTGAGT CAGCAAAGGT ATTGTCCTCA
64 CATCTCTGGC TATTAAAGTA TTTTCTGTTG TTGTTTTTCT CTTTGGCTGT TTTCTCTCAC
124 ATTGCCCTTCT CTAAGCTAC AGTCTCTCCT TTCTTTTCTT GTCCCTCCCT GGTGTTGGTAT
184 GTGACCTAGA ATTACAGTCA GATTTTCAGAA AATGATTCTC TCATTTTGCT GATAAGGACT
244 GATTCGTTTT ACTGAGGGAC GGCAGAACTA GTTTCCTATG AGGGCATGGG TGAATACAAC
304 TGAGGCTTCT CATGGGAGGG AATCTCTACT ATCCAAAAT ATTAGGAGAA AATTGAAAAT
364 TTCCAACTCT GTCTCTCTCT TACCTCTGTG TAAGGCAAAT ACCTTATTCT TGTGGTGT
424 TTGTAACCTC TTCAAACCTT CATTGATTGA ATGCCTGTTT TGGCAATACA TTAGGTTGGG
484 CACATAAGGA ATACCAACAT AAATAAAACA TTCTAAAAGA AGTTTACGAT CTAATAAAGG
544 AGACAGGTAC ATAGCAAACCT AATTCAAAGG AGCTAGAAGA TGGAGAAAAT GCTGAATGTG
604 GACTAAGTCA TTCAACAAAG TTTTCAGGAA GCACAAAGAG GAGGGGCTCC CCTCACAGAT
664 ATCTGGATTA GAGGCTGGCT GAGCTGATGG TGGCTGGTGT TCTCTGTTGC AGAAGTCAAG
724 ATGGCCAAAG TTCCAGACAT GTTTGAAGAC CTGAAGAACT GTTACAGGTA AGGAATAAGA
784 TTTATCTCTT GTGATTTAAT GAGGGTTTCA AGGCTCACCA GAATCCAGCT AGGCATAACA
844 GTGGCCAGCA TGGGGGCAGG CCGGCAGAGG TTGTAGAGAT GTGTACTAGT CCTGAAGTCA
904 GAGCAGGTTT AGAGAAGACC CAGAAAAACT AAGCATTGAG CATGTTAAAC TGAGATTACA
964 TTGGCAGGGA GACCGCCATT TTAGAAAAAT TATTTTTGAG GTCTGCTGAG CCCTACATGA
1024 ATATCAGCAT CAACTTAGAC ACAGCCTCTG TTGAGATCAC ATGCCCTGAT ATAAGAATGG
1084 GTTTTACTGG TCCATTCTCA GGAAACTTG ATCTCATTCA GGAACAGGAA ATGGCTCCAC
1144 AGCAAGCTGG GCATGTGAAC TCACATATGC AGGCAAATCT CACTCAGATG TAGAAGAAAG
1204 GTAAATGAAC ACAAAGATAA AATTACGGAA CATATTAAC TAACATGATG TTTCCATTAT
1264 CTGTAGTAAA TACTAACACA AACTAGGCTG TCAAAATTTT GCCTGGATAT TTTACTAAGT
1324 ATAAATATG AAATCTGTTT TAGTGAATAC ATGAAAGTAA TGTGTAACAT ATAATCTATT
1384 TGTTAAAAAT AAAAAGGAAG TGCTTCAAAA CCTTTCTTTT CTCTAAAGGA GCTTAACATT
1444 CTTCCTGAA CTTCAATTAA AGCTCTTCAA TTTGTTAGCC AAGTCCAATT TTTACAGATA
1504 AAGCACAGGT AAAGCTCAA GCCTGTCTTG ATGACTACTA ATTCCAGATT AGTAAGATAT

Fig. 3

```

1564 GAATTACTCT ACCTATGTGT ATGTGTAGAA GTCCTTAAAT TTCAAAGATG ACAGTAATGG
1624 CCATGTGTAT GTGTGTGACC CACAACATC ATGGTCATTA AAGTACATTG GCCAGAGACC
1684 ACATGAAATA ACAACAATTA CATTCTCATC ATCTTATTTT GACAGTGAAA ATGAAGAAGA
1744 CAGTTCCTCC ATTGATCATC TGTCTCTGAA TCAGGTAAGC AAATGACTGT AATCTCATG
1804 GGACTGCTAT TCTTACACAG TGGTTTCTTC ATCCAAAAGAG AACAGCAATG ACTTGAATCT
1864 TAAATACTTT TGTTTTACCC TCACTAGAGA TCCAGAGACC TGTCTTTCAT TATAAGTGAG
1924 ACCAGCTGCC TCTCTAAACT AATAGTTGAT GTGCATTGGC TTCTCCCAGA ACAGAGCAGA
1984 ACTATCCCAA ATCCCTGAGA ACTGGAGTCT CCTGGGGCAG GCTTCATCAG GATGTTAGTT
2044 ATGCCATCCT GAGAAAGCCC CGCAGGCCGC TTCACCAGGT GTCTGTCTCC TAACGTGATG
2104 TGTTGTGGTT GTCTTCTCTG ACACCAGCAT CAGAGGTTAG AGAAAGTCTC CAAACATGAA
2164 GCTGAGAGAG AGGAAGCAAG CCAGCTGAAA GTGAGAAGTC TACAGCCACT CATCAATCTG
2224 TGTTATTGTG TTTGGAGACC ACAAATAGAC ACTATAAGTA CTGCCTAGTA TGTCTTCAGT
2284 ACTGGCTTTA AAAGCTGTCC CCAAAGGAGT ATTTCTAAAA TATTTTGAGC ATTGTTAAGC
2344 AGATTTTTAA CCTCCTGAGA GGAACATAAT TGGAAAGCTA CCACTACTA CAATCATTGT
2404 TAACCTATTT AGTTACAACA TCTCATTTTT GAGCATGCAA ATAAATGAAA AAGTCTTCCT
2464 AAAAAAATCA TCTTTTTATC CTGGAAGGAG GAAGGAAGGT GAGACAAAAG GGAGAGAGGG
2524 AGGGAAGCCT AATGAAACAC CAGTTACCTA AGACCAGAAT GGAGATCCTC CTCACTACCT
2584 CTGTTGAATA CAGCACCTAC TGAAAGAACT TTCATTCCCT GACCATGAAC AGCCTCTCAG
2644 CTTCTGTTTT CTTTCCCTAC AGAAATCCTT CTATCATGTA AGCTATGGCC CACTCCATGA
2704 AGGCTGCATG GATCAATCTG TGTCTCTGAG TATCTCTGAA ACCTCTAAA CATCCAAGCT
2764 TACCTTCAAG GAGAGCATGG TGGTAGTAGC AACCAACGGG AAGGTTCTGA AGAAGAGACG
2824 GTTGAGTTA AGCCAATCCA TCACTGATGA TGACCTGGAG GCCATCGCCA ATGACTCAGA
2884 GGAAGGTAAG GGGTCAAGCA CAATAATATC TTTCTTTTAC AGTTTTAAGC AAGTAGGGAC
2944 AGTAGAATTT AGGGGAARAT TAAACGTGGA GTCAGAATAA CAAGAAGACA ACCAAGCATT
3004 AGTCTGGTAA CTATACAGAG GAAAATTAAT TTTTATCCTT CTCCAGGAGG GAGAAATGAG
3064 CAGTGGCCTG AATCGAGAAT ACTTGCTCAC AGCCATTATT TCTTAGCCAT ATFGTAAAGG
3124 TCGTGTGACT TTTAGCCTTT CAGGAGAAAG CAGTAATAAG ACCACTTACG AGCTATGTTT
3184 CTCTCATACT AACTATGCCT CCTTGGTCAT GTTACATAAT CTTTTCGTGA TTCAGTTTCC
3244 TCTACTGTAA AATGGAGATA ATCAGAATCC CCCACTCATT GGATTGTTGT AAAGATTAAG
3304 AGTCTCAGGC TTTACAGACT GAGCTAGCTG GGCCCTCCTG ACTGTTATAA AGATTAAATG
3364 AGTCAACATC CCCTAACTT C TGACTAGAA TAATGTCTGG TACAAAGTAA GCACCCAATA
3424 AATGTTAGCT ATTACTATCA TTATTATTAT TATTTTATTT TTTTTTTTTG AGATGGAGTC
3484 TGGCTCTGTC ACCCAGGCTG GAGTGCAGTG GCACAATCTC GGCTCACTGC AAGCTCTGCC
3544 TCCTGGGTTT ATGCCATTCT CCTGCCTCAG CCTCCCGAGT AAGCTGGGAA TACAGGCACC
3604 CGCCACTGTT CCCGGCTAAT TTTTGTATT TTTAGTAGAG ACGGAGTTT ACCGTGGTCT
3664 CCATCTCCTC GTGATCCACC CACCTTGGCC TCCCAAAGTG CCGGGATTAC AGGCGTGAGC
3724 CACCGCGCCC GGCCATTAT TATTATTATT ACTACTACTA CTACCTATAT GAATACTACC
3784 AGCAATACTA ATTTATTAAT GACTGGATTA TGTCTAAACC TCACAAGAAT CCTACCTTCT
3844 CATTTTACAT AAAAGGAAAC TAAGCTCATT GAGATAGGTA AACTGCCCAA TGGCATACAT
3904 CTGTAAGTGG GAGAGCCTCA AATCTAATTC AGTTCTACCT GAGTAAAAAA ATCATGGTTT
3964 CTCCTCCATC CTTTACTGT ACAAGCCTCC ACATGAACTA TAAACCCAAT ATTCCTGTTT
4024 TTAAGATAAT ACCTAAGCAA TAACGCATGT TCACCCTAGAA GGTTTTAAAA TGTAACAAAA
4084 TATAAGAAAA TAAAAATCAC TCATATCGTC AGTGAGAGTT TACTACTGCC AGCACTATGG
4144 TATGTTTCTT TAAATCTTT GCTATACACA TACCTACATG TGAACAAATA TGTCTAACAT
4204 CAAGACCACA CTATTPACAA CTTTATATCC AGCTTTTCTT ACTTAGCAAT GTATTGAGGA
4264 CATTTTAGAG TGCCCGTTTT TCACCATTAT AAGCAATGCA ACAATGAACA TCTGTATAAA
4324 TAAATATTCA TTTCTCTCAC CTTTTATTTT CTTAGAATAT ATTCCTAGAA GTAGAATTTT
4384 CCAGAGCCAT GAGGATTTGT GACGCTATTG ATATGTGCCA CTTTGCCTC TCTGTGACAT
4444 ATATAATTAT TTTAATGCA TTCATTTTTT TCTCAGAGTG CATTCGTTTG AAAACATAGA
4504 CGGGAAATAC TGGTAGTCTT CCTTGTCTAGT TAGAAACACC CAAACAATGA AAAATGAAAA

```

Fig. 3 (cont.)

ES 2 356 167 T3

4564 AGTTGCACAA ATAGTCTCTA AAAACAATGA AACTATTGCC TGAGGAATTG AAGTTTAAAA
4624 AGAAGCACAT AAGCAACAAC AAGGATAATC CTAGAAAACC AGTTCTGCTG ACTGGGTGAT
4684 TTCACTTCTC TTTGCTTCCT CATCTGGATT GGAATATTCC TAATACCCCC TCCAGAACTA
4744 TTTCCCTGT TTGTACTAGA CTGTGTATAT CATCTGTGTT TGTACATAGA CATTAACTCTG
4804 CACTTGTGAT CATGGTTTTA GAAATCATCA AGCCTAGGTC ATCACCTTTT AGCTTCTCTGA
4864 GCAATGTGAA ATACAACCTT ATGAGGATCA TCAAATACGA ATTCATCCTG AATGACGCC
4924 TCAATCAAAG TATAATTGCA GCCAATGATC AGTACCTCAC GGCTGCTGCA TTACATAATC
4984 TGGATGAAGC AGGTACATTA AAATGGCACC AGACATTTCT GTCATCCTCC CCTCCTTCA
5044 TTTACTTATT TATTTATTTT AATCTTTCTG CTGCAAAAA ACATACCTCT TCAGAGTTCT
5104 GGGTTGCACA ATTCTTCCAG AATAGCTTGA AGCACAGCAC CCCCATAAAA ATCCCAAGCC
5164 AGGGCAGAAG GTTCAACTAA ATCTGGAAGT TCCACAAGAG AGAAGTTTCC TATCTTTGAG
5224 AGTAAAGSGT TGTGCACAAA GCTAGCTGAT GTACTACCTC TTTGGTTCTT TCAGACATTC
5284 TTACCCTCAA TTTTAAAACT GAGGAACTG TCAGACATAT TAAATGATTT ACTCAGATTT
5344 ACCCAGAAGC CAATGAAGAA CAATCACTCT CCTTTAAAA GTCTGTTGAT CAAACTCACA
5404 AGTAACACCA AACCAGGAAG ATCTTTATTA TCTCTGATAA CATATTTGTG AGGCAAAACC
5464 TCCAATAAGC TACAAATATG GCTTAAAGGA TGAAGTTTAG TGTCCAAAA CTTTTATCAC
5524 ACACATCCAA TTTTCATGGC GGACATGTTT TAGTTTCAAC AGTATACATA TTTTCAAAGG
5584 TCCAGAGAGG CAATTTTGCA ATAAACAAGC AAGACTTTTT CTGATTGGAT GCACTTCAGC
5644 TAACATGCTT TCAACTCTAC ATTTACAAAT TATTTTGTGT TCTATTTTTC TACTTAATAT
5704 TATTTCTGCA ATTTTCCCAA TATTGACATC GTGTATGTAT TTGCCATTTT TAATATCACT
5764 AGACAATTC AATCAGGTTGC TACGTTGGTC CCTTGGGTTT ACTCTAATA GCTTGATTGC
5824 AAATATCTT GTATATATTA TTGTTTTTTC TCCTATCTTG TAATTTCTTT GAGCACATCC
5884 CAAAGAGGAA TGCCTAGATC AATGGGCACA AATAATTTGA CAGCTCTTAT TAAACATTAT
5944 TCTGTAAGTA AAAACTGAAC TACTTTTCAG TACTACTAGC AACATATGAG TGTATCAGCT
6004 TCCTAAACCC CTCCATGTTA GGTCAATTAT AACTTATGAT CTAACAAATT ACAGGGTCTT
6064 ATCCCACTAA TGAAATTATA AGAGATTCAA CACTTATTCA GCCCCGAAGG ATTCATTCAA
6124 CGTAGAAAAT TCTAAGAACA TTAACCAAGT ATTTACCTGC CTAGTGAGTG TGAAGACAT
6184 TGTGAAGGAC ACAAAGATGT ATAGAATTC ATTCCTGACT TCCAGGTATT TACACCATAG
6244 GTGGGGACCT AACTACACAC ACACACACAC ACACACACAC ACCATGCACA
6304 CACAATCTAC ATCAACACTT GATTTTATAC AAATACAATG AATTTACTTT CTTTTGGTT
6364 CTCTCTTCA CCAGTGAAT TTGACATGGG TGCTTATAAG TCATCAAAGG ATGATGCTAA
6424 AATTACCGTG ATTCTAAGAA TCTCAAAAAC TCAATTGTAT GTGACTGCC AAGATGAAGA
6484 CCAACAGTG CTGCTGAAGG TCAGTTGTCC TTTGTCTCCA ACTTACCTTC ATTTACATCT
6544 CATATGTTG TAAATAAGCC CAATAGGCAG ACACCTCTAA CAAGGTGACA CTGTCTCTT
6604 TCCTTCTAC CACAGCCCC ACCTACCCAC CCCACTCCA TTGATTCCAG AGGCGTGCCT
6664 AGGCAGGATC TATGAGAKAA TATAACAGAG AGTAAGAGGA AAATTACCTT CTTTCTTTTT
6724 CCTTCCCTG CCTGACCTTA TTCACCTCC ATCCAGAGC ATCCATTTAT TCCATTGATC
6784 TTTACTGACA TCTATTATCT GACCTACACA ATACTAGACA TTAGGACAAT GTGGCCTGCC
6844 TCCAAGAAC TCAAATAAGC CAACTGAGAT CAGAGAGGAT TAATCACCTG CCAATGGGCA
6904 CAAAGCAACA AGCTGGGAGC CAAGTCCAA AATGGGGCCT GCTGCTTCCA GTTCCCCTCT
6964 CTCTGCATTG ATGTCAGCAT TATCCTTCGT CCCAGTCTG TCTCCACTAC CACTTTCCCC
7024 CTCAAACACA CACACACACA ACAGCCTTAG ATGTTTTCTC CACTGATAAG TAGGTGACTC
7084 AATTTGTAAG TATATAATCC AAGACCTTCT ATTCCCAAGT AGAATTTATG TGCCTGCCCTG
7144 TGCTTTTCTA CCTGGATCAA GTGATGTCTA CAGAGTAGGG CAGTAGCTTC ATTCATGAAC
7204 TCATTCAACA AGCATTATTC ACTGAGAGCC TTGTATTTTT CAGGCATAGT GCCAACAGCA
7264 GTGTGGACAG TGGTGCATCA AAGCCTCTAG TCTCATAGAA CTTAGTCTTC TGGAGGATAT
7324 GGAAAACAGA CAACCCAAAC AACCAACAAA AGAGCAAGAT GCTGCAAAAA AAAAAAAAT
7384 GAATAGGGTG CTAAGATAGA GAAAAGTGGG AGAGTGCTAT TTAGACAAAG TGGTAAAAAC
7444 AAAGCCCTT GTGAGATGAG AGCTGCCGAC AGAGGGGGCG GGTCATGGTT GTGGGTTTTT
7504 GGGTAGGACA TTCAGAGGAG GGGGCGGGTC GTGGTTGTGG GTTTTTGGGT AGGACATTCA

Fig. 3 (cont.)

```

7564 GAGGAGGGGG CGGGTCGTGG TTGTGGGTTT TTGGGTAGGA CATTGAGAGG AGGGGGCGGG
7624 TCGTGGTTGT GGGTTTTTGG GTAGGACATT CAGAGGAGGG GGCGGGTCGT GGTGTGGGT
7684 TTTTGGGACA TTCAGAGGAG TCTGAATGCA CCCAGGCCA CAACTTCAAG ATGGTAAAGG
7744 ACAGCTCCAA GGATCAGAAG AAGCATTCTT GGAAGTGGGG CATTGAGAGA AGGAGGAAAA
7804 ATATGCAGAG ACTAGTGCTT GCAGAGCTTG CATTGAGATT TCATTGAGG TACAATGAAA
7864 ACCCATTAAAT GGGTTTCACA CAGTGAATG GCCTGACCTC ACTTATATTT CCTAAAATAG
7924 AAAACAGATC AGAAGGAAGG CAATAGAGAA GCAGAAAGTC CAATGAGGAG GTTTCACAGC
7984 AGTCATGGGG GTGGGGTAAG GAAAGAAGT GGAAGAAAC AGACAGAATT GGGTTATATT
8044 TTGGAGATAG AACCAACAGA AGGAAGAGGA GAAACAACAT TTAGTGAGAA GGGAAAAAGT
8104 AGGAGAGGAA TAGGTTTGGG AAATAAATCC TGCTGACATT GGAAACCCCA AGGAAGCCTC
8164 AAAAGTATAT TTAGTGTCTT TAGATTTAAA AGAATAGGAA AGAAGCATCT CAACTTGGAA
8224 TTTGAAATCT ATTTTCCAT AAAAGTATTG TTAATTTCTA CTCATACTCA CAAGAAAAGT
8284 ACATTCTAAA GAGTATATTG AAAGAGTTTA CTGATATACT TAGGAATTTT GTGTGTATGT
8344 GTGTGTGTGT ATGTGTGTGT GTGTGTTTAA CCTTCAATG TTGACTTAAA TACTGAGATA
8404 AATGTCATCT AAATGCTAAA TTGATTCCC AAAGGTATGA TTTGTTCACT TGGAGATCAA
8464 AATGTTTAGG GGGCTTAGAA TCACTGTAGT GCTCAGATT GATGCAAAAT GTCTTAGGCC
8524 TATGTTGAAG GCAGGACAGA AACAATGTTT CCCTCCTACC TGCCTGGATA CAGTAAGATA
8584 CTAGTGTAC TGACAATCTT CATAACTAAT TTAGATCTCT CTCCAATCAA CTAAGGAAAT
8644 CAACTCTTAT TAATAGACTG GGCCACACAT CTACTAGGCA TGTAATAAAT GCTTGCTGAA
8704 TGAACAAATG AATGAAGAGC CTATAGCATC ATGTTACAGC CATAGTCTTA AAGTGGTGT
8764 TCTCATGAAG GCCAAATGCT AAGGGATTGA GCTTCAGTCC TTTTCTAAC ATCTTGTCT
8824 CTAACAGAAT TCTCTTCTTT TCTTCATAGG AGATGCCTGA GATACCCAAA ACCATCACAG
8884 GTAGTGAGAC CAACCTCCTC TTCTTCTGGG AAACCTCACGG CACTAAGAAC TATTTACAT
8944 CAGTTGCCCA TCCAAACTTG TTTATPGCCA CAAAGCAAGA CTACTGGGTG TGCTTGGCAG
9004 GGGGGCCACC CTCTATCACT GACTTTCAGA TACTGGAAAA CCAGGCGTAG GTCTGGAGTC
9064 TCACTTGTCT CACTTGTGCA GTGTGACAG TTCATATGTA CCATGTACAT GAAGAAGCTA
9124 AATCCTTTAC TGTTAGTCAT TTGCTGAGCA TGTACTGAGC CTTGTAATTC TAAATGAATG
9184 TTTACTACTCT TTGTAAGAGT GGAACCAACA CTAACATATA ATGTTGTTAT TTAAGAACA
9244 CCCTATATTT TGCATAGTAC CAATCATTTT AATTATTATT CTTCATAACA ATTTTAGGAG
9304 GACCAGAGCT ACTGACTATG GCTACCAAAA AGACTCTACC CATATTACAG ATGGGCAAAAT
9364 TAAGGCATAA GAAAACCTAAG AAATATGCAC AATAGCAGTT GAAACAAGAA GCCACAGACC
9424 TAGGATTTCA TGATTTCAAT TCAACTGTTT GCCTTCTGCT TTTAAGTTGC TGATGAACTC
9484 TTAATCAAAT AGCATAAGTT TCTGGGACCT CAGTTTTATC ATTTTCAAAA TGGAGGGAAT
9544 AATACCTAAG CCTTCTGCC GCAACAGTTT TTTATGCTAA TCAGGGAGGT CATTGTTGTA
9604 AAATACTTCT CGAAGCCGAG CCTCAAGATG AAGGCAAAGC ACGAAATGTT ATTTTTTAAT
9664 TATTATTTAT ATATGTATTT ATAAATATAT TTAAGATAAT TATAATATAC TATATTTATG
9724 GGAACCCCTT CATCCTCTGA GTGTGACCAG GCATCCTCCA CAATAGCAGA CAGTGTTTTC
9784 TGGGATAAGT AAGTTTGATT TCATTAATAC AGGGCATTTT GGTCCAAGTT GTGCTTATCC
9844 CATAGCCAGG AAACCTCTGCA TTCTAGTACT TGGGAGACCT GTAATCATAT AATAAATGTA
9904 CATTAAATAC CTGAGCCAG TAATTGGTCC GATCTTTGAC TCTTTTGCCA TTAACCTTAC
9964 CTGGGCATTC TTGTTTCATT CAATCCACC TGCAATCAAG TCCTACAAGC TAAAATTAGA
10024 TGAACCTAAC TTTGACAACC ATGAGACCAC TGTTATCAA ACTTTCTTTT CTGGAATGTA
10084 ATCAATGTTT CTCTAGGTT CTAATAATG TGATCAGACC ATAATGTTAC ATTATTATCA
10144 ACAATAGTGA TTGATAGAGT GTTATCAGTC ATAACTAAAT AAAGCTTGCA ACAAATTTCT
10204 CTGACACATA GTTATTCATT GCCTTAATCA TTATTTTACT GCATGGTAAT TAGGGACAAA
10264 TGGTAAATGT TTACATAAAT AATTGTATTT AGTGTACTT TATAAAATCA AACCAAGATT
10324 TTATATTTTT TTCTCCTCTT TGTTAGCTGC CAGTATGCAT AAATGGCATT AAGAATGATA
10384 ATATTTCCGG GTTCACTTAA AGCTCATATT ACACATACAC AAAACATGTG TTCCCATCTT
10444 TATACAAACT CACACATACA GAGCTACATT AAAACAACCT AATAGGCCAG GCACGGTGGC
10504 TCAGACCTGT AATCCAGCA CTTTGGGAGG

```

Fig. 3 (cont.)

ES 2 356 167 T3

-1933 AGAAAGAAAG AGAGAGAGAA AGAAAAGAAA GAGGAAGGAA GGAAGGAAGG AAGAAAAGACA
-1873 GGCTCTGAGG AAGGTGGCAG TTCCTACAAC GGGAGAACCA GTGGTTAATT TGCAAAGTGG
-1813 ATCCTGTGGA GGCANNCAGA GGAGTCCCCT AGGCCACCCA GACAGGGCTT TTAGCTATCT
-1753 GCAGGCCAGA CACCAAATTT CAGGAGGGCT CAGTGTTAGG AATGGATTAT GGCTTATCAA
-1693 ATTCACAGGA AACTAACATG TTGAACAGCT TTTAGATTTT CTGTGGAAAA TATAACTTAC
-1633 TAAAGATGGA GTTCTTGTGA CTGACTCCTG ATATCAAGAT ACTGGGAGCC AAATTAAAAA
-1573 TCAGAAGGCT GCTTGGAGAG CAAGTCCATG AAATGCTCTT TTTCCCACAG TAGAACCTAT
-1513 TTCCCTCGTG TCTCAAATAC TTGCACAGAG GCTCACTCCC TTGGATAATG CAGAGCGAGC
-1453 ACGATACCTG GCACATACTA ATTTGAATAA AATGCTGTCA AATTCCCATT CACCCATTCA
-1393 AGCAGCAAAC TCTATCTCAC CTGAATGTAC ATGCCAGGCA CTGTGCTAGA CTGGCTCAA
-1333 AAAGATTTCA GTTTCCTGGA GGAACCAGGA GGGCAAGGTT TCAACTCAGT GCTATAAGAA
-1273 GTGTTACAGG CTGGACACGG TGGCTCAGC CTGTAATCCC AACATTTGGG AGGCCGAGGC
-1213 GGGCAGATCA CAAGGTCAGG AGATCGAGAC CATCCTGGCT AACATGGTGA AACCTGTCT
-1153 CTAATAAAAA TACAAAAAAT TAGCCGGGCG TTGGCGGCAG GTGCCTGTAG TCCAGCTGC
-1093 TGGGGAGGCT GAGGCAGGAG AATGGTGTGA ACCCGGGAGG CGGAACTTGC AGGGGGCCGA
-1033 GATCGTGCCA CTGCACTCCA GCCTGGGCGA CAGAGTGAGA CTCTGTCTCA AAAAAAAAAA
-973 AAAAGTGTTA TGATGCAGAC CTGTCAAAGA GGCAAAGGAG GGTGTTCCCTA CACTCCAGGC
-913 ACTGTTTATA ACCTGGACTC TCATTTCATC TACAAATGGA GGGCTCCCCT GGCAGATCC
-853 CTGGAGCAGG CACTTTGCTG GTGTCTCGGT TAAAGAGAAA CTGATAACTC TTGGTATTAC
-793 CAAGAGATAG AGTCTCAGAT GGATATTCTT ACAGAAACAA TATTCCCCTT TTCAGAGTT
-733 CACCAAAAAA TCATTTTAGG CAGAGCTCAT CTGGCATTGA TCTGGTTCAT CCATGAGATT
-673 GGCTAGGGTA ACAGCACCTG GTCTTGCAAG GTTGTGTGAG CTTATCTCCA GGGTTCACCC
-613 AACTCCGTC AAGGCCTGAA CCCTGCATAC CGTATGTTCT CTGCCCCAGC CAAGAAAGGT
-553 CAATTTTCTC CTCAGAGGCT CCTGCAATTG ACAGAGAGCT CCCGAGGCAG AGAACAGCAC
-493 CCAAGGTAGA GACCCACACC CTCAATACAG ACAGGGAGGG CTATTGGCCC TTCATTGTAC
-433 CCATTTATCC ATCTGTAAGT GGAAGATTTC CTAAACTTAA GTACAAAGAA GTGAATGAAG
-373 AAAAGTATGT GCATGTATAA ATCTGTGTGT CTCCACTTTT GTCCCACATA TACTAAATTT
-313 AAACATTCTT CTAACGTGGG AAAATCCAGT ATTTAATGT GGACATCAAC TGCACAACGA
-253 TTGTACAGAA AACAAATGCAT ATTTGCATGG TGATACATTT GCAAAATGTG TCATAGTTTG
-193 CTACTCCTTG CCTTCCATG AACAGAGAA TTATCTCAGT TTATTAGTCC CCTCCCCTAA
-133 GAAGCTTCCA CCAATACTCT TTTCCCCTTT CCTTTAACTT GATTGTGAAA TCAGGTATTCT
-73 AACAGAGAAA TTTCTCAGCC TCCTACTTCT GCTTTTGAAA GCTATAAAAA CAGCGAGGGA
-13 GAAACTGGCA GATACCAAAC CTCTTCGAGG CACAAGGCAC AACAGGCTGC TCTGGGATTC
48 TCTTCAGCCA ATCTTCATTG CTCAAGTATG ACTTTAATCT TCCTTACAAC TAGGTGCTAA
108 GGGAGTCTCT CTGTCTCTCT GCCTCTTTGT GTGTATGCAT ATTCTCTCTC TCTCTCTCTT
168 TCTTTCTCTG TCTCTCCTCT CCTTCCCTCT TGCTCCTCT CTCAGCTTTT TGCAAAAATG
228 CCAGGTGTAA TATAATGCTT ATGACTCGGG AAATATTCTG GGAATGGATA CTGCTTATCT
288 AACAGCTGAC ACCCTAAAGG TTAGTGTCAA AGCCTCTGCT CCAGCTCTCC TAGCCAATAC
238 ATTGCTAGTT GGGGTTTGGT TTAGCAAATG CTTTTCTCTA GACCCAAAGG ACTTCTCTTT
308 CACACATTCA TTCATTTACT CAGAGATCAT TTTCTTGCAT GACTGCCATG CACTGGATGC
468 TGAGAGAAAT CACACATGAA CGTAGCCGTC ATGGGGAAGT CACTCATTTT CTCCTTTTTA
528 CACAGGTGTC TGAAGCAGCC ATGGCAGAAG TACCTGAGCT CGCCAGTGAA ATGATGGCTT
588 ATTACAGGTC AGTGGAGACG CTGAGACCAG TAACATGAGC AGGTCTCCTC TTTCAAGAGT
648 AGAGTGTAT CTGTGCTTGG AGACCAGATT TTTCCCCTAA ATTGCCTCTT TCAGTGGCAA
708 ACAGGGTGCC AAGTAAATCT GATTTAAAGA CTACTTTCCC ATTACAAGTC CCTCCAGCCT
768 TGGACCTGG AGGCTATCCA GATGTGTTGT TGCAAGGGCT TCCTGCAGAG GCAAAATGGGG
828 AGAAAAGATT CCAAGCCCAC AATACAAGGA ATCCCTTTGC AAAGTGTGGC TTGGAGGGAG
888 AGGGAGAGCT CAGATTTTAG CTGACTCTGC TGGGCTAGAG GTTAGGCCTC AAGATCCAAC
948 AGGGAGCACC AGGGTGCCCA CCTGCCAGGC CTAGAATCTG CCTTCTGGAC TGTCTGCGC

Fig. 4

ES 2 356 167 T3

1008 ATATCACTGT GAAACTTGCC AGGTGTTTCA GGCAGCTTTG AGAGGCAGGC TGTTTGCAGT
 1068 TTCTTATGAA CAGTCAAGTC TTGTACACAG GGAAGGAAAA ATAAACCTGT TTAGAAGACA
 1128 TAATTGAGAC ATGTCCCTGT TTTTATTACA GTGGCAATGA GGATGACTG TTCTTTGAAG
 1188 CTGATGGCCC TAAACAGATG AAGGTAAGAC TATGGGTTTA ACTCCCAACC CAAGGAAGGG
 1248 CTCTAACACA GGGAAAGCTC AAAGAAGGGA GTTCTGGGCC ACTTTGATGC CATGGTATTT
 1308 TGTTTTAGAA AGACTTTAAC CTCTTCCAGT GAGACACAGG CTGCACCCT TGCTGACCTG
 1368 GCCACTTGGT CATCATATCA CCACAGTCAC TCACCTAACGT TGGTGGTGGT GGCCACACTT
 1428 GGTGGTGACA GGGGAGGAGT AGTGATAATG TTCCCATTTT ATAGTAGGAA GACAACCAAG
 1488 TCTTCAACAT AAATTTGATT ATCCTTTTAA GAGATGGATT CAGCCTATGC CAATCACTTG
 1548 AGTTAAACTC TGAAACCAAG AGATGATCTT GAGAATAAC ATATGTCTAC CCCTTTTGAG
 1608 TAGAATAGTT TTTTGCTACC TGGGGTGAAG CTTATAACAA CAAGACATAG ATGATATAAA
 1668 CAAAAAGATG AATTGAGACT TGAAAGAAAA CCATTCACTT GCTGTTTGAC CTTGACAAGT
 1728 CATTTTACCC GCTTTGGACC TCATCTGAAA AATAAAGGGC TGAGCTGGAT GATCTCTGAG
 1788 ATTCCAGCAT CCTGCAACT CCAGTCTGA AATATTTTCA GTTGTAGCTA AGGGCATTTG
 1848 GGCAGCAAAT GGTCATTTTT CAGACTCATC CTTACAAAGA GCCATGTTAT ATTCCTGCTG
 1908 TCCCTTCTGT TTTATATGAT GCTCAGTAGC CTTCTAGGT GCCCAGCCAT CAGCCTAGCT
 1968 AGGTCAGTTG TGCAGGTTGG AGGCAGCCAC TTTTCTCTGG CTTTATTTTA TTCCAGTTTG
 2028 TGATAGCCTC CCCTAGCCTC ATAATCCAGT CCTCAATCTT GTTAAAAACA TATTTCTTTA
 2088 GAAGTTTTAA GACTGGCATA ACTTCTGGC TGCAGCTGTG GGAGGAGCCC ATTTGGCTTGT
 2148 CTGCTGGCC TTTGCCCCCT ATTGCCTCTT CCAGCAGCTT GGCTCTGCTC CAGGCAGGAA
 2208 ATTCCTCTCT GCTCAACTTT CTTTTGTGCA CTTACAGGTC TCTTTAACTG TCTTTCAAGC
 2268 CTTTGAACCA TTATCAGCCT TAAGGCAACC TCAGTGAAGC CTTAATACGG AGCTTCTCTG
 2328 AATAAGAGGA AAGTGTAAC ATTTACAAA AAGTACTCTC ACAGGATTTG CAGAATGCCT
 2388 ATGAGACAGT GTTATGAAAA AGGAAAAAAA AGAACAGTGT AGAAAAATTG AATACCTGCT
 2448 GAGTGAGCAT AGGTGAATGG AAAATGTTAT GGTCACTGTC ATGAAAAAGC AAATCATAGT
 2508 GTGACAGCAT TAGGGATACA AAAAGATATA GAGAAGGTAT ACATGTATGG TGTAGGTGGG
 2568 GCATGTACAA AAAGATGACA AGTAGAATCG GGATTTATTC TAAAGAATAG CCTGTAAGGT
 2628 GTCCAGAAGC CACATTCTAG TCTTGAGTCT GCCTCTACCT GCTGTGTGCC CTTGAGTACA
 2688 CCCTTAACCT CTTGAGCTT CAGAGAGGGA TAATCTTTTT ATTTATTTT ATTTTATTTT
 2748 GTTTTGTTTT GTTTTGTTTT GTTTTATGAG ACAGAGTCTC ACTCTGTTGC CCAGGCTGGA
 2808 GTGAGTGGT ACAATCTTGG CTTACTGCAT CCTCCACCTC CTGAGTTCAA GCGATTCTCC
 2868 TTCCTCAGTC TCCTGAATAG CTAGGATTAC AGGTGCACCC CACCACACC AGCTAATTTT
 2928 TGFATTTTAA GTAGAGAAGG GGTTTCGCCA TGTTCGCCAG GCTGGTTTTG AAGTCTGTAC
 2988 CTAAATGATT CATCCACCTC GGCTTCCCAA AGTGTGGGA TTACAGGCAT GAGCCACCAC
 3048 GCCTGGCCCA GAGAGGGATG ATCTTTAGAA GTCGGGATT CTTTCAAGCC CTTTCTCTCT
 3108 CTCTGAGCTT TCTACTCTCT GATGTCAAAG CATGGTTCCT GGCAGGACCA CCTCACCAGG
 3168 CTCCCTCCCT CGCTCTCTCC GCAGTGCTCC TTCCAGGACC TGGACCTCTG CCCTCTGCAT
 3228 GGCGGCATCC AGCTACGAAT CTCGGACCAC CACTACAGCA AGGGCTTCAG GCAGGCCGCG
 3288 TCAGTTGTTG TGGCCATGGA CAAGCTGAGG AAGATGCTGG TTCCCTGCCC ACAGACCTTC
 3348 CAGGAGAATG ACCTGAGCAC CTTCTTTCCC TTCATCTTTG AAGAAGGTAG TTAGCCAAGA
 3408 GCAGGCAGTA GATCTCCACT TGTGTCTCTT TGGAAATCAT CAAGCCCCAG CCAACTCAAT
 3468 TCCCCAGAG CCAAAGCCCT TTAAGGTAG AAGGCCCAGC GGGGAGACAA AACAAAGAAG
 3528 GCTGGAAACC AAAGCAATCA TCTCTTTAGT GGAAACTATT CTAAAGAAG ATCTTGATGG
 3588 CTAAGTACAT TTGCAACTCC CTCACTCTTT CTCAGGGGCC TTTCACTTAC ATTGTCACCA
 3648 GAGGTTTCGTA ACCTCCCTGT GGGCTAGTGT TATGACCATC ACCATTTTAC CTAAGTAGCT
 3708 CTGTTGCTCG GCCACAGTGA GCAGTAATAG ACCTGAAGCT GGAACCCATG TCTAATAGTG
 3768 TCAGGTCCAG TGTCTTAGC CACCCCACTC CCAGCTTCAT CCCTACTGGT GTTGTATCA
 3828 GACTTTGACC GTATATGCTC AGGTGTCTCT CAAGAAATCA AATTTTGCCA CCTCGCTCA
 3888 CGAGGCCTGC CTTCTGATT TTATACCTAA ACAACATGTG CTCCACATTT CAGAACCCTAT
 3948 CTTCTTCGAC ACATGGGATA ACGAGGCTTA TGTGCACGAT GCACCTGTAC GATCACTGAA

Fig. 4 (cont.)

ES 2 356 167 T3

4008 CTGCACGGCTCGGGACTCACAGCAAAAAG CTTGGTGATG TCTGGTCCAT ATGAACTGAA
4068 AGCTCTCCAC CTCCAGGGAC AGGATATGGA GCAACAAGGT AAATGGAAAC ATCCTGGTTT
4128 CCCTGCCTGG CCTCCTGGCA GCTTGCTAAT TCTCCATGTT TTAAACAAAG TAGAAAGTTA
4188 ATTTAAGGCA AATGATCAAC ACAAGTGAAA AAAAATATTA AAAAGGAATA TACAACTTT
4248 GGTCCTAGAA ATGGCACATT TGATTGCACT GGCCAGTGCA TTTGTTAACA GGAGTGTGAC
4308 CCTGAGAAAT TAGACGGCTC AAGCACTCCC AGGACCATGT CCACCCAAGT CTCTGGGCA
4368 TAGTGCAGTG TCAATTCTTC CACAATATGG GGTCATTTGA TGGACATGGC CTAAGTGCCT
4428 GTGGGTCTC TCTTCCTGTT GTTGAGGCTG AAACAAGAGT GCTGGAGCGA TAATGTGTCC
4488 ATCCCCCTCC CCAGTCTTCC CCCCTTGCCC CAACATCCGT CCCACCCAAT GCCAGGTGCT
4548 TCCTGTAGG GAAATTTTAC CGCCAGCAG GAACTTATAT CTCTCCGCTG TAACGGGCAA
4608 AAGTTCAAG TGCGGTGAAC CCATCATTAG CTGTGGTGAT CTGCCTGGCA TCGTGCCACA
4668 GTAGCCAAAG CCTCTGCACA GGAGTGTGGG CAACTAAGGC TGCTGACTTT GAAGGACAGC
4728 CTCACTCAGG GGGAAAGCTAT TTGCTCTCAG CCAGGCCAAG AAAATCCTGT TTTTGGAA
4788 TCGGGTAGTA AGAGTGATCC CAGGGCCTCC AATTGACACT GCTGTGACTG AGGAAGATCA
4848 AAATGAGTGT CTCTCTTTGG AGCCACTTTC CCAGCTCAGC CTCTCCTCTC CCAGTTTCTT
4908 CCCATGGGCT ACTCTCTGTT CTTGAAACAG TTCTGGTGCC TGATTTCTGG CAGAAGTACA
4968 GCTTCACTC TTTCTTTTCC TTCCACATTG ATCAAGTTGT TCCGCTCCTG TGGATGGGCA
5028 CATTGCCAGC CAGTGACACA ATGGCTTCTT TCCTTCTTTC CTTTCCGATT TAAATGTAG
5088 ACCCTCTTTC ATTCTCCGTT CCTACTGCTA TGAGGCTCTG AGAAACCCTC AGGCCTTTGA
5148 GGGGAAACCC TAAATCAACA AAATGACCCCT GCTATTGTCT GTGAGAAGTC AAGTTATCCT
5208 GTGTCTTAGG CCAAGGAACC TCACTGTGGG TTCCACAGA GGCTACCAAT TACATGTATC
5268 CTACTCTCGG GGCTAGGGGT TGGGGTGACC CTGCATGCTG TGTCCCTAAC CACAAGACCC
5328 CCTTCTTTCT TCAGTGGTGT TCTCCATGTC CTTTGTACAA GGAGAAGAAA GTAATGACAA
5388 AATACCTGTG GCCTTGGGCC TCAAGGAAA GAATCTGTAC CTGTCCGCG TGTTGAAAGA
5448 TGATAAGCCC ACTCTACAGC TGGAGGTAAG TGAATGCTAT GGAATGAAGC CCTTCTCAGC
5508 CTCTGCTAC CACTTATTCC CAGACAATTC ACCTTCTCCC CGCCCCATC CCTAGGAAAA
5568 GCTGGGAACA GGTCTATTTG ACAAGTTTGT CATTAAATGTA AATAAATTTA ACATAATTTT
5628 TAACTGCGTG CAACCTTCAA TCTGCTGCA GAAAATAAA TCATTTTGCC GATGTTATTA
5688 TGCTTACCA TAGTTACAAC CCCAACAGAT TATATATTGT TAGGGCTGCT CTCATTTGAT
5748 AGACACCTTG GGAAATAGAT GACTTAAAGG GTCCCATTAT CACGTCCACT CCACTCCCAA
5808 AATCACCACC ACTATCACCT CCAGCTTCT CAGCAAAAAGC TTCATTTCCA AGTTGATGTC
5868 ATTCTAGGAC CATAAGGAAA AATACAATAA AAAGCCCCTG GAAACTAGGT ACTTCAAGAA
5928 GCTCTAGCTT AATTTTACC CCCCCAAAA AAAAAAATC TCACCTACAT TATGCTCCTC
5988 AGCATTTGGC ACTAAGTTTT AGAAAAGAAG AAGGGCTCTT TTAATAATCA CACAGAAAGT
6048 TGGGGGCCCA GTTACAACCT AGGAGTCTGG CTCCTGATCA TGTGACCTGC TCGTCAGTTT
6108 CCTTCTGGC CAACCCAAAG AACATCTTTC CCATAGGCAT CTTTGTCCCT TGCCCCACAA
6168 AAATTCTTCT TTCTCTTTCG CTGCAGAGTG TAGATCCCAA AAATTACCCA AAGAAGAAGA
6228 TGAAAAGCG ATTTGTCTTC AACAGATAG AAATCAATAA CAAGCTGGAA TTTGAGTCTG
6288 CCAAGTCCC CAACTGGTAC ATCAGCACCT CCAAGCAGA AAACATGCC GTCTTCTGG
6348 GAGGGACCAA AGGCGGCCAG GATATAACTG ACTTACCCT GCAATTTGTG TCTTCTAAA
6408 GAGAGCTGTA CCCAGAGAGT CCTGTGCTGA ATGTGACTC AATCCCTAGG GCTGGCAGAA
6468 AGGGAACAGA AAGGTTTTTG AGTACGGCTA TAGCCTGGAC TTTCTGTGTG TCTACACCAA
6528 TGCCCAACTG CTTGCCTTAG GGTAGTGCTA AGAGGATCTC CTGTCCATCA GCCAGGACAG
6588 TCAGCTCTCT CTTTCAGGG CCAATCCCCA GCCCTTTTGT TGAGCCAGGC CTCTCTCACC
6648 TCTCTACTC ACTTAAAGCC CGCCTGACAG AAACCACGGC CACATTTGGT TCTAAGAAAC
6708 CCTCTGTCAT TCGTCCAC ATTCTGATGA GCAACCGCTT CCCTATTTAT TTATTTATTT
6768 GTTTGTTTGT TTTGATTCAT TGGTCTAATT TATTCAAAGG GGGCAAGAAG TAGCAGTGTG
6828 TGTAAAAGAG CCTAGTTTTT AATAGCTATG GAATCAATC AATTTGGACT GGTGTGCTCT
6888 CTTAAATCA AGTCCTTAA TTAAGACTGA AAATATATAA GCTCAGATTA TTTAAATGGG
6948 AATATTTATA AATGAGCAA TATCATACTG TTCAATGGTT CTGAAATAAA CTTCACTGAA

Fig. 4 (cont.)

```

7008 GAAAAAAAAA AAAGGGTCTC TCCTGATCAT TGACTGTCTG GATTGACACT GACAGTAAGC
7068 AAACAGGCTG TGAGAGTTCT TGGGACTAAG CCCACTCCTC ATTGCTGAGT GCTGCAAGTA
7128 CCTAGAAATA TCCTTGGCCA CCGAAGACTA TCCTCCTCAC CCATCCCCTT TATTTGTTG
7188 TTCAACAGAA GGATATTCAG TGCACATCTG GAACAGGATC AGCTGAAGCA CTGCAGGGAG
7248 TCAGGACTGG TAGTAACAGC TACCATGATT TATCTATCAA TGCACCAAAC ATCTGTTGAG
7308 CAAGCGCTAT GTACTAGGAG CTGGGAGTAC AGAGATGAGA ACAGTCACAA GTCCCTCCTC
7368 AGATAGGAGA GGCAGCTAGT TATAAGCAGA ACAAGGTAAC ATGACAAGTA GAGTAAGATA
7428 GAAGAACGAA GAGGAGTAGC CAGGAAGGAG GGAGGAGAAC GACATAAGAA TCAAGCCTAA
7488 AGGGATAAAC AGAAGATTTC CACACATGGG CTGGGCCAAT TGGGTGTGGG TTACGCCTGT
7548 AATCCCAGCA CTTTGGGTGG CAGGGGCAGA AAGATCGCTT GAGCCAGGA GTTCAAGACC
7608 AGCCTGGGCA ACATAGTGAG ACTCCCATCT CTACAAAAAA TAAATAAATA AATAAAACAA
7668 TCAGCCAGGC ATGCTGGCAT GCACCTGTAG TCCTAGCTAC TTGGGAAGCT GACACTGGAG
7728 GATTGCTTGA GCCCAGAAGT TCAAGACTGC AGTGAGCTTA TCCGTTGACC TGCAGGTCGA
7788 C

```

Fig. 4 (cont.)

ES 2 356 167 T3

```

-5988 GTCGACCTGG AGGTCAACGG ATCTGAGAGG AGAGTAGCTT CTTGTAGATA ACAGTTGGAT
-5928 TATATAACCAT GTCCTGATCC CCTTCATCAT CCAGGAGAGC AGAGGTGGTC ACCCTGATAG
-5868 CAGCAAGCCT GGGGGCTGCA GCTTGGTGGG TAGAGGTACT CAGGGGTACA GATGTCTCCA
-5808 AACCTGTCTT GCTGCCTTAG GGAGCTTCTA ATAAGTTGAT GGATTGGTT AAAATTAACT
-5748 TGGCTACTTG GCAGGACTGG GTCAGTGAGG ACCAACAAAA AGAAGACATC AGATTATACC
-5688 CTGGGGGTTT GTATTTCTTG TGTTCCTTC TCTTCTTGT ACTAAAATAT TTACCCATGA
-5628 CTGGGAAAGA GCAACTGGAG TCTTTGTAGC ATTATCTTAG CAAAAATTTA CAAAGTTTGG
-5568 AAAACAATAT TGCCCATATT GTGTGGTGTG TCCTGTGACA CTCAGGATTC AAGTGTGGC
-5508 CGAAGCCACT AAATGTGAGA TGAAGCCATT ACAAGGCAGT GTGCACATCT GTCCACCCAA
-5448 GCTGGATGCC AACATTTTAC AAATAGTGCT TGCCTGACAC AAATGCAGTT CCAGGAGGCC
-5388 CAAATGAAAA TGTTTGTACT GAAATTTGTT AAAGCTTCCC GACAAACTAG ATTTATCAGT
-5328 AAGGATTGTT TTCTGCAAGG GGGATGAAAC TTGTGGGGTG AGCCATTTGG GCTGAGGAGG
-5268 AGGGAGGTTG GAGCTGAGAA ATGTGGAGAC AATTTCCCTT TAGAAGGACT GAATCTCCCT
-5208 GCCTCTCTGG GGTGCGGCAG CCAGCAGGAT CCAATGGTGT ATATGTCTCC CCAGCTCCCC
-5148 ATTCAGTGAT ATCATGTCAG TAGCTTGAAT TTATCCGTGG TGGGAGTATT ATGTCATGGA
-5088 AATTGGCAAA TGGAAACTTT TATTGGAGAT TCAATTGTTA AACTTTTACC AGCACAACAC
-5028 TGCCCTGCCT TCAGAGTCAA TGACCCTATC CAAGTTTAAAT CCATCTGTCC ACTGTCTCCA
-4968 ACACGATCTT TATAAAACAC ACCTGACAAC ATTACCCTTT TATTCAGTTT TTTAAAAGAT
-4908 AAGTTTCCAG CTCATCGGGG TGGCTTTAAA GGCCATTTCT CCTCTGGACC TCACCCAACT
-4848 TTTCAAATCA CTTTCCCTAC CCCTACCTCT AAATGCTACT CAAACTCCAG CCATCCTGAA
-4788 TAATAAGACT TTTGAAAAGT AGATTATGGG CTGGGCACAG TGGCTCACAC CTGTAATCCC
-4728 AGCACTTGG GAGGCCAAGA TGGGTGGATC ACCTGAGGTC GGGAGTTCGA GACCAGCCTG
-4668 ACTAACATAG TGAAACCTTG TCTCTACTAA AAATACAAAA TTAGTTGGGG GTGGTGGCAC
-4608 AAGCCTGTAA TCCCAGCTAC TCAGGAGGTT GAGGCAGGGG AATTGCTTGA ACCTGGGAGG
-4548 CGGAGGTTGC GGTGAGCCTA GATTGCTCCA CTGCACCTCA GCCTGGGCAA CAAGAGCGAA
-4488 ACTCCATGTC AAAAAAATAA ATAAATAAAT AAAGTAGATT ACATCAGATA CCTCTGGCCT
-4428 AGGTGTGTTA TGACCAACTC TCCTGCTGAG AATAACTAGA AAAGCTAGAC AAAACATATT
-4368 TCCAAAAGAT CTCTTTGGAG GCATCAGAGA ATGGCCAAGG CTGTAAGGAA CTGCCCTGAGC
-4308 CCAGAGAGGT GGAGCCCAGC ACTGGTGCCC TTTACTCCTG GGGACATGTG CTGGTTTCAA
-4248 AAACCTCAGC TGAGCTTTTG AGCATTTCATG GAACCTGGTG GGGGAGATGA AATTTGTACC
-4188 TTAATCCTG CCTACAGGGA GGGTCCCTGA TAATCCCCAC CCAATTTGGA AATCTGGGTC
-4128 AGCCTTCACA GGTACTGAAG CCCTCCTCTG AATGATCTCA AGTCCTGCTA GGGTAGAGGT
-4068 TACCTGCTTT TGAAAGGCTC CTGGCCTACC TGTGCAGCAG GAGCAAAAGT GAACCATCTC
-4008 AGGGTACAGA TAACAATCAT CCAGAGCCTT GAATGACCTC TACTGTGCTT AATATATAGT
-3948 ATTCAGCAGT CAGTAAAAG GATTTAGGCA CATGCAAGAT GACCTGTGTA TCAGGGAGAA
-3888 ATAGGCAATA AATTGAGATC CAGCAGGGAT TTGAATCATG GATTTGAATC AGGGGCAGCC
-3828 TTCGAAAGAA CTATGGAGAA TATACTCAGA TTTAAAACAT AAGATTGGAA TTTTTGGCAG
-3768 AGAACTAACA ACTGTACAAA AAAGGAACCA AATGGAAATC CTAGAACTGA AAGATGCAAT
-3708 TAACCGATGT TGAGAAATAG CCAACATCTA TTGAACACTT CCCATGTGGA CAGCTGTGCT
-3648 AAACACTTTA CAGGCATCAA CATAAGATGT GTCCCCTTAC AGCAGTGCAG TGTCCCTCCT
-3588 AAGACATGGA CAGCCTGGTT TCCCTATCTC TCTGCTTCAT CAAAACCCCT TTACGTGGGG
-3528 CTTAGACACT CTTGTTGTCT CTAGTGCTA GTAGCACAGG GCTCAGCACA TGGGAAGCCAC
-3468 TAGATACAAT TTGATGACCA GGACCTCCGA TGAAAGCCAT GGGTGTGAT TGGGAAGGCA
-3408 TTGTCTTTTA TGTGCTATGG TCTTAAAGCT TCATCCAGGA AGCAGAATC GGGGGGTGCT
-3348 GAGGACCCAG AACCGAGAAT AAGATTAGTC AGAGATTTC TGTGGGCAGA AATCATAAGG
-3288 ACGCCAAC TG TTTGGGTGAG ATAAGACGAA ACCAAGAGTG GACTTGTGGC CAGAAGCGTG
-3228 AGGAAGAGGG AGAGAGCTTC CCTTGTCCCC TTTCTTCTC TCCCTAAGCC ACAGTGATTG
-3168 ACAGCCCCC CGCTTGGAG TCAGAGCAGG CTTGAGACTG GACTGGGAAA GGAGGGTGGG
-3108 TCAGGATACA GAGCAGGAAG GCTGGGAGTG CAGGGCAGGA GCAAGGGGCT GGGGCATTCA
-3048 TTGTGCCTGA TCTCTCCAC TTTACCTGGG GTAAAGAAGC ATATGCAAAA GCCACGGTGT

```

Fig. 5

-2988 GAGTATTTCC CAAGTGCAG GGTGAGGCA TGATTCATCA CGTGCAGCAT TTCATTCAAT
 -2928 CCTTATAGTA ACCGATGATG TGGCTTCTAT TATTAGCTCT ATCAGATAAT GAAACTGAGA
 -2868 CCAAGACAGG CTCTGCACAT TGTGTGGGGT AATGACACAG GGGGATTCAG ACCTAGACTC
 -2808 CATAACTCCT GCCCCAGGGA CCACCCCCAC CCTCACCCCTG TGCATGTGGA CAAAGGACAG
 -2748 ACTGGGCCAC TTCTCAGGAC ACAGCGGGGA AATGACACAG AGCAGGGAGG TTCCAGGAGC
 -2688 CCCGAGCGTC TTTTCTCCAG GAGAATACTC TCTGAATTCA GACTGGGGTC AGAGAAACAT
 -2628 TTACCCAGGA GCCGCAGTGT GGGTGGGGCT TTTTACTTGA AACGCTGTCT GAAGGCAGTG
 -2568 GCAGGATGAA CTCTCCACCC TACCTTGGCA AGCCACTTCT CTTCTGCAAT CTGTAAGGAC
 -2508 ATTGTTGAGA GAATTATGGT CTTCCAAATC CGGAGGGTTG AAGAAAGACA AATAGGAGAG
 -2448 AACCTATCAT AGTCAGGTGC TAGCTGCCTT CTCTTTCAGA GAGTGTGAGA ATAAAGTGAT
 -2388 ACACTTGATT ATTAGCAAACT ACTTTGGAAA TTTTAAACGC TAATATTCAA CACACTCTGG
 -2328 AAGAGGCAAA TAAGTAGACA GGTTCATATA CATCATCTCC TTCAGCTAGT CCTCACAAAA
 -2268 ACAAACAAT GAATAACAA AATTCTTCTT TGGCCCTCAT AGGAAGACAC TGTTTCTTGA
 -2208 ACGTGTTC AAAAAAGGATGG GTGACTCACT CAAGGTCACA CTGTTTATGA GGACAGTACA
 -2148 GGAATACAGA CATGCCATTT TGCCTGAAAA AATCCATCAC CCAGGGAGGT GACACAATTT
 -2088 TGCAGAAATG TTCTATTTCC TCTGAAGGAT ACATTCTTTA AACCTTTGGG AAATTCATTC
 -2028 ATAGTCTTCC TCCTTTGAAG GATTACTCTC TGGACACAAA GTGTTTGATT CTGATTTGTT
 -1968 GGTTGGAAGA TGTGTTGGTT GAGAGAAAGA TTCTGATTTG TTGGTTGAAA ATAGACTCAT
 -1908 CAAGATCAAC TGCTGTAGTA GTAAATATTT TGACATTTTG TCTGTATTCC TGTGCTGCCC
 -1848 TCACAAGCTG CATCACCTTG AGTGAGTCAT TCATACTTTT TTGTTTGT TTGTTTGGGA
 -1788 GATGGAGTCT TACTCTGTTG CCTAGGCTGG AGTGCGGTGG CGTGATCTTG GCTCACTGGG
 -1728 ACCTCCATCT CCTGGGTTCA AGTGATCCTC CTGCTCAGC CTCCCGAGTA GCTGGGATTA
 -1668 CAGGCACATG CCACCATCCC TGCTAATTTT TGCATTTTCA GTAGAGACGG AGTTTCACCA
 -1608 TGTGTTGTCAG GTTGGTCTTG AACTCCTGAC CTCAGGTGAT CCGCCACCT CAGCCTCCCC
 -1548 AAGTGTGTTG ATTACAGGTG TGAGCCACCG TGCCAGCCC AGCCATCATT TTTGAAACAC
 -1488 GTTTGAGAAA TAGTGTCTTC CTTTGAGGGC CAAGGAGACA TTTTTTTGT TATTGTCTTT
 -1428 GTTTTGTGTA GGACTAGCTG AAGGGGTGA TGTATATTAA CCTGCCTACT TATTGTCTTC
 -1368 TTCCAGAGT GTGATGAATA TTAGGGTTTA AAGTTTCTGA AGCATTTGTT AATAAAGCCC
 -1308 GGGGCTGGAG GTCAGAAGAC CTGGATTTCT CTGCATACTT TTGCCATCAG CAAGCTGTGT
 -1248 GGCCTTGGAC AGATCCCTTT TTTGTCTAAA TCTTCTGAG TCTTCTGAA AACAAATGCCA
 -1188 GGTTGGGACA GGATGATTGC CAAGCTCCCG TCCAGCTCTA AAACACTGCA ACGTATGCTT
 -1128 CTGCACCAGC ACTGTCCATC CTGTAGATCA TGCAGAAATT CTCTTCAACT TTTTCTTACC
 -1068 CATAAAATAG GAGCATGCTT ACCTTTTTCC TAATGTTCCA GGCCCGGGT CTAGATATTTG
 -1008 TAAGTAAGGA AGTTAATGTG TATCAGAGCC CATTATGGGC CAGAAGTCT CCTCTTCCCT
 -948 CCTACACCTG CTTCCTCCCT CCTCCTCC CTCTTCCCT TCCTTCCCTC CATCCATTTG
 -888 TGAAGAAGAC ATGATCACCC TCATTCTGAG AGTGAAGAGA CAGAGGCTCA ACTAATGAAA
 -828 TGATTTGTTT AAGGTCACAC GGGTGGCACA AGGCAAGTGG CAGAGGTTGA ATTTAGACCC
 -768 ATTCTGTGCC AAATGCTGAG TTTATGTCAT CGTCCCGAGA CCATAACTTT AAAGATGTAA
 -708 GATAGTGGGA AAAGAGTTGA TTTCAAAGCA CTTCTCAGAA GGACTCACTT TACATCAGGG
 -648 GTCAGCAGAC TCAGGCCAAA TCCGGTCCAT TCCCGCTTT TGCAAAAGAAA GTTGTAGTGG
 -588 AACACAGCTA GGCTTATGTA TTTATGGATT GCCAACGTCC TTTTGTGAAA CAGACAGCTG
 -528 AGCTGAGTAA TCGTGGCGCA CAAAACCTAA AATATTTACT ATCTCGTCTT TTACAGAATG
 -468 TTTGCCAATC TATGGTCCCG AGTCCAAGGC TGTCCATTTT TCAAAGAACA CAAAGTGACA
 -408 TGAGACTGTC CCATGTGCAG GGAGCCCTAT CATTTTATTA TGARAAAACG GCCTTTCTGC
 -348 TCAAACTGTG TTTTAAAAA GTCAACAAAC AGACTCTGGG TACCTGTGAG GAACAGTAGG
 -288 GAGTTTGGTT TCCATTGTGC TCTTCTTCCC AGGAACTCAA TGAAGGGGAA ATAGAAATCT
 -228 TAATTTGGG GAAATGCAC AGGGGAAAAA GGGGAGGGAA TCAGTTACAA CACTCCATTTG
 -168 CGACACTTAG TGGGGTTGAA AGTGACAACA GCAAGGGTTT CTCTTTTGG AAATGCGAGG
 -108 AGGGTATTTT CGCTTCTCGC AGTGGGGCAG GGTGGCAGAC GCCTAGCTTG GGTGAGTGAC
 -48 TATTTCTTTA TAAACCACAA CTCTGGGCCC GCAATGGCAG TCCACTGCTT GCTGCAGTCA

Fig. 5 (cont.)

ES 2 356 167 T3

13 CAGAAATGGAA ATCTGCAGAG GCCTCCGCAG TCACCTAATC ACTCTCCTCC TCTTCCTGTT
 73 CCATTCAGAG ACGATCTGCC GACCCTCTGG GAGAAAATCC AGCAAGATGC AAGCCTTCAG
 133 GTAAGGCTAC CCCAAGGAGG AGAAGGTGAG GGTGGATCAG CTGGAGACTG GAAACATATC
 193 ACAGCTGCCA GGGCTGCCAG GCCAGAGGGC CTGAGAATG GGTGGGGCT GGAGAGGATG
 253 TCCATTATTC AAGAAAGAGG CTGTTACATG CATGGGCTTC AGGACTTGTG TTTCAAATA
 313 TCCCAGATGT GGATAGTGGC ACCGAGGGC TGTCTTACTT TCCCAGAGAC TCAGGAACCC
 373 AGTGAGTAAT AGATGCATGC CAAGGAGTGG GACTGCGATT CAGGCCTAGT TGAATGTGCT
 433 GACAGAGAAG CAGAGAGGGG CACCAGGGG ACAGCCCGAA GGCCAGACT GATATGGGCA
 493 AGGCCTGTCT GTGCTGACAT GTCGGAGGGT CCCACTCTCC AGGGACCTTG GTTCCCCCGT
 553 CTGTGACATC TGTGACATGA GAGTCACGAT AACTCCTTGT GTGCCTTACA GGGTGTGTTG
 613 GAAAATTAAA TGCACAGATA ATAGCGTAAC AGTATCCGT GCATTGTAAA GAGCCTGAAA
 673 ACCATTATGA TTTGAAAATG GAATCGGCTT TGTGAGACCA TCACTATTGT AAAGATGTGA
 733 TGCTGATAGA AATGACAGGA CTGCTTGTGC ATGCCCTCTG CAGTGTGACA TTCCAGCAGT
 793 GAAATCATGT TGGGGTACT TCTCCCCAC TCTGACCTT ATGTTTGTCT GGGCCGAGGC
 853 TGCAAGTCGG GCTCTGTGGG TGTATGAGTG ACAAGTCTCT CCCTCCAGA TATGGGGACT
 913 GTCTGCTTCC CTAGGTTGCC TCTCCCTGCT CTGATCAGCT AGAAGCTCCA GGAGATCCTC
 973 CTGGAGGCCC CAGCAGGTGA TGTTTATCCC TCCAGACTGA GGCTAAATCT AGAACTAGG
 1033 ATAATCACAA ACAGGCCAAT GCTGCCATAT GCAAAGCACT TTGGTTTGCC TGGCCACCCC
 1093 TCGTCCAGCA TGTGGGCTCT TCAGAGCACC TGATGAGGTG GGTACAGTTA GCCACACTTC
 1153 ACAGGTGAAG AGGTGAGGCA CAGGTCCCAG GTCAGGCTGG CCGGAGCTCT GTTTATTACG
 1213 TCTCACAGCT TTGAGTCCTG CTCTCAACCA GAGAGGCCCT TTACCAAGAA GAAAGGATTG
 1273 GGACCCAGAA TCAGGTCACT GGCTGAGGTA GAGAGGAAGC CGGGTTGTTT CCAAGGGTAG
 1333 CTGCTCCTGC AGGACTCTGA GCAGGTCACC AGCTAATGGA GGAAAGGCTC TAGGGAAAGA
 1393 CCCTTCTGGT CTCAGACTCA GAGCGAGTTA GCTGCAAGGT GTTCCGTCTC TTGAAACTTC
 1453 TACCTAGGTG CTATGGTAGC CACTAGTCTC AGGTGGCTAT TTAAATTTAT ACTTAAATGA
 1513 ATGAAAATAG AAGAAAATTT AAAATCCAGA CCCTTGGTCA CACTATCCAC ATTTAAAGAG
 1573 GTCAATAGCC ACATGTGGTT AGTGGCCACC CTATTGGGCA GTGCAGCTAC AGAACATTTT
 1633 TGCATCCAG AAAGTCTTT TGGATGTTGC TGCTCTACAG CATGCTTTGC TGAACAGAA
 1693 GTGCCTTCCC TGGGAATCTC AGATGGGAAG CAAGTAAGGA GGGGAGTCAA ATGTGGGCTC
 1753 ACTGCTCACC AGCTGTGAGG GTTGGGCTG CCTCTAACC ATTGTCAGCC TCAGTCTTCT
 1813 CATCCATGCA TGCCGTGGGT ATAATAAAAT ACTATACCCC TGGAAAGAGCT GGATGCAAAT
 1873 TTGACAAGTT CTGGGGGACA CAGGAAGGTG CCAAGCACA GGCTGGGCAC ATGGTGGCTG
 1933 TGCCTACAG CTGAGTCTTT TTCCTTTTCA GAATCTGGGA TGTTAACCAG AAGACCTTCT
 1993 ATCTGAGGAA CAACCAACTA GTTGTGGAT ACTTGCAAG ACCAAATGTC AATTTAGAAG
 2053 GTGAGTGTT GCCAGGAAAG CCAATGTATC TGGGCATCAC GTCACCTTGC CCGTCTGTCT
 2113 GCAGCAGCAT GGCCTGCCTG CACAAACCCT AGGTGCAATG TCCTAATCCT TGTTGGGCTC
 2173 TTGTATTCAA GTTGAAGCT GGGAGGGCCT GGCTACTGAA GGGCACATAT GAGGGTAGCC
 2233 TGAAGAGGGT GTGGAGAGGT AGAGTCTAGG TCAGAGGTCA GTGCCTATAG GCAAGTGGTC
 2293 CCAGGGCCAC AGCTGGGAAG GGCAAATACC AGAAGGCAAG GTTGACCATT CCCTCCTCA
 2353 AGTGCCTATT AAGGCTCCAT GTTCTATGT TGTCAAACC CTAACCTCAAT CCCAAATTA
 2413 TCCACCATGT ATAAGGTTGA GCTATGTCTC TTATCCTGG ACACCATACT CAGCCATATC
 2473 TGGTCCACAC ATTAACAGCT GGATGACCTT GAAGAAGCTT CACCCACTCT GTTCTCAGC
 2533 TTTCCCTTCA GTGGGATGAT ATCAACTGGA CAACAGGATG TGCGATTCTT TTAGTTCAG
 2593 CCTTCCAGGA TGTTTTCACT CCCCTGTTG TTGTTGTAGG ATGGTATTAC CTCCACCTC
 2653 CCACCTTCCC TATGCCCTGG TTCTGTCTCC TGTGCTCGC TCTGAAAGTG GATGAGACCT
 2713 ACAATTCCTG TCCTGGTAGT TCTCCTAATG AACACACTGA AGCACGAGGA AGCTGAGATT
 2773 TTTGTTGCTA CATGAGAGCA TGGAGGCCTC TTAGGGAGAG AGGAGGTTCA GAGACTCCTA
 2833 GGCTCCTGGT GGAGCCCCAC TCATGGCCTT GTTCATTTT CCTGCCCTC AGCAACTC
 2893 CTATTGACCT GGAGCACAGG TATCCTGGGG AAAGTGAGGG AAATATGGAC ATCACATGGA

Fig. 5 (cont.)

```

2953 ACAACATCCA GGAGACTCAG GCCTCTAGGA GTAAC TGGGT AGTGTGCATC CTGGGAAAG
3013 TGAGGGAAAT ATGGACATCA CATGGAACAA CATCCAGGAG ACTCAGGCCT CTAGGAGTAA
3073 CTGGGTAGTG TGCATCCTGG GGAAAGTGAG GGAAATATGG ACATCACATG GAACAACATC
3133 CAGGAGACTC AGGCCTCTAG GAGTAACTGG GTAGTGTGCA TCCTGGGGAA AGTGAGGGAA
3193 ATATGGACAT CACATGGAAC AACATCCAGG AGACTCAGGC CTCTAGGAGT AACTGGGTAG
3253 TGTGCTTGGT TTAATCTTCT ATTTACCTGC AGACCAGGAA GATGAGACCT CTCTGCCCTT
3313 CTGACCTCGG GATTTTAGTT TTGTGGGGAC CAGGGGAGAT AGAAAAATAC CCGGGGTCTC
3373 TTCATTATG CTGCTTCCTC TTCTATTAAC CTGACCCCTC CCTCTGTTCT TCCCCAGAAA
3433 AGATAGATGT GGTACCCATT GAGCCTCATG CTCTGTTCTT GGAATCCAT GGAGGGAAGA
3493 TGTGCCTGTC CTGTGTCAAG TCTGGTGATG AGACCAGACT CCAGCTGGAG GTAAAAACAT
3553 GCTTTGGATC TCAAATCACC CCAAAACCCA GTGGCTTCAA ACAACCAAAA TTTTCTTA
3613 TGATTCTGTG GGTGACCAG GATTAGCTGG GTAGTCTGT TCCATGTGGT GGAACATGCT
3673 GGGGTCACCT TGAAGCTGC ATTCAGCAGA GTGCCAGGCT TCGCTGGGC ATCCAAGGTG
3733 GTCCTCATC CTCCAGGCTC TCTTCCATG TGATCTCTCA GTGTTAAGA GTTAGTTGGA
3793 GCTTCCTTAC AGCATGGCGG CTGACTTCCA AAAGGGATTA TTCCAAAAG AGCCTCAACA
3853 TGCAGGCGCT TATTATGACT TCTGCTTGA TCATCCTATT GGCCAAAGCC AGTCACGTGG
3913 CTAAGTCTAG CCCCCTGTGA GAGGAGACTG CATAAGAGTG TGAACACCAG GAGACACGCT
3973 CACTGGGGC CACCCTGTA ACCATCTACC ACAGGACCTG AATCTCTGTG TGCTACTCCC
4033 TTGCTCAAGG GCCCCCCTAC CCACGCAGAC CTGCTGTCTT CTAGCAAAGC CCATCCTCAG
4093 GACCTTTCTC TTCCAATCCT TATTGACTCA AATTGATTAG TTGGTGCTCC ACCCAGAGCC
4153 CTGTGCTCCT TTATCTCATG TAATGTTAAT GGGTTTCCCA GCCCTGGGAA AACATGGCTT
4213 TGTCTCAGGG GCTTGCTGGA TGCAACCTTA ACCTCAATGT GAGTGGCCAT ACTGTGGCAC
4273 TGTCCCATCC CTCACCAGGG ACACTGTTCT GGAGGGTGAC TGCCTGTTCT GTGAGGAGTG
4333 GGGATGGCTA GGACATTGCA TGGAACACAC CACCACCCA TCTTCTCAGA GCTCAAACCC
4393 TGACAGAACA CCAGCTCCAC AGGCCTTGGC TTCTGCTGAT GGTGCCGTGT ATTTACCAGA
4453 CTTAGTGGTC CAAGGCCAGA GTGGCAGATT TCCCAAAGTC AAGGTGTGAC AGTGGGACAG
4513 CCTCTTTGTG TCTTTGCTGT CCTAAGAAAC CTGGGCCAGG CCAGGCGCAG TGGCTCACGC
4573 CTTGTAATCC CAGCACTTTG AGAGGCCAAG GTGGGCAGAT CACGAGGTCA GGAGTTGAG
4633 ACCAGCCTGG CCAACATTGG TGAAACCCTG TCTCTATTAA AAATAGAAA CATTAGACAG
4693 GTGTGGTGGT GCATGCCTGT AATCCCAGCT ACTCAGGAGG CTGAGGCAGG AGAATCGCTT
4753 GAACCCAGGA GGTGGAGGTT GCAGTGAGCC GAGATTGTGC CACTGCCTC CAGCCTAGGC
4813 GACAGAGCAA GACTCCGTCT CGGGAAAATT AATTAATAAA TAAATAAAC TAGGTCCCAG
4873 AGTCCCACAG AATGGCAGAC AGGAGCACCT GGGGGCTTTT AGGGTATGGC ATTTCCCCTG
4933 TACTAACTCT GGGCTGTCCA GAGGCGATT CATGGCGTGG AGTGGAGAGG GAGGCAGCAC
4993 AGGACTTCTT AGGCCTCAGC TCTCACCTGC CCATCTTTG ATTTCCAGGC AGTTAACATC
5053 ACTGACCTGA GCGAGAACAG AAAGCAGGAC AAGCGCTTCG CCTTCATCCG CTCAGACAGT
5113 GGCCCACCA CCAGTTTGA GTCTGCCGCC TGCCCCGGTT GGTTCCTCG CACAGCGATG
5173 GAAGCTGACC AGCCCGTCAG CCTCACCAAT ATGCCTGACG AAGGCGTCAT GGTACCCAAA
5233 TTCTACTTCC AGGAGGACGA GTAGTACTGC CCAGGCCTGC CTGTTCCCAT TCTTGCATGG
5293 CAAGGACTGC AGGGACTGCC AGTCCCCTG CCCCAGGGCT CCCGGCTATG GGGGCACTGA
5353 GGACCAGCCA TTGAGGGGTG GACCCTCAGA AGGCGTCACA ACAACCTGGT CACAGGACTC
5413 TGCTCTCTCT TCAACTGACC AGCCTCCATG CTGCCTCCAG AATGGTCTTT CTAATGTGTG
5473 AATCAGAGCA CAGCAGCCCC TGCACAAAGC CCTTCCATGT CGCCTCTGCA TTCAGGATCA
5533 AACCCCGACC ACCTGCCCAA CCTGCTCTCC TCTTGCCACT GCCTCTTCTT CCCTCATTCC
5593 ACCTTCCCAT GCCCTGGATC CATCAGGCCA CTTGATGACC CCCAACCAAG TGGCTCCCAC
5653 ACCCTGTTTT ACAAAAAGA AAAGACCAGT CCATGAGGGA GGTTTTTAAG GGTTTGTGGA
5713 AAATGAAAAT TAGGATTTCA TGATTTTTTT TTTTCAGTCC CCGTGAAGGA GAGCCCTTCA
5773 TTTGGAGATT ATGTTCTTTC GGGGAGAGGC TGAGGACTTA AAATATCTCT GCATTTGTGA
5833 AATGATGGTG AAAGTAAAGT GTAGCTTTTC CCTTCTTTTT CTCTTTTTTT TGTGATGTC
5893 CAACTTGTA AAATTAAGA TTATGGTACT ATGTTAGCCC CATAATTTTT TTTTCTTTT

```

Fig. 5 (cont.)

```

"5953" TAAACACTT CCATAATCTG GACTCCTCTG TCCAGGCACT GCTGCCCAGC CTCCAAGCTC
6013 CATCTCCACT CCAGATTTTT TACAGCTGCC TGCAGTACTT TACCTCCTAT CAGAAGTTTC
6073 TCAGCTCCCA AGGCTCTGAG CAAATGTGGC TCCTGGGGGT TCTTCTTCC TCTGCTGAAG
6133 GAATAAATTG CTCCTTGACA TTGTAGAGCT TCTGGCACTT GGAGACTTGT ATGAAAGATG
6193 GCTGTGCCTC TGCCTGTCTC CCCACCAGGC TGGGAGCTCT GCAGAGCAGG AAACATGACT
6253 CGTATATGTC TCAGGTCCCT GCAGGGCCAA GCACCTAGCC TCGCTCTTGG CAGGTACTCA
6313 GCGAATGAAT GCTGTATATG TTGGGTGCAA AGTTCCTAC TTCCTGTGAC TTCAGCTCTG
6373 TTTTACAATA AAATCTTGAA AATGCCTATA TTGTTGACTA TGTCTTGGC CTTGACAGGC
6433 TTTGGGTATA GAGTGCTGAG GAAACTGAAA GACCAATGTG TYTTYCTTAC CCCAGAGGCT
6493 GCGCCTGGC CTCTTCTCTG AGAGTTCTTT TCTTCCTTCA GCCTCACTCT CCCTGGATAA
6553 CATGAGAGCA AATCTCTCTG CGGGG

```

Fig. 5 (cont.)

Genotipos de construcciones de SNP de IL1B

Construcciones/Nº del SNP	1	2	3	4	5	6	7	8	9	10	14	15	17
pGL3-IL1BS	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B1C	C	T	G	C	A	T	A	G	G	G	C	T	C
pGL3-IL1BS3KM	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B2C	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B3A	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B4T	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B5G	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B6C	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B7G	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B8A	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B9A	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B10C	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B14T	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B15C	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL3-IL1BL3KM	T	T	G	C	A	T	A	G	G	G	C	T	C
pGL-B17T	T	T	G	C	A	T	A	G	G	G	C	T	C

Figura 6

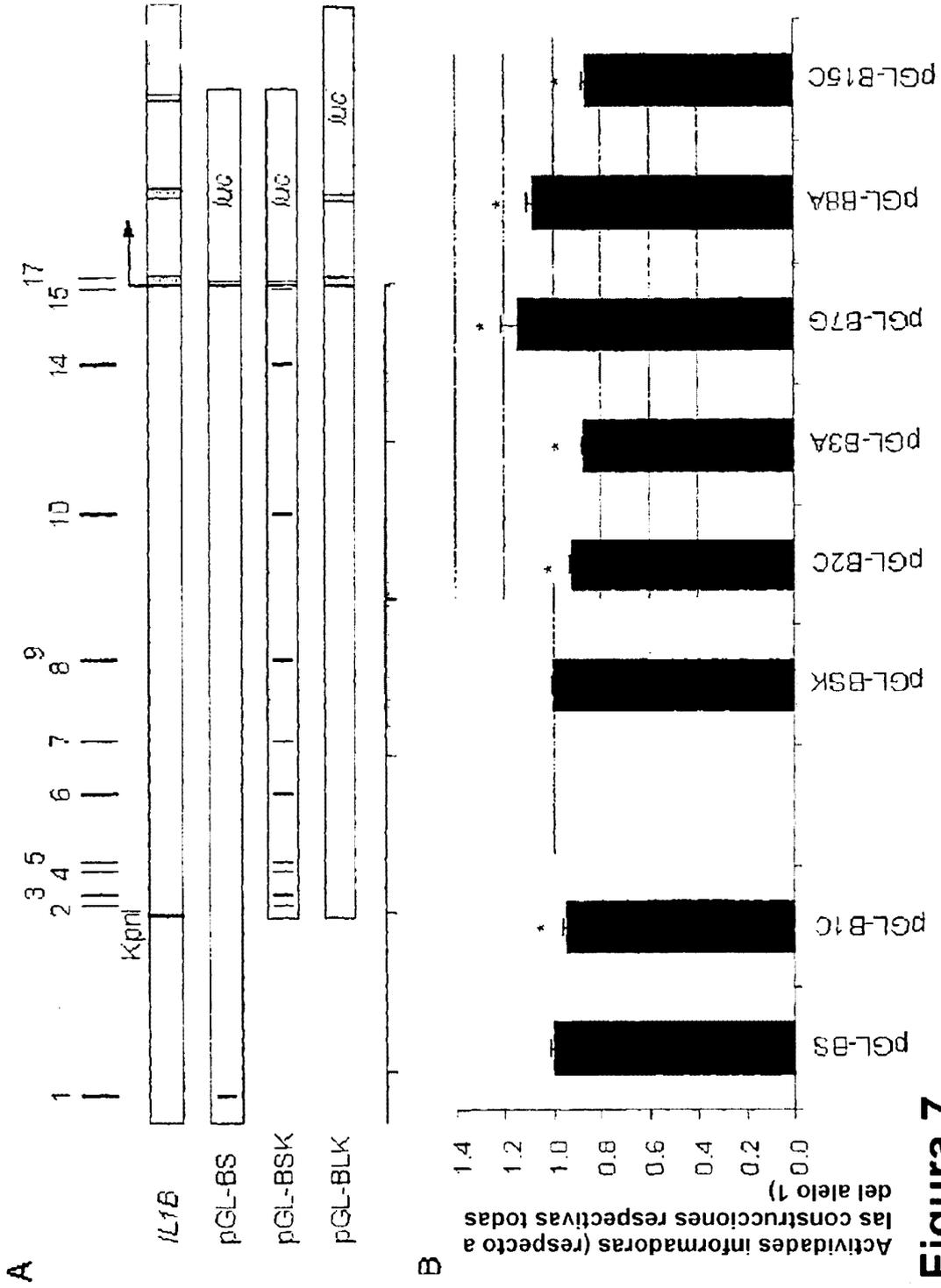


Figura 7 Construcciones de SNP para el promotor de IL1B

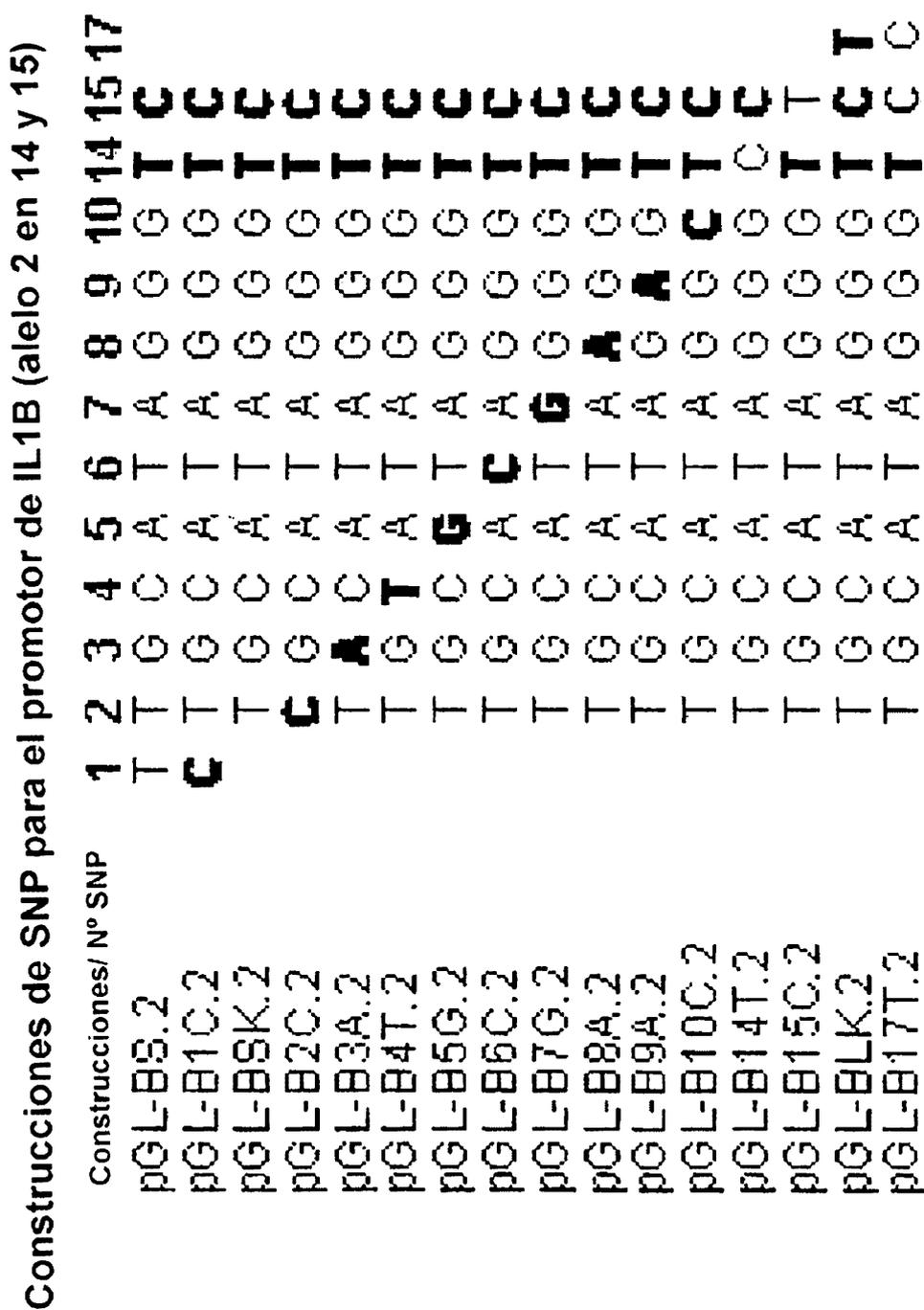


Figura 8

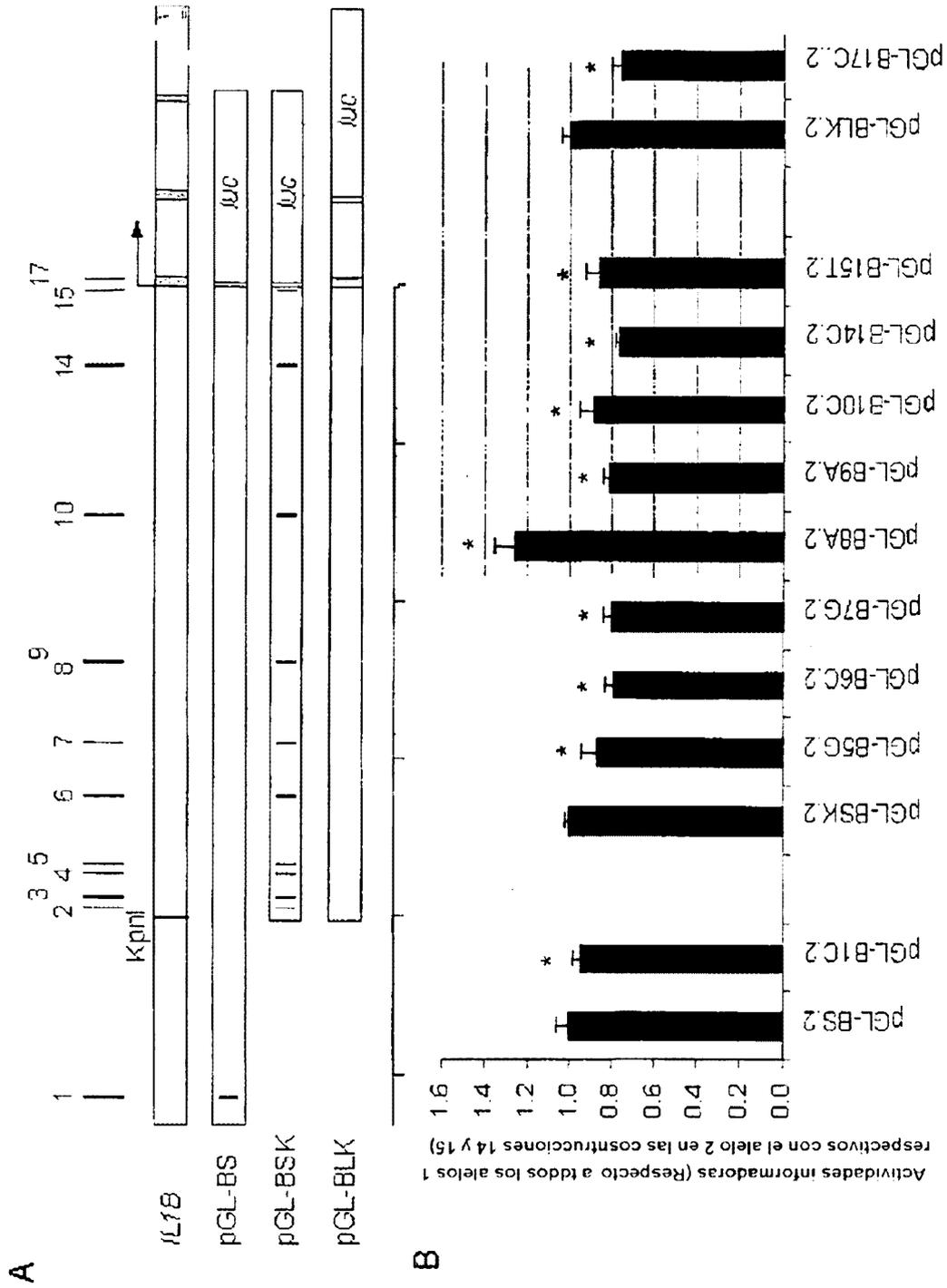


Figura 9
Construcciones de SNP para el promotor de 1L-1B (alelo 2 en 14 y 15)

Figura 10

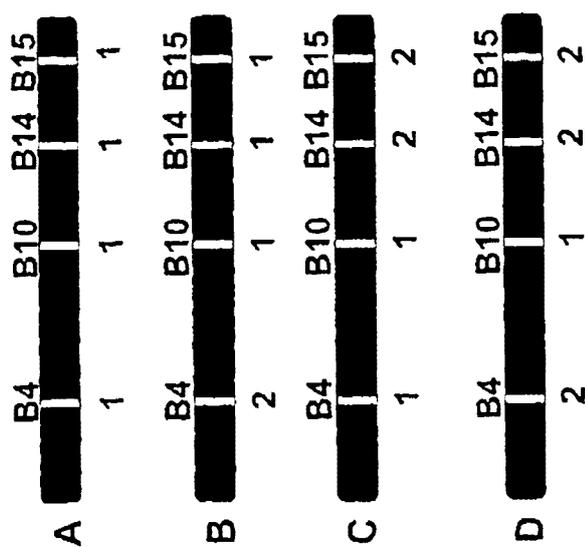
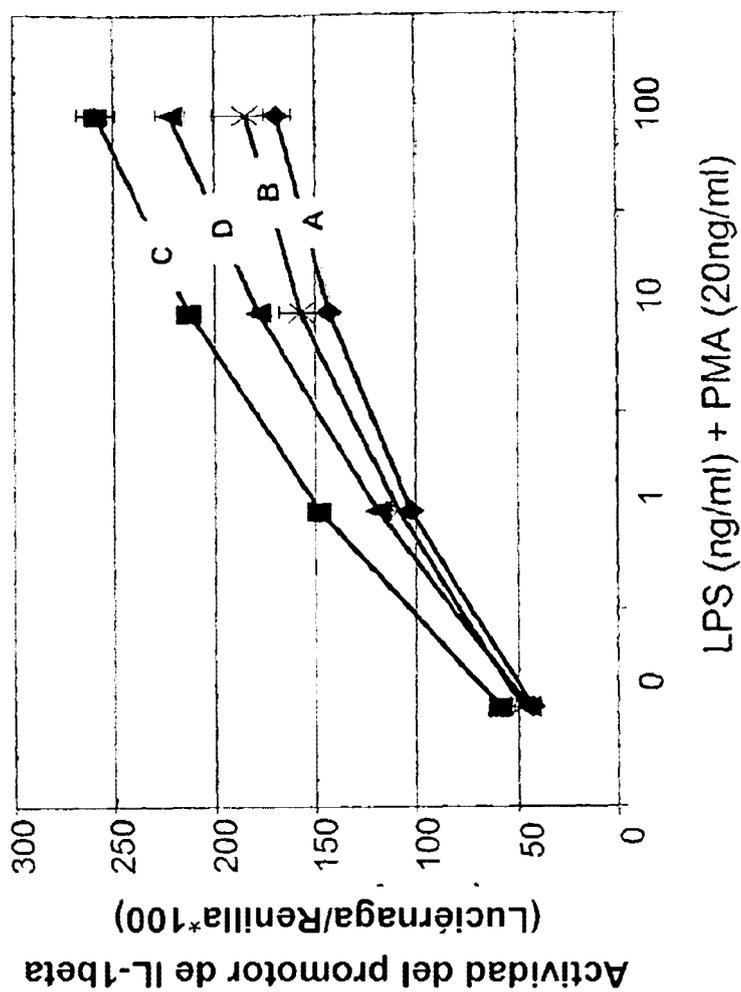


Figura 11

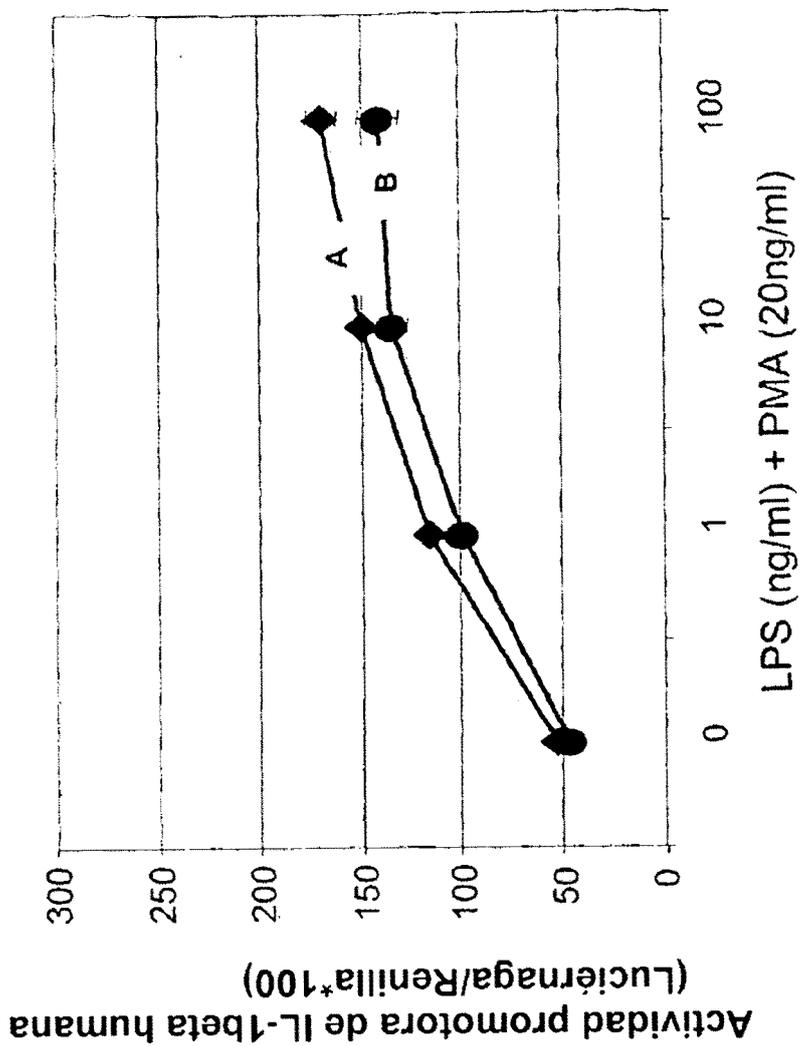


Figura 12

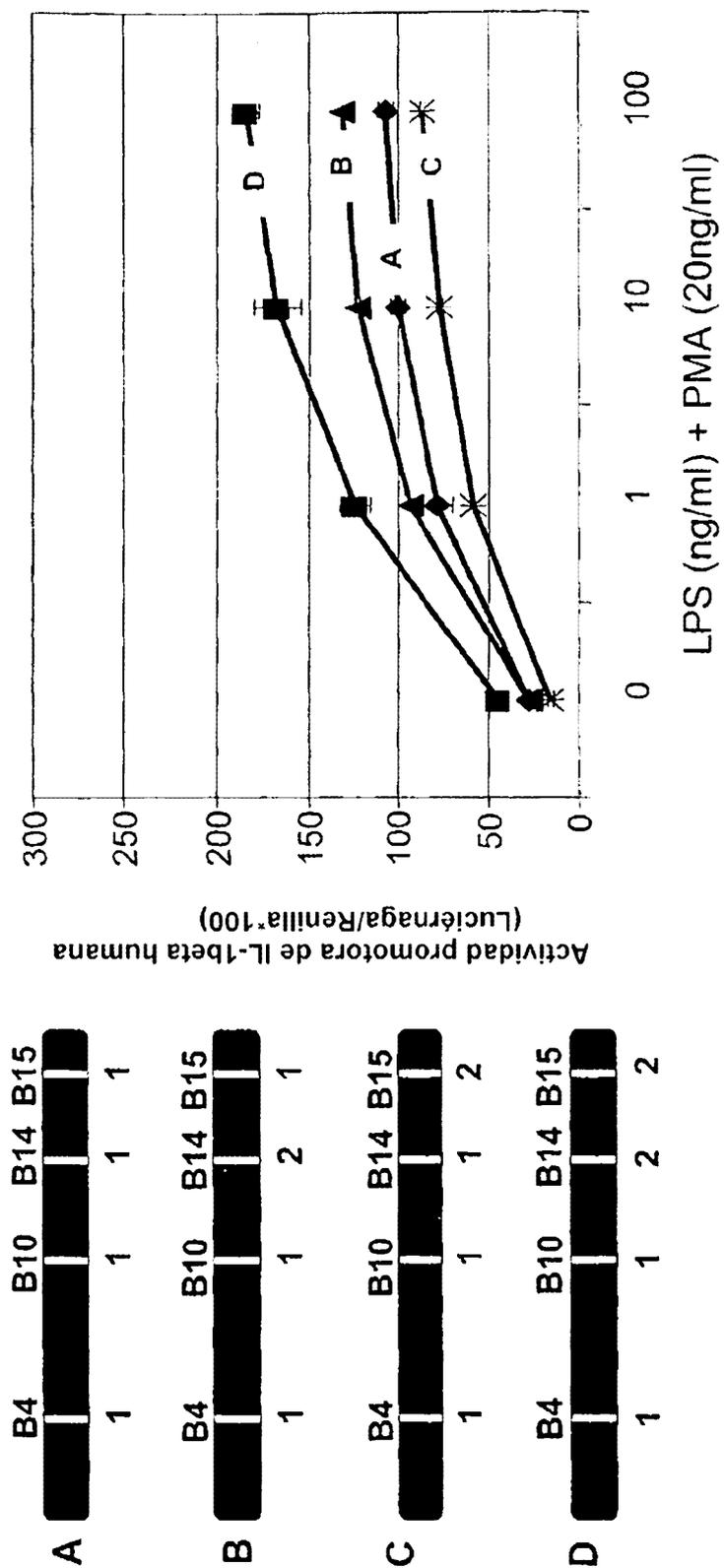
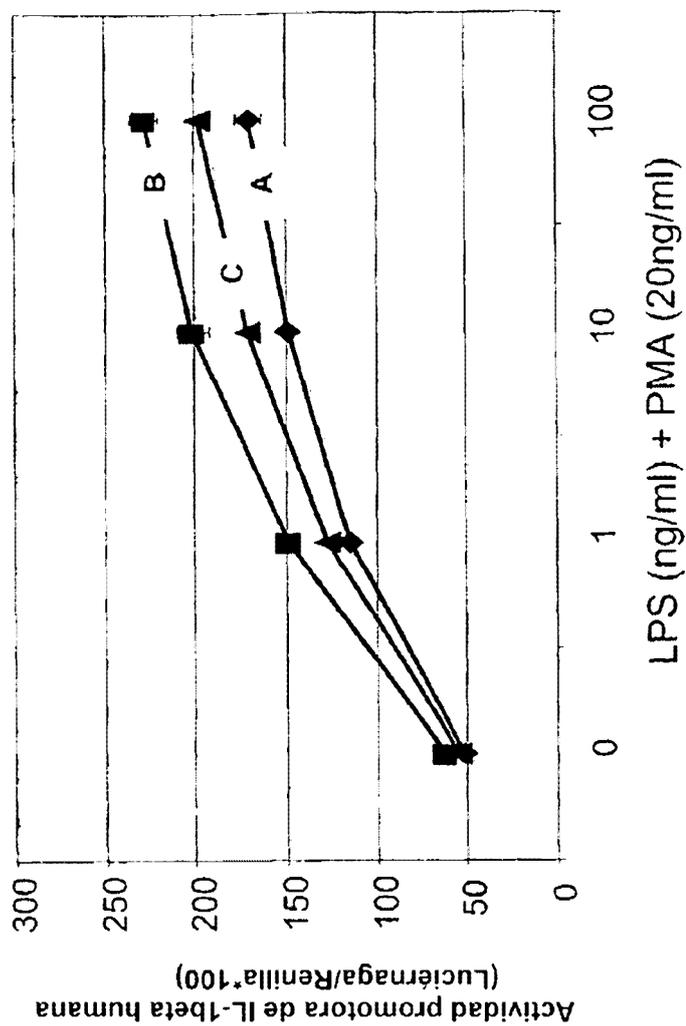


Figura 13



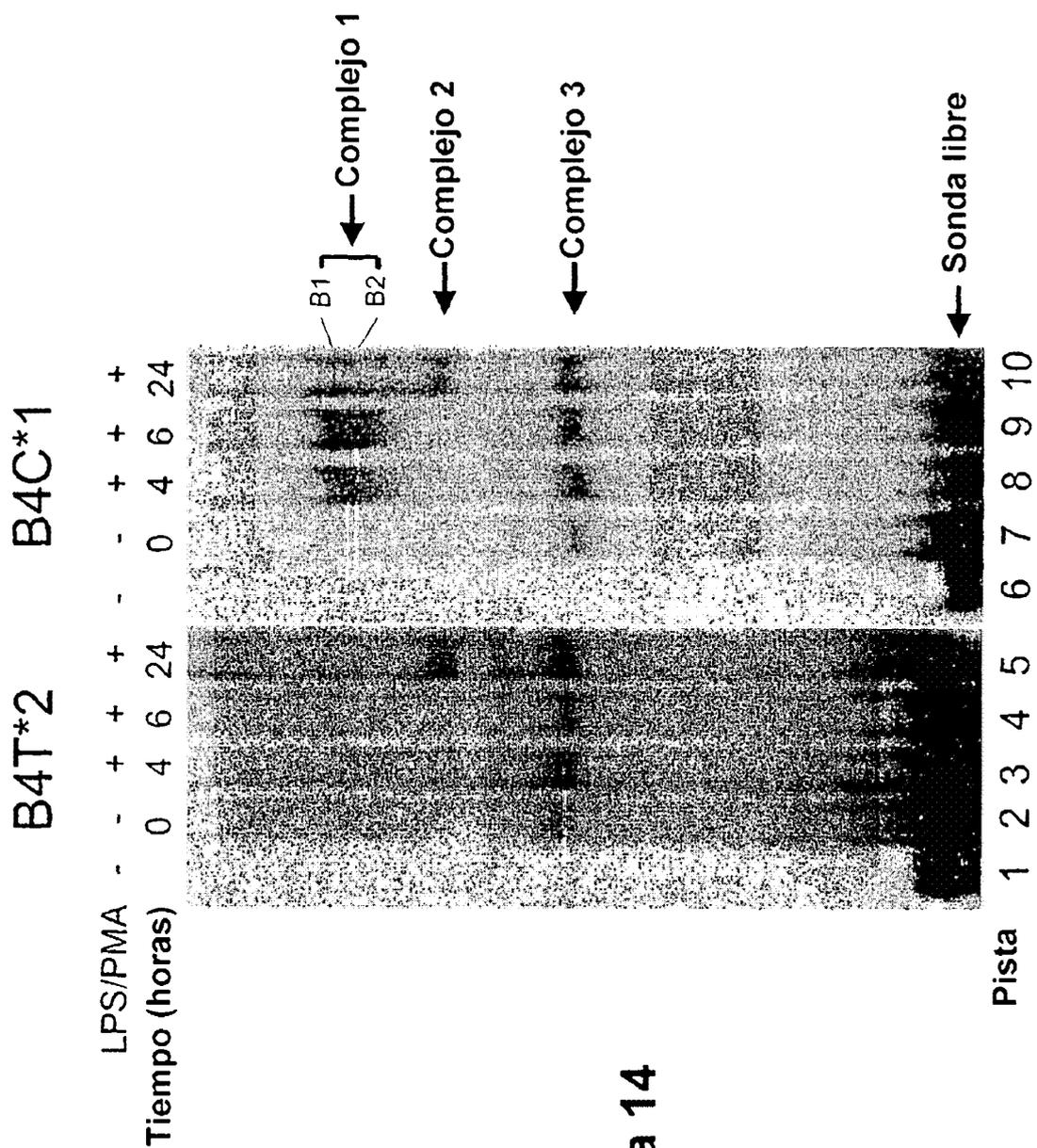


Figura 14

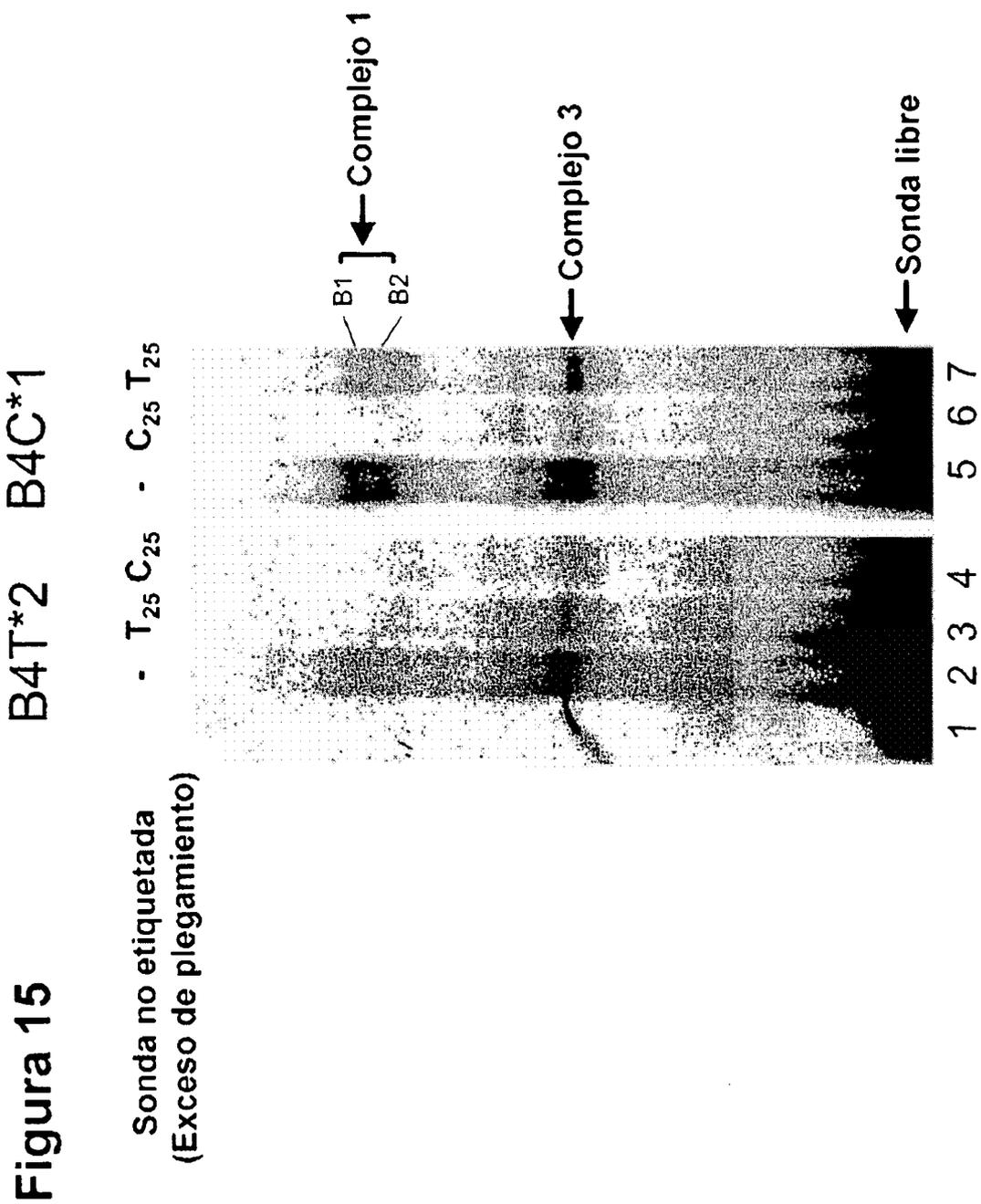


Figura 16

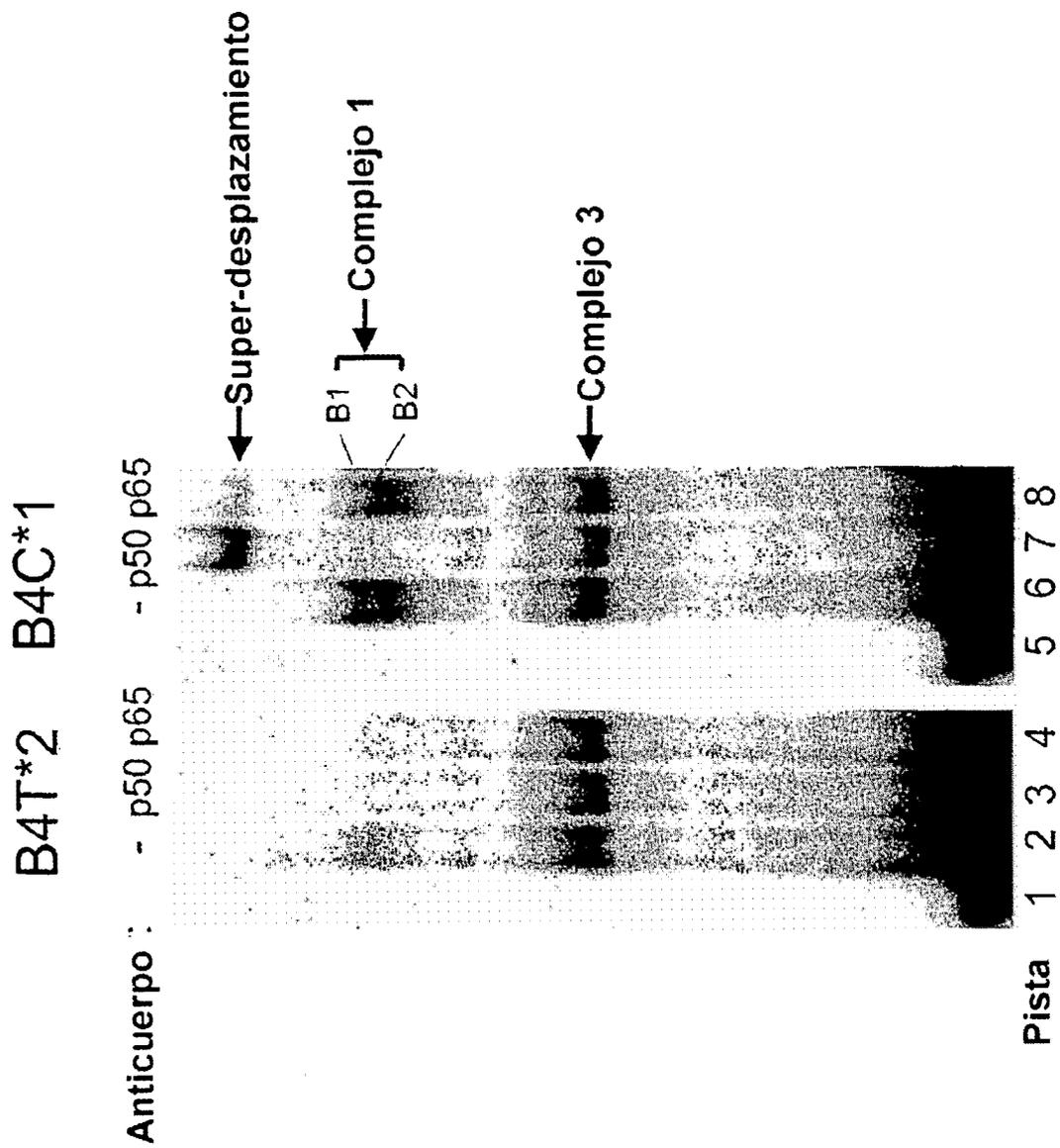


Figura 17

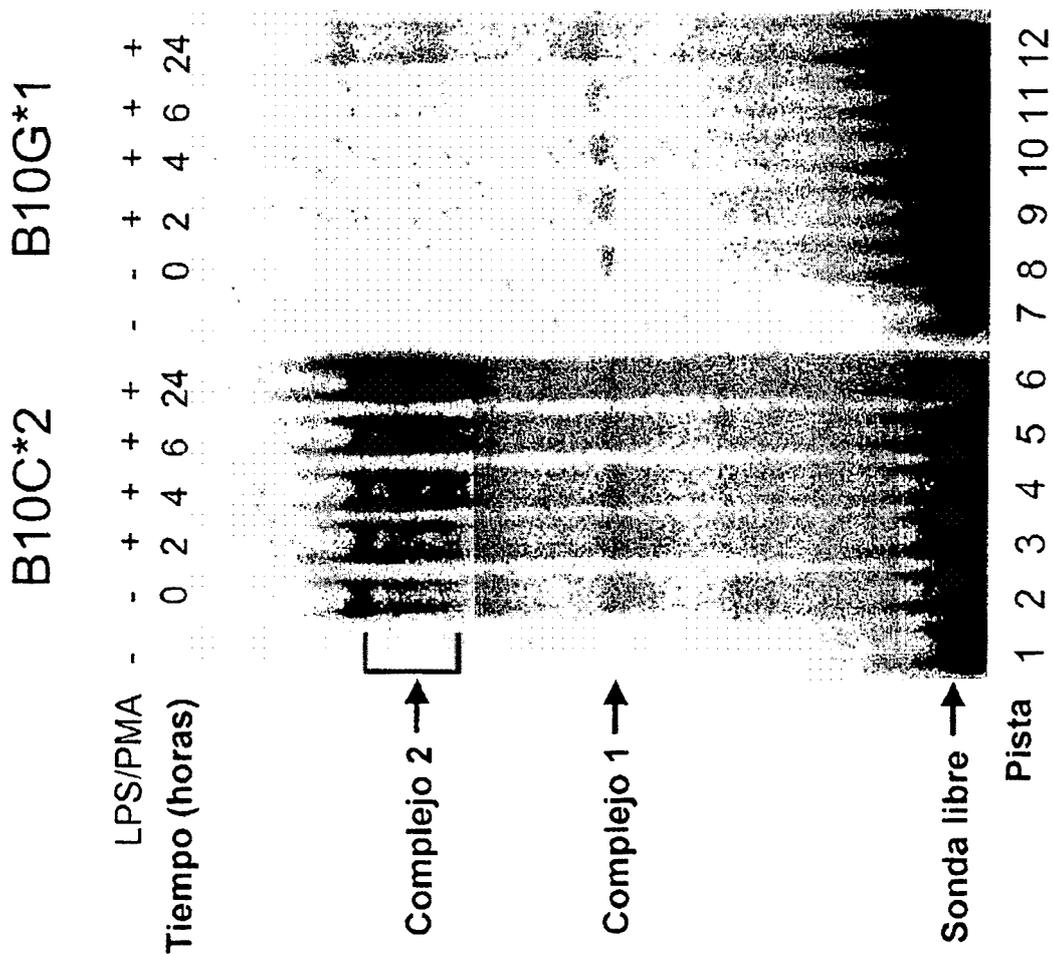


Figura 18

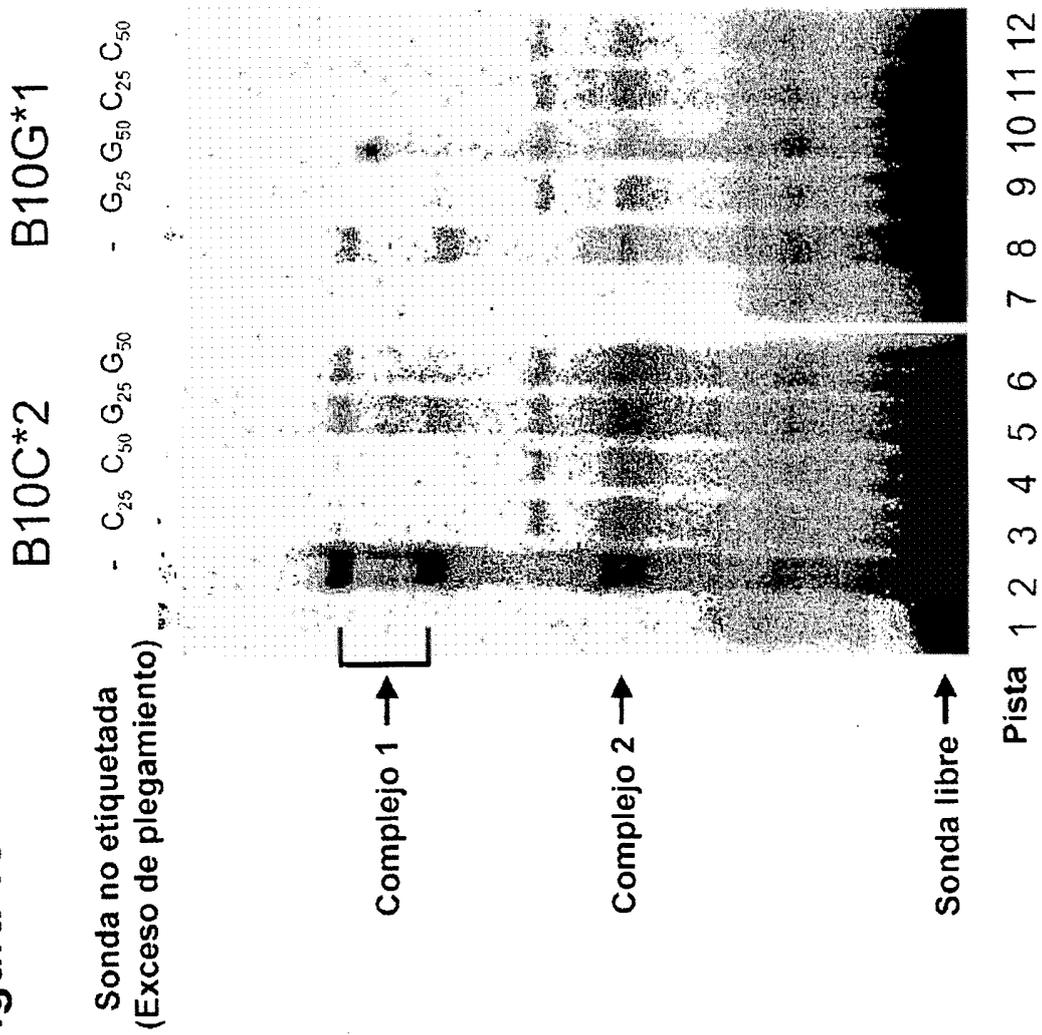
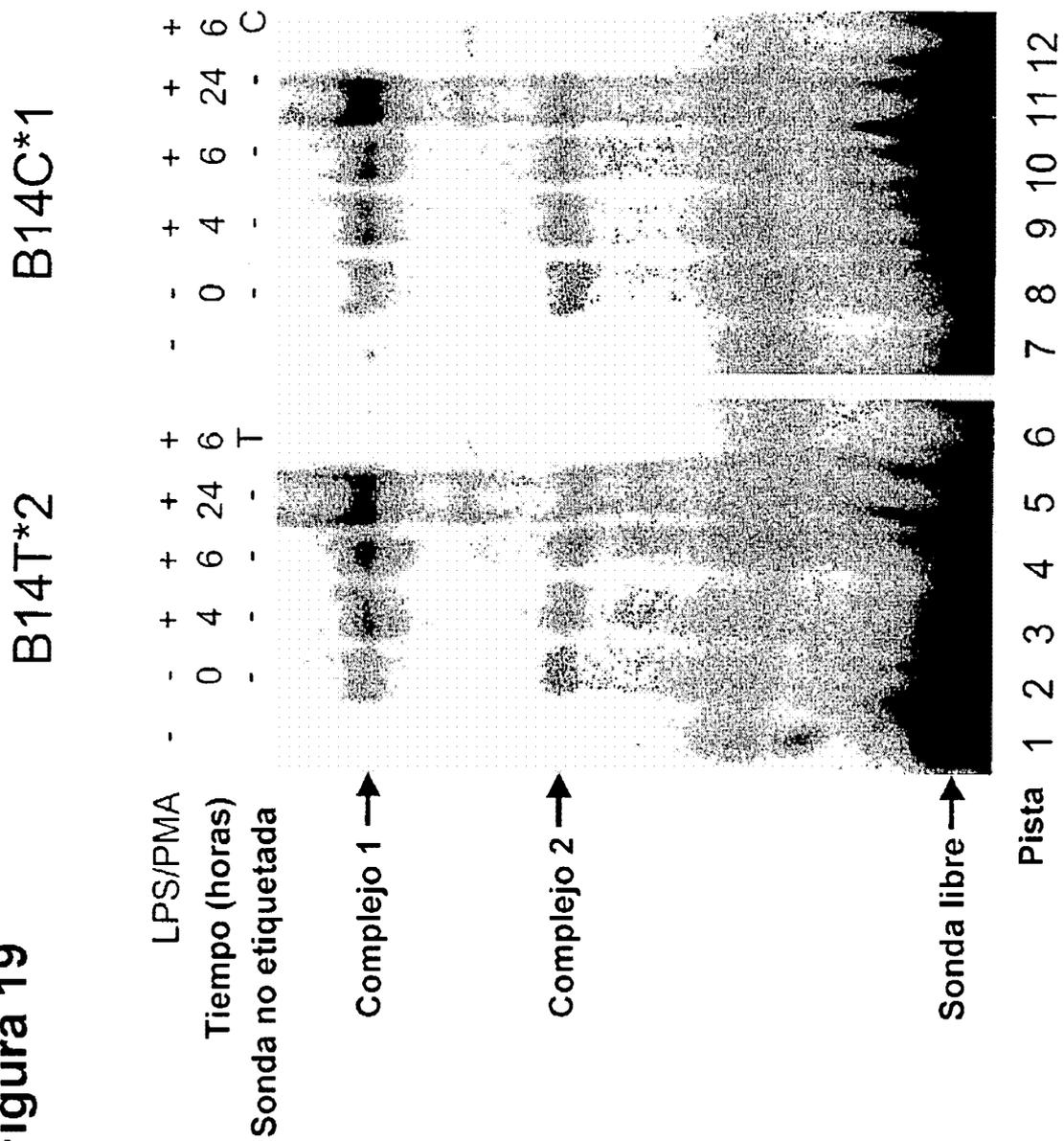


Figura 19



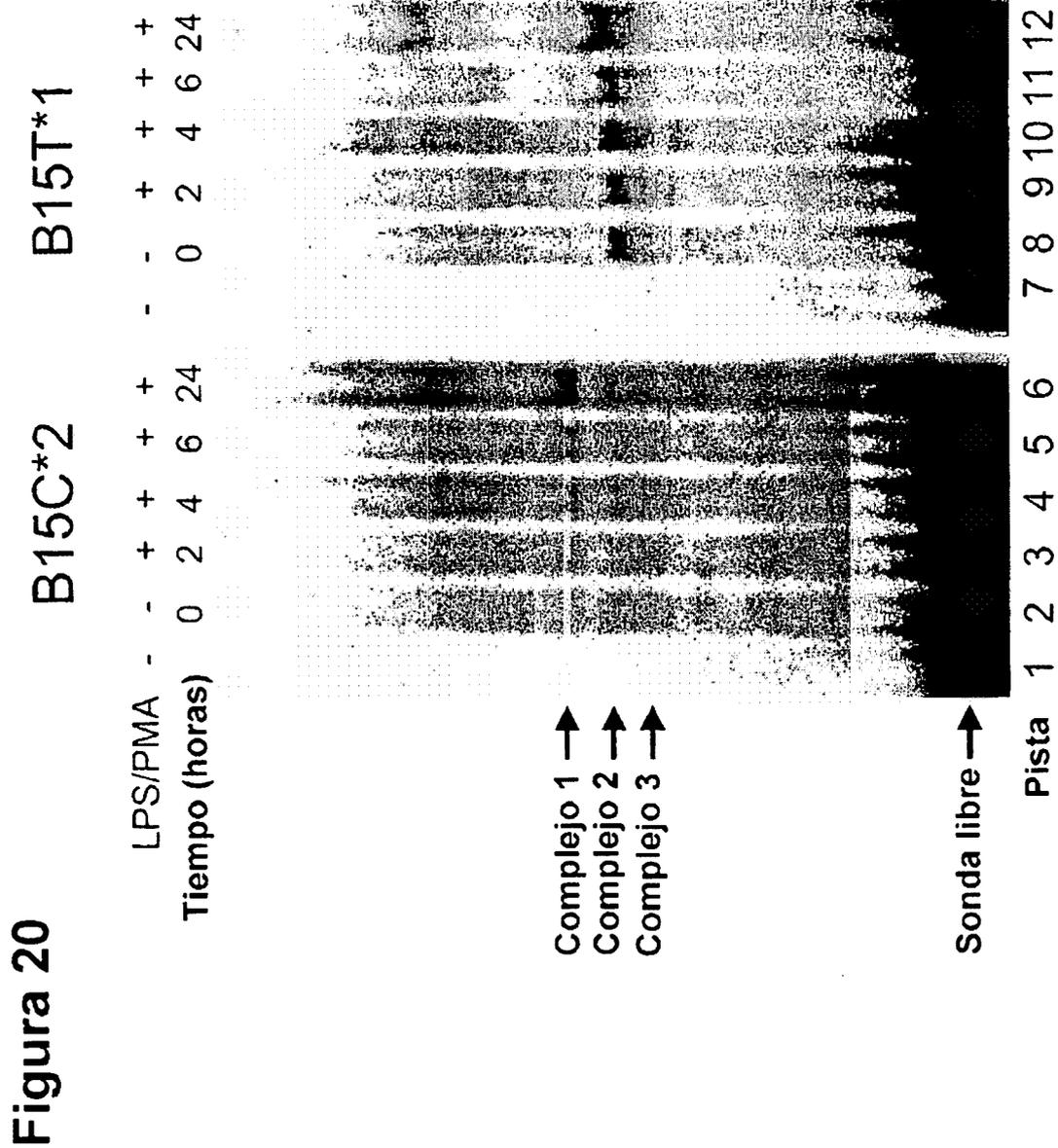


Figura 21

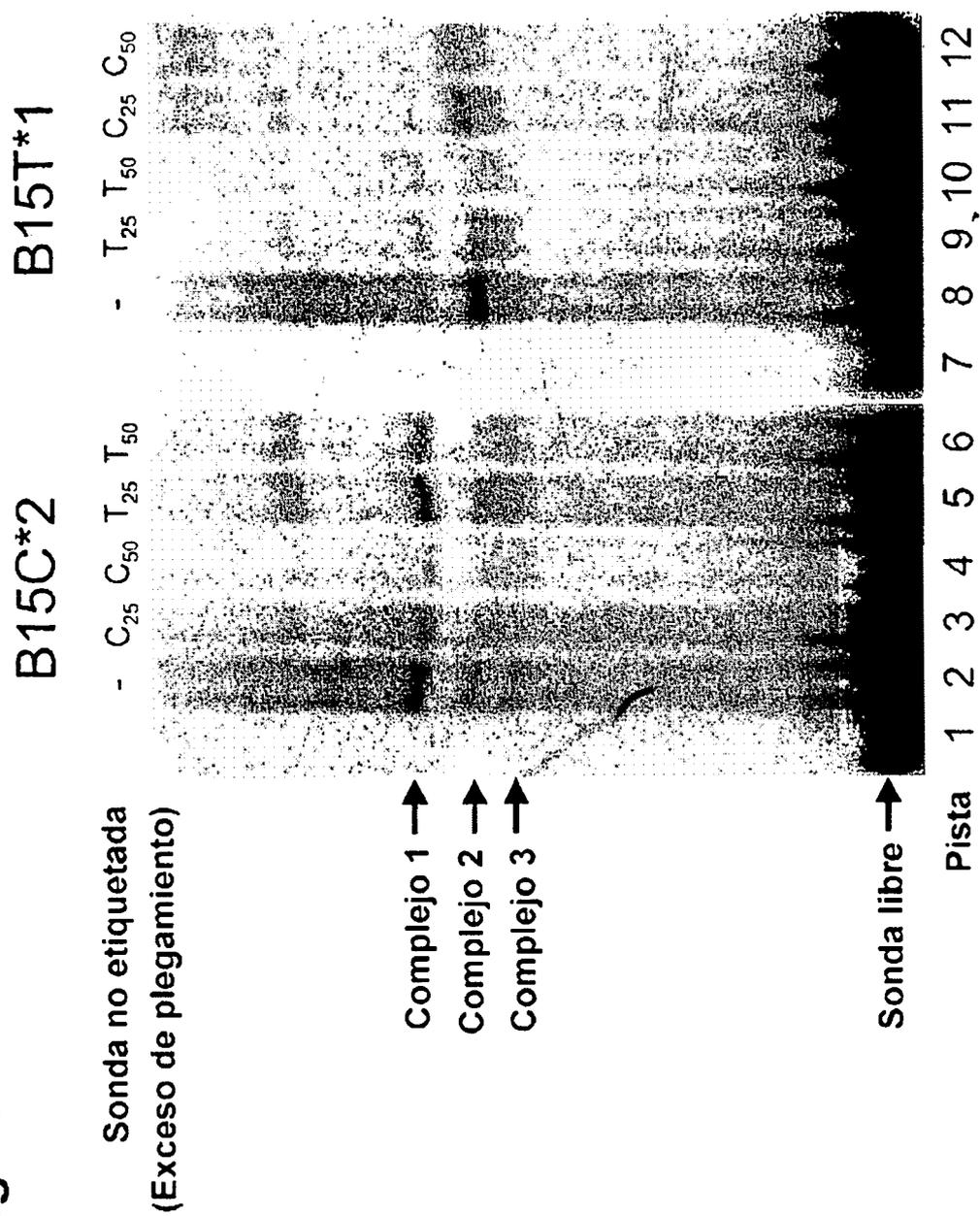


Figura 22

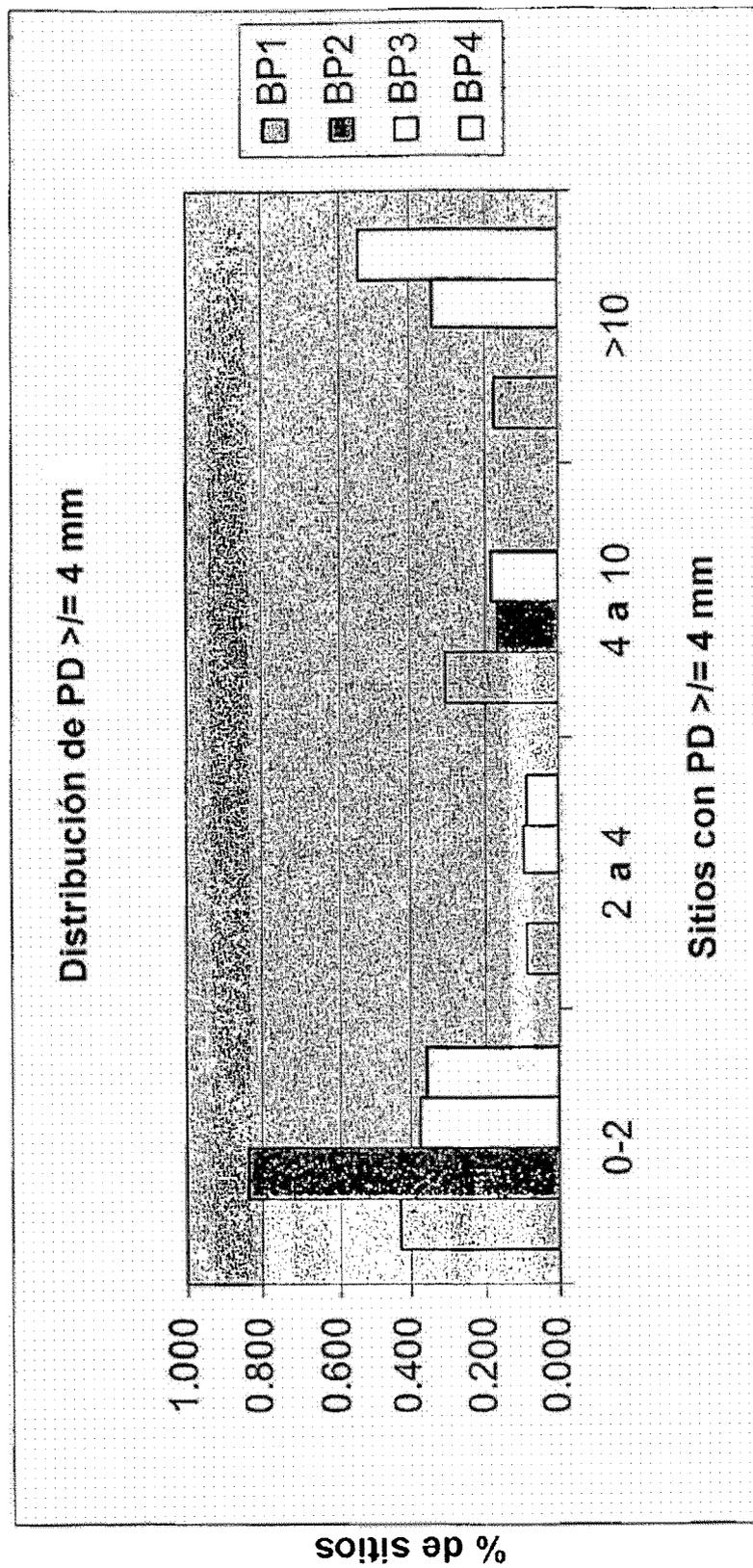


Figura 23

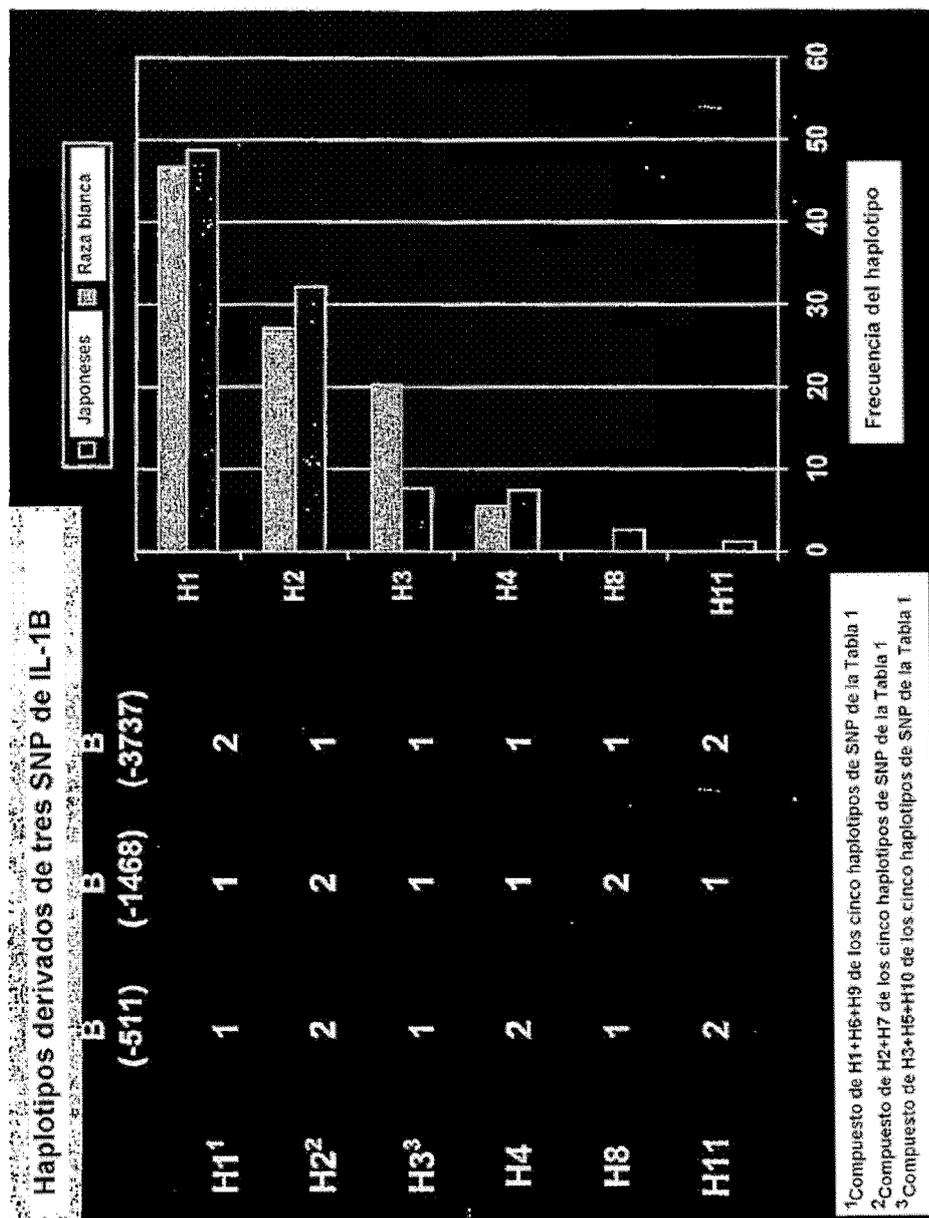


Figura 24

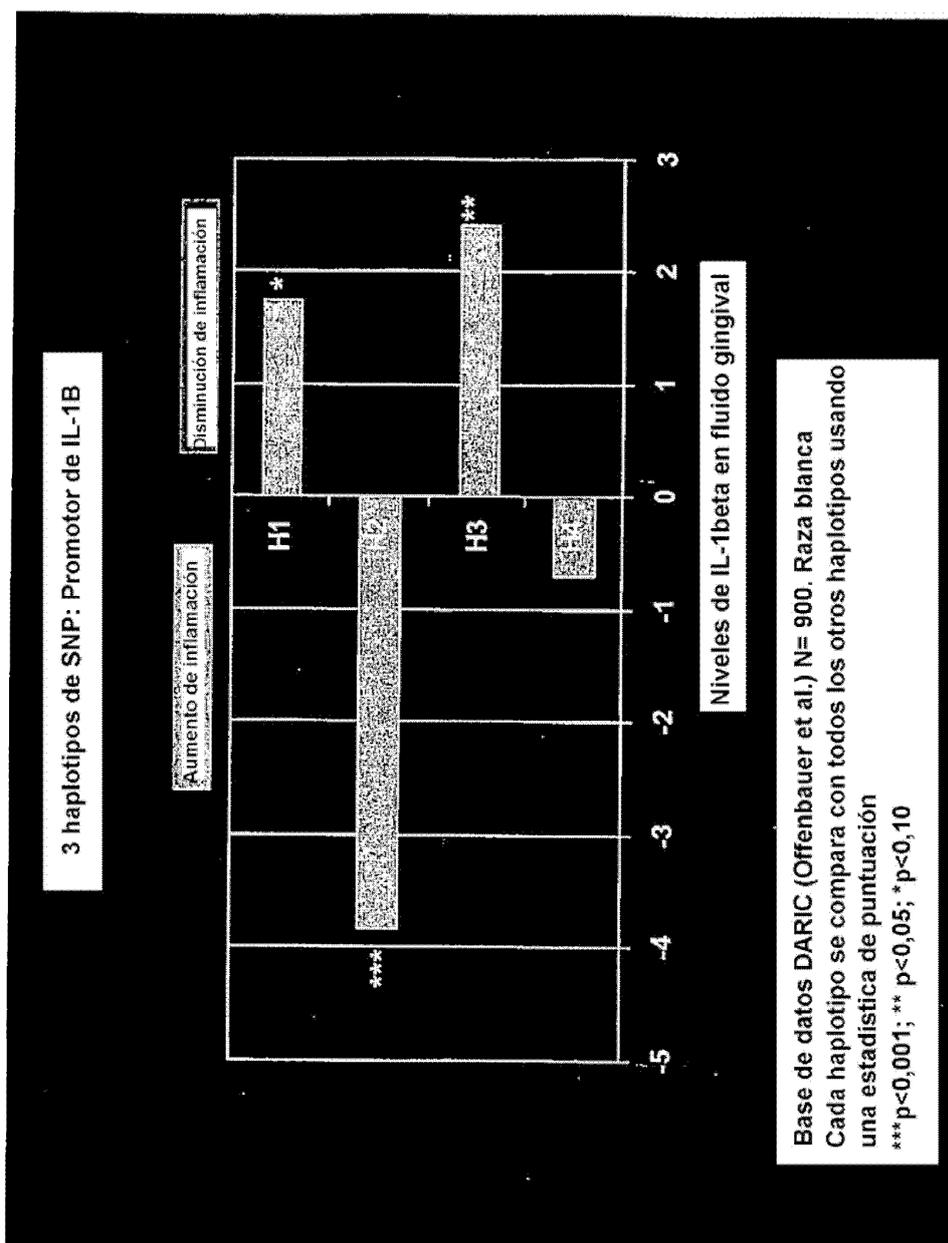


Figura 25

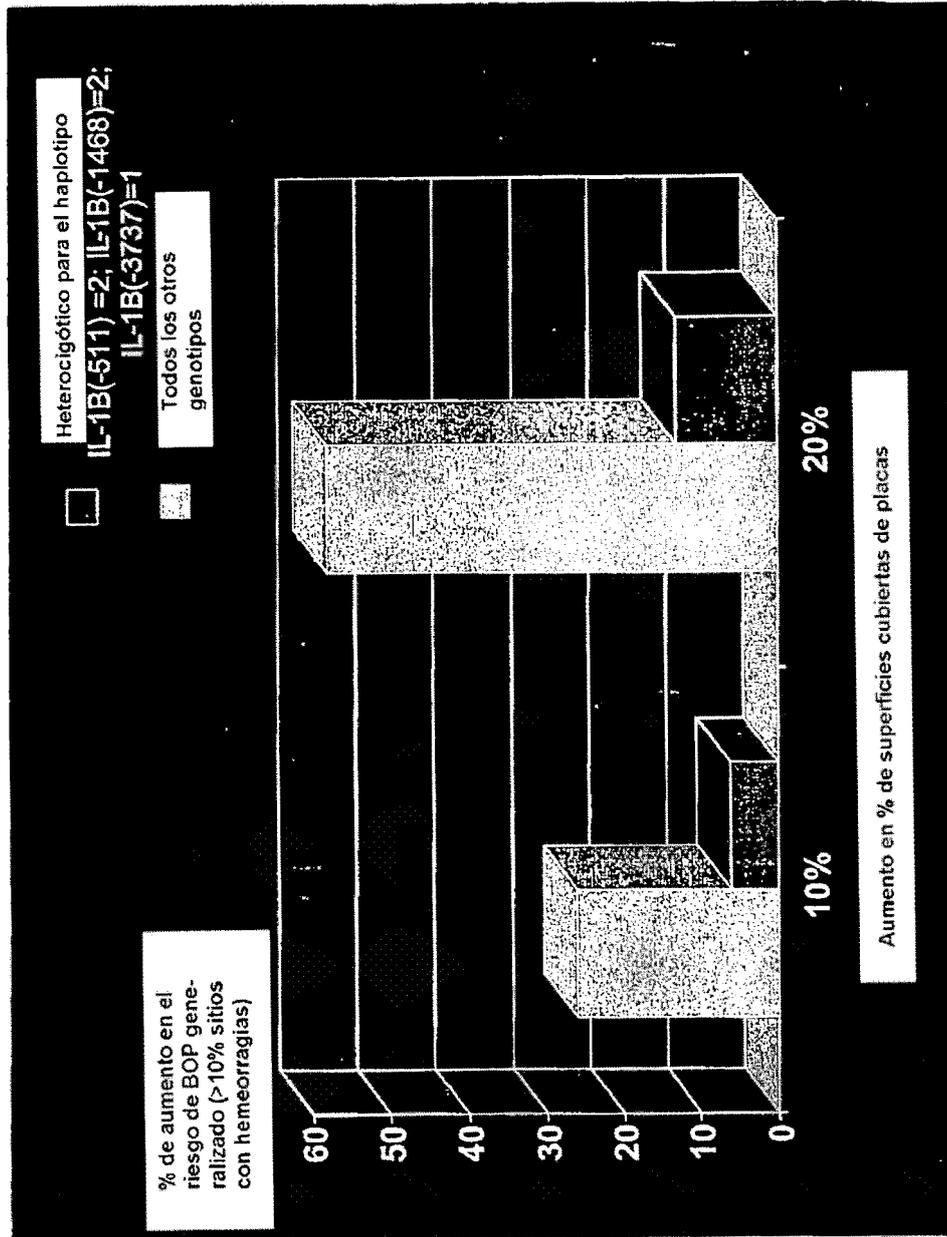
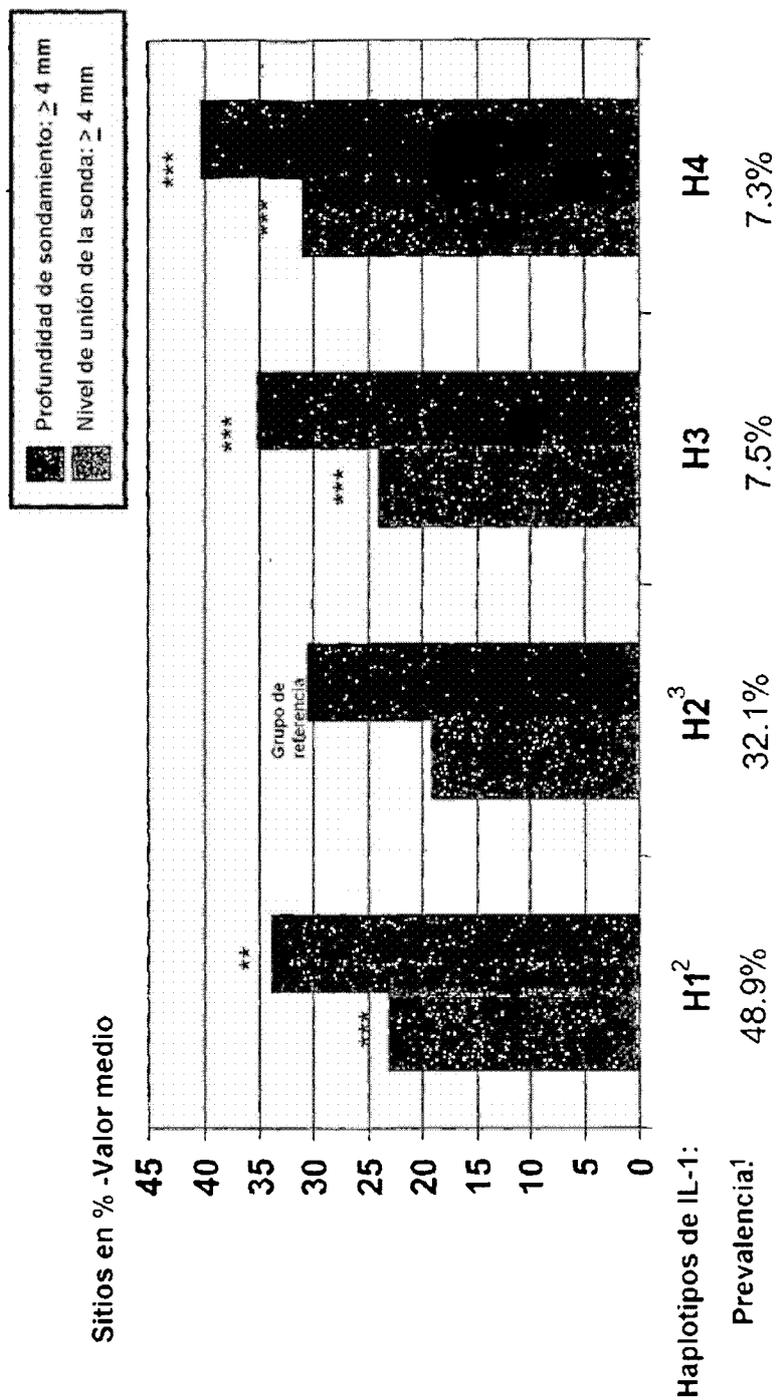


Figura 26

Los haplotipos de IL-1 están asociados con la gravedad de la periodontitis en japoneses



***p<0,001 comparado con el grupo de referencia (H2)

¹% de haplotipos totales representados por los haplotipos japoneses indicados. Los haplotipos se obtuvieron basándose en tres SNP: IL-1B(-511), IL-1B(-1468), y IL-1B(-3737).

²Cada barra representa el % de la media con PPD ≥ 4 mm o PAI ≥ 4 mm para individuos que llevan el haplotipo indicado más un haplotipo H2. Por tanto, las diferencias entre H1 (H1/H2) y H2 (H2/H2) representa el efecto del incremento de una sola copia del haplotipo H1 respecto al parámetro clínico.

³El grupo de referencia incluye la combinación de dos haplotipos, es decir, H2/H2.