

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 641 600**

51 Int. Cl.:

G06F 19/24 (2011.01)

G06F 19/20 (2011.01)

C12Q 1/68 (2006.01)

G06F 19/18 (2011.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **18.06.2010 E 10166561 (0)**

97 Fecha y número de publicación de la concesión europea: **02.08.2017 EP 2397561**

54 Título: **Métodos de determinación de la presencia o la ausencia de segmentos genéticos**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
10.11.2017

73 Titular/es:

**PROGENIKA BIOPHARMA, S.A. (100.0%)
Parque Tecnológico de Zamudio, Ibaizabal
Bidea-Edificio 801-A-2a Planta
48160 Derio (Bizcaia), ES**

72 Inventor/es:

**GARCÍA, DAVID;
LÓPEZ, MÓNICA;
TEJEDOR, DIEGO;
OCHOA, JORGE;
MARTÍNEZ, ANTONIO y
SIMÓN, LAUREANO**

74 Agente/Representante:

VALLEJO LÓPEZ, Juan Pedro

ES 2 641 600 T3

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

DESCRIPCIÓN

Métodos de determinación de la presencia o la ausencia de segmentos genéticos

5 **Campo de la invención**

La presente invención se refiere a métodos para determinar la presencia o la ausencia de uno o más segmentos genéticos funcionales. Se desvelan métodos de selección, optimización y uso de sondas oligonucleotídicas en la determinación de la presencia o la ausencia de uno o más segmentos genéticos funcionales.

10

Antecedentes de la invención

La determinación de la presencia o la ausencia de un segmento genético funcional en una muestra o en un individuo es de gran importancia en varios campos, desde la medicina forense hasta la sistemática, el diagnóstico y la gestión sanitaria. Por ejemplo, se cree que las eliminaciones que implican genes supresores de tumores desempeñan un papel importante en el desarrollo de varios tipos de cáncer (Dutrillaux *et al.*, 1990; Hensel *et al.*, 1990). En el caso de los grupos sanguíneos humanos, la ausencia de ciertos segmentos tiene grandes implicaciones para los fenotipos del grupo sanguíneo, lo que, a su vez, tiene un efecto sobre la compatibilidad transfusional. Por ejemplo, RhDel humano es causado por una eliminación de 1.013 pb que incluye el exón 9 del gen RHD (Chang *et al.*, 1998).

20

Se conoce una serie de métodos y se usan para evaluar la presencia o la ausencia de segmentos genéticos específicos. Estos incluyen la secuenciación (Sanger y Coulson, 1975; Smith *et al.*, 1986) basada en enfoques (Kitano *et al.*, 2000), así como el uso de la electroforesis o la transferencia de Southern (Southern, 1975) para detectar marcadores tales como polimorfismos de longitud del fragmento de restricción (RFLP, Botstein *et al.*, 1980) u otros marcadores basados en la reacción en cadena de la polimerasa (PCR, Mullis *et al.*, 1986; Mullis y Faloona, 1987), tales como los amplicones de los cebadores específicos de secuencias (Olerup y Zetterquist, 1992) o polimorfismos conformacionales monocatenarios de productos de PCR (Jin *et al.*, 1993).

25

“Statistical Algorithms Reference Guide”, Affymetrix, describe los algoritmos estadísticos de Affymetrix usados en el análisis de expresión de matrices de sondas GeneChip®.

30

Bolstad *et al.*, “Bioinformatics”, 2003, Vol. 19, n.º 2, pág. 185-193, describe una comparación de métodos de normalización para datos de matrices de oligonucleótidos de alta densidad basados en la varianza y el sesgo.

35

He y Zhou, “Applied and Environmental Microbiology”, 2008, Vol. 74, n.º 10, pág. 2957-2966, describe la evaluación empírica de un nuevo método de cálculo de la relación señal-ruido para el análisis de datos de micromatrices.

El documento WO2006/075254 describe un método *in vitro* de genotipificación de variaciones genéticas en un individuo, y productos de uso en el método.

40

Liu *et al.*, “Microarrays: Optical Technologies and Informatics, Proceedings of SPIE”, 2001, Vol. 4266, pág. 56-67, describe algoritmos basados en rangos para el análisis de micromatrices.

Zhao *et al.*, *Physiol. Genomics*, 2004, Vol. 19, pág. 93-105, describe el análisis de micromatrices de expresión génica tras la constricción aórtica transversa en ratones.

45

Liu *et al.*, “Bioinformatics”, 2003, Vol. 19, n.º 18, pág. 2397-2403, describe algoritmos para micromatrices de genotipificación a gran escala.

Di *et al.*, “Bioinformatics”, 2005, Vol. 21, n.º 9, pág. 1958-1963, describe algoritmos basados en modelos dinámicos para el rastreo y la genotipificación de SNP de 100 K en micromatrices oligonucleotídicas. Varias patentes y/o solicitudes de patente describen también métodos de detección de la ausencia o de eliminaciones de ciertos segmentos genéticos funcionales, o su uso para ciertos objetivos. Por ejemplo, la solicitud de patente WO 2010/008071 desvela un método de detección de la presencia o la ausencia de un exón que tenga cualquier mutación génica en el mismo. El método es un método de detección de mutaciones génicas utilizando una técnica de PCR cuantitativa. La patente de EE.UU. n.º 6599701 proporciona métodos de caracterización de organismos mediante la identificación de la presencia, de la ausencia, del tamaño o del polimorfismo de secuencias de regiones intrónicas. El documento WO 2001/018245 proporciona métodos de identificación de una alteración en un gen de interés, en particular, en la región de histocompatibilidad principal, utilizando la amplificación de reacción en cadena de la polimerasa (LR-PCR) de largo alcance del ADN diana que incluye todo o una parte de un elemento móvil humano.

50

55

60

La presente invención tiene por objeto proporcionar métodos para la selección y el uso de conjuntos de sondas que permitan detectar la presencia o la ausencia de segmentos funcionales, y también, en ciertos casos, tanto para detectar variantes alélicas ubicadas dentro de sus secuencias afines como para determinar la presencia o la ausencia de segmentos funcionales de forma simultánea, con un alto rendimiento.

65

Divulgación de la invención

En términos generales, la presente invención proporciona métodos y productos para determinar la presencia o la ausencia de un segmento genético de interés, tal como un exón, un intrón o un promotor, en una muestra que contiene ADN. Los presentes inventores han encontrado que se puede emplear un conjunto de sondas que interroguen secuencias de interés dentro del segmento genético de interés en una etapa de formación usando una pluralidad de muestras de referencia de estado presente/ausente conocido, etapa de formación que establece límites "sin determinación" que permiten la posterior clasificación de las muestras como "presentes" o "ausentes" con un alto grado de confianza estadística. En algunos casos, las sondas incluyen sondas específicas de alelos que permiten la determinación simultánea de un genotipo en uno o más sitios polimórficos junto con la determinación de la presencia o la ausencia de un segmento genético. Los métodos de la invención pueden proporcionar, por lo tanto, ahorros de eficiencia considerables en comparación con, por ejemplo, la secuenciación completa.

Por consiguiente, en un primer aspecto, la presente invención proporciona un método de determinación de la presencia o la ausencia de un segmento genético de interés, tal como un exón, un intrón o un promotor, en una muestra que contiene ADN, comprendiendo el método:

(i) poner en contacto un primer conjunto de sondas que comprende una pluralidad de réplicas de al menos una sonda oligonucleotídica que interroga a una primera secuencia afín dentro de dicho segmento de interés con:

- (a) una pluralidad de muestras de referencia en las que el segmento genético de interés esté ausente; y
- (b) una pluralidad de muestras de referencia en las que el segmento genético de interés esté presente;

en condiciones que permitan que se produzca la hibridación de la sonda y la secuencia afín;

(ii) medir la intensidad de hibridación del conjunto de sondas de cada una de las muestras de referencia, obteniéndose de este modo, de la etapa (i)(a), un primer grupo de valores de intensidad de hibridación para las muestras de referencia en las que el segmento genético de interés está ausente y, de la etapa (i)(b), un segundo grupo de valores de intensidad de hibridación para las muestras de referencia en las que el segmento genético de interés está presente; y

(iii) establecer una región "sin determinación" de los valores de intensidad de hibridación que se encuentra en la región entre dichos primer y segundo grupos y que está limitada por un límite sin determinación inferior ("L_{NC}") y un límite sin determinación superior ("U_{NC}"), en la que el L_{NC} y el U_{NC} representan límites de confianza estadística para asignar un valor de intensidad de la hibridación a dichos primer y segundo grupos, respectivamente, y se calculan de la siguiente manera:

$$L_{NC} = \frac{T_{I/B} \cdot [2 \cdot a(I/B)_n - aMediana I/B] - a(I/B)_n^2}{T_{I/B} - aMediana I/B}$$

$$U_{NC} = \frac{T_{I/B} \cdot [pMediana I/B - 2 \cdot p(I/B)_1] + p(I/B)_1^2}{pMediana I/B - T_{I/B}}$$

donde:

$$T_{I/B} = \frac{pMediana I/B \cdot a(I/B)_n - aMediana I/B \cdot p(I/B)_1}{pMediana I/B + a(I/B)_n - aMediana I/B - p(I/B)_1}$$

I/B es la intensidad de hibridación del conjunto de sondas;

aMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que el segmento genético de interés está ausente;

pMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que el segmento genético de interés está presente;

a(I/B)_n es el mayor valor de I/B medido de las muestras de referencia en las que el segmento genético de interés está ausente; y

p(I/B)₁ es el menor valor de I/B medido de las muestras de referencia en las que el segmento genético de interés está presente;

iv) poner en contacto el primer conjunto de sondas con al menos una muestra de ensayo que contenga ADN en condiciones que permitan que se produzca la hibridación de la sonda y la secuencia afín;

- v) medir la intensidad de la hibridación del conjunto de sondas para la al menos una muestra de ensayo; y
- vi) comparar la intensidad de hibridación medida con el L_{NC} y el U_{NC} ;

en el que una intensidad de hibridación medida por debajo del L_{NC} indica que dicho segmento genético de interés está ausente en la muestra de ensayo y una intensidad de hibridación medida por encima del U_{NC} indica que dicho segmento genético de interés está presente en la muestra de ensayo; y en el que el ADN de las muestras de referencia y/o la muestra de ensayo se ha marcado con un marcador fluorescente; y en el que la medición de la intensidad de hibridación comprende medir la señal de fluorescencia del marcador fluorescente en cada ubicación de la sonda oligonucleotídica; y en el que la intensidad de hibridación para cada conjunto de sondas se mide como I/B, donde:

I se determina como la media o la mediana de la señal de fluorescencia medida combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente, tras el recorte de las lecturas extremas; y

B se determina como la media o la mediana de la señal de fluorescencia de fondo combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente, tras el recorte de las lecturas extremas.

En algunos casos, dicha secuencia afin comprende un sitio polimórfico y en el que dicho primer conjunto de sondas comprende una pluralidad de réplicas de al menos dos, al menos tres o al menos cuatro sondas oligonucleotídicas específicas de alelos que distinguen los alelos de dicho sitio polimórfico. En algunos casos, dicha muestra de ensayo comprende dicho segmento genético de interés, comprendiendo además el método la genotipificación de dicha muestra de ensayo para identificar un alelo en dicho sitio polimórfico.

En algunos casos, se emplea al menos un segundo conjunto de sondas además de dicho primer conjunto de sondas y en el que el segundo conjunto de sondas comprende una pluralidad de réplicas de al menos una sonda oligonucleotídica que interroga a una segunda secuencia afin dentro de dicho segmento de interés, comprendiendo el método además establecer una segunda región sin determinación limitada por los límites L_{NC} y U_{NC} para el segundo conjunto de sondas. En algunos casos, la muestra de ensayo se clasifica como aquella que tiene o que no tiene dicho segmento de interés basado en un consenso estricto entre las determinaciones del primer (PS_1) y segundo (PS_2) conjunto de sondas, de acuerdo con un consenso estricto como el mostrado en la Tabla 1:

(Tabla 1)

Valor de I/B de PS_1	Valor de I/B de PS_2	Determinación de segmento genético
$I/B \text{ de } PS_1 > U_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Presente
$I/B \text{ de } PS_1 < L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Ausente
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Sin determinación
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Sin determinación
$I/B \text{ de } PS_1 > U_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación
$I/B \text{ de } PS_1 < L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación

En algunos casos, la muestra de ensayo se clasifica como aquella que tiene o que no tiene dicho segmento de interés basado en un consenso mayoritario entre las determinaciones del primer (PS_1) y segundo (PS_2) conjunto de sondas en ausencia de determinaciones contradictorias, de acuerdo con un consenso mayoritario como el mostrado en la Tabla 2:

(Tabla 2)

Valor de I/B de PS_1	Valor de I/B de PS_2	Determinación de segmento genético
$I/B \text{ de } PS_1 > U_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Presente
$I/B \text{ de } PS_1 < L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Ausente
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Presente
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Ausente
$I/B \text{ de } PS_1 > U_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Presente
$I/B \text{ de } PS_1 < L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Ausente

En algunos casos, las sondas oligonucleotídicas están unidas a un soporte sólido, opcionalmente, en el que:

- (a) las sondas oligonucleotídicas están unidas a un soporte sólido esencialmente plano en forma de una matriz; o
- (b) las sondas oligonucleotídicas están unidas a partículas, tales como perlas, cilindros o cualquier otra partícula de tamaño micrométrico o nanométrico.

En algunos casos:

- (a) la muestra de ensayo comprende ADN amplificado a partir de ADN genómico de un sujeto de ensayo, ADN que opcionalmente se ha fragmentado y/o marcado con un marcador detectable; y/o
- (b) cada una de las muestras de referencia comprende ADN amplificado a partir de ADN genómico de un sujeto de referencia, ADN que opcionalmente se ha fragmentado y/o marcado con un marcador detectable.

En algunos casos, el segmento genético de interés comprende un exón del gen RHD humano. En algunos casos, la sonda oligonucleotídica se selecciona de las sondas mostradas en la Tabla 4. En algunos casos, el método comprende además la genotipificación de la muestra de ensayo para identificar al menos un alelo en un sitio de polimorfismo de un solo nucleótido ("SNP") en el gen RHD humano.

En algunos casos, dicho primer conjunto de sondas se selecciona:

- (i) proporcionando una pluralidad de conjuntos de sondas candidatos, comprendiendo cada conjunto de sondas candidato una pluralidad de réplicas de al menos una sonda oligonucleotídica que interroga a una secuencia afín dentro de un segmento genético de interés;
- (ii) poniendo en contacto los conjuntos de sondas candidatos con:

- (a) una pluralidad de muestras de referencia en las que el segmento genético de interés esté ausente; y
- (b) una pluralidad de muestras de referencia en las que el segmento genético de interés esté presente;

en condiciones que permitan que se produzca la hibridación de la sonda y la secuencia afín;

(ii) midiendo la intensidad de hibridación del conjunto de sondas de cada una de las muestras de referencia, obteniéndose de este modo, de la etapa (ii)(a), para cada conjunto de sondas candidato, un primer grupo de valores de intensidad de hibridación para las muestras de referencia en las que el segmento genético de interés está ausente y, de la etapa (ii)(b), un segundo grupo de valores de intensidad de hibridación para las muestras de referencia en las que el segmento genético de interés está presente; y

(iv) puntuando cada uno de los conjuntos de sondas candidatos dividiendo la amplitud del intervalo entre el mayor valor de intensidad de hibridación medido del primer grupo y el menor valor de intensidad de hibridación medido del segundo grupo entre la suma de la mediana o de la media de los valores de intensidad de hibridación del primer y segundo grupo, obteniéndose así una puntuación del conjunto de sondas para cada uno de los conjuntos de sondas candidatos;

de modo que una puntuación del conjunto de sondas superior indica que es más probable que el conjunto de sondas candidato sea óptimo para determinar la presencia o la ausencia del segmento genético de interés, y de modo que las muestras de referencia se marcan con un marcador fluorescente y la intensidad de hibridación se mide como I/B, donde:

I se determina como la media o la mediana de la señal de fluorescencia medida combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente, tras el recorte de las lecturas extremas; y

B se determina como la media o la mediana de la señal de fluorescencia de fondo combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente, tras el recorte de las lecturas extremas.

En algunos casos, la puntuación de conjunto de sondas (S) para cada conjunto de sondas candidato se calcula de la siguiente manera:

$$S = \frac{A}{M} = \frac{p(I/B)_1 - a(I/B)_n}{p\text{Mediana } I/B + a\text{Mediana } I/B}$$

donde:

- aMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que el segmento genético de interés está ausente;
- pMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que el segmento genético de interés está presente;
- a(A/B)_n es el mayor valor de I/B medido de las muestras de referencia en las que el segmento genético de interés está ausente; y

$p(I/B)_1$ es el menor valor de I/B medido de las muestras de referencia en las que el segmento genético de interés está presente.

La presente invención incluye la combinación de los aspectos y las características preferidos que se describen excepto cuando dicha combinación sea claramente inadmisibles o esté indicado que se evite expresamente. Estos y otros aspectos y realizaciones de la invención se describen con más detalle a continuación y con referencia a los ejemplos y a las figuras adjuntas.

Descripción de las figuras

La **Figura 1** muestra la frecuencia de las muestras en el conjunto de muestras de referencia en función de su proporción de I/B para los conjuntos de sondas óptimos **A**) PS9 y **B**) PS11 (ID de genotipificación del grupo sanguíneo BCV053 y BCV055, respectivamente). Las barras de la izquierda muestran aquellas muestras en las que el exón 5 de RHD está ausente; las barras de la derecha muestran aquellas en las que el exón 5 de RHD está presente.

La **Figura 2** muestra en detalle la región de transición determinada a partir de las proporciones de I/B de muestras en las que se sabe que el exón 05 de RHD está ausente (lado de la izquierda) a las proporciones de I/B de muestras en las que se sabe que está presente (lado de la derecha) para conjuntos de sondas óptimos **A**) PS9 y **B**) PS11 (ID de genotipificación de grupo sanguíneo BCV053 y BCV055, respectivamente). El área sombreada muestra la región "sin determinación" limitada por el límite inferior (L_{NC}) y el límite superior (U_{NC}), con el umbral de la proporción de I/B ponderado ($T_{I/B}$) indicado por la línea vertical discontinua.

Descripción detallada de la invención

La presente invención se refiere a un método de genotipificación y a un análisis de datos para determinar la presencia o la ausencia de uno o más segmentos funcionales (por ejemplo, promotor, exón, intrón), partes de los mismos y/o combinaciones de los mismos, en particular, locus de ADN genómico de un individuo.

El método hace uso de la detección o de la falta de detección de una o más secuencias de nucleótidos específicas dentro de los segmentos funcionales. Las secuencias de nucleótidos específicas, denominadas en lo sucesivo secuencias afines, consisten normalmente, pero no exclusivamente, en tramos de 19-27 pares de bases. En algunos casos, el nucleótido de la posición central dentro de la secuencia afin desempeña un papel fundamental en el proceso de detección y, por lo tanto, en la determinación de la presencia frente a la ausencia de segmentos funcionales.

En ciertos casos, el método de la invención puede denominarse hibridación específica del alelo, y puede hacer uso de sondas oligonucleotídicas sintéticas, normalmente de 10 a 50 nucleótidos de longitud, preferentemente de 19 a 27 nucleótidos de longitud, cuyas secuencias están diseñadas para ser complementarias a la secuencia afin interrogada. La complementariedad de secuencias permite el apareamiento del ADN genómico y moléculas de sonda oligonucleotídica. Se puede hacer que se produzca el apareamiento específico, es decir, el apareamiento de las sondas con su secuencia afin y con ninguna otra secuencia, en condiciones apropiadas, que incluyen, pero sin limitación, el tiempo de incubación, la temperatura de incubación, la concentración de sonda y secuencias afines y la mezcla. El apareamiento específico a sondas permite la detección de secuencias afines en una mezcla de secuencias afines y no afines. La detección o la ausencia de detección de secuencias afines, a su vez, permite determinar la presencia frente a la ausencia de segmentos funcionales.

Las sondas oligonucleotídicas sintéticas pueden usarse para la detección de regiones no variantes conservadas y/o variantes alélicas particulares en el ADN genómico de un individuo. A menudo, las variantes alélicas son polimorfismos de un solo nucleótido (SNP), es decir, posiciones de nucleótidos en las que la composición de ADN puede variar entre individuos.

En algunos casos, las sondas oligonucleotídicas sintéticas descritas en el presente documento se diseñan y se usan para detectar la presencia o la ausencia de segmentos funcionales y, también, para detectar variantes alélicas ubicadas dentro de sus secuencias afines y para determinar la presencia o la ausencia de segmentos funcionales.

Dado un determinado nucleótido de una determinada posición de un locus de ADN genómico, se pueden diseñar moléculas oligonucleotídicas sintéticas, denominadas en lo sucesivo sondas, para detectar dicho nucleótido en una muestra de ensayo. Las sondas pueden diseñarse en pares de manera que un miembro del par de sondas sea complementario a una cadena de la secuencia afin, mientras que el otro miembro del par de sondas sea complementario a la otra cadena de la secuencia afin. Las sondas también pueden diseñarse en conjuntos de manera que tengan diferentes longitudes y sean complementarias a una cadena o las dos cadenas de la secuencia afin. El número de sondas usado para interrogar un segmento no variante es al menos uno y, a menudo, dos o cuatro.

El conjunto de al menos una, a menudo, dos o cuatro, sondas que interroga a una secuencia afín se denomina conjunto de sondas. Los parámetros evaluados en el diseño de las sondas con el fin de aumentar al máximo la sensibilidad y la especificidad de la detección de la secuencia afín son su composición de nucleótidos, su secuencia de nucleótidos, su longitud y la cadena de ADN con la que se complementan. Entre las secuencias diseñadas, el proceso de selección de sondas es en gran medida empírico.

Los SNP normalmente se definen por dos nucleótidos alternativos en una sola posición, donde cada uno de los dos nucleótidos alternativos define una de dos variantes alélicas.

Dado el polimorfismo de un solo nucleótido (SNP) en una posición particular de un locus de ADN genómico, y dadas dos variantes alélicas A y B definidas por el nucleótido particular en dicha posición, las sondas pueden diseñarse en pares para detectar la variante A y/o la variante B en una muestra de ensayo. El diseño es tal que un miembro del par de sondas es complementario en su totalidad a una cadena de la secuencia afín de la variante alélica A, y, en parte, a la secuencia afín de la variante alélica B, mientras que el otro miembro del par sonda es complementario en su totalidad a la misma cadena de la secuencia afín de la variante alélica B y, en parte, a la secuencia afín de la variante alélica A. En la mayoría de los casos, la complementariedad en parte se refiere a la complementariedad total, excepto en la posición de nucleótido central de la secuencia afín. Se pueden diseñar más de dos sondas, de modo que la complementariedad sea a una o a las dos cadenas de la secuencia afín y/o las sondas tengan diferentes longitudes.

Dado que la mayoría de los SNP consisten en dos variantes alélicas, el número de sondas usadas para interrogar un SNP es de al menos dos y, a menudo, de cuatro. Cuando se usan cuatro sondas, normalmente se complementan dos con cada variante alélica, siendo cada una de las dos para una variante alélica dada complementaria a cualquiera de las cadenas de ADN de la secuencia afín o comprendiendo cada una de las dos para una variante alélica dada una longitud de nucleótidos única. El conjunto de al menos dos, a menudo, cuatro, sondas que interroga a un SNP se denomina conjunto de sondas. Los parámetros evaluados en el diseño de las sondas con el fin de aumentar al máximo la sensibilidad y la especificidad de la detección de la secuencia afín son su composición de nucleótidos, su secuencia de nucleótidos, su longitud y la cadena de ADN con la que se complementan. Entre las secuencias diseñadas, el proceso de selección de sondas es en gran medida empírico.

Se puede usar al menos un conjunto de sondas para interrogar un segmento no variante y/o un SNP y, por lo tanto, se usa al menos un conjunto de sondas para determinar la presencia o la ausencia de un segmento funcional. Una determinación del genotipo, es decir, detección o no detección del segmento funcional, o detección de la presencia o la ausencia de un segmento funcional, se realiza a partir de valores numéricos que son una medida de la abundancia de secuencias afines en la muestra de ensayo. Para reducir al mínimo la variación entre ensayos en la medición de un solo segmento no variante y/o un SNP puede ser, lo que tiene el potencial de conducir a una determinación errónea, y a menudo es deseable usar más de un conjunto de sondas para realizar dicha determinación. El uso de más de un conjunto de sondas puede ayudar a evitar o a reducir al mínimo el efecto de los eventos genéticos tales como las inserciones, las eliminaciones, las conversiones, las estructuras 3D y similares. Como se contempla en el presente documento, no se aplica ninguna restricción en particular sobre la ubicación de múltiples secuencias afines en un segmento genético de interés, es decir, pueden estar superpuestas o no superpuestas, contiguas o no contiguas.

De acuerdo con cualquier aspecto de la presente invención, las sondas pueden estar unidas a un soporte sólido químicamente funcionalizado. Un ejemplo de un soporte sólido es una superficie de vidrio plana, sobre la que se colocan moléculas de sonda por deposición de contacto. Otro ejemplo de un soporte sólido es una perla de polímero de tamaño micrométrico, a la que se unen moléculas de sonda por conjugación. Otro ejemplo de un soporte sólido es una partícula de tamaño nanométrico a la que están unidas las moléculas de sonda. Una descripción ilustrativa en el presente documento se refiere al procedimiento realizado en el que las sondas se inmovilizan sobre una superficie de vidrio plana. La fijación de una sonda a la superficie se realiza en múltiples ubicaciones individuales (en el caso de las partículas, la fijación sería a múltiples conjuntos de partículas), en lo sucesivo denominadas características de replicación o "réplicas". El número de características de replicación para cada sonda suele ser de diez, aunque puede variar.

De acuerdo con cualquier aspecto de la presente invención, los segmentos funcionales o sus partes que abarcan secuencias afines pueden amplificarse, por ejemplo, mediante PCR, usando como molde ADN genómico en una muestra de referencia y/o en una muestra de ensayo. Los segmentos funcionales amplificados o sus partes pueden marcarse (por ejemplo, con un marcador fluorescente) para permitir su detección y, opcionalmente, fragmentarse para facilitar su apareamiento con las sondas oligonucleotídicas.

De acuerdo con cualquier aspecto de la presente invención, los segmentos funcionales marcados y fragmentados o sus partes pueden incubarse en condiciones que aumenten al máximo la sensibilidad y especificidad del apareamiento con las sondas unidas al soporte sólido. La cantidad de segmentos funcionales apareados con una sonda o sus partes puede determinarse indirectamente a partir de la medición del marcador, normalmente, un fluorocromo, unido al soporte sólido. Esta medición se denomina en el presente documento intensidad de señal. A modo de ejemplo, la fluorescencia emitida por el fluorocromo puede recogerse por medio de un dispositivo de

detección de la fluorescencia, tal como un escáner confocal. La intensidad de la señal conectada con cada característica de la sonda se puede cuantificar mediante un software específico y transformarse en valores numéricos que, por ejemplo, pueden variar entre 0 y 65.000.

5 En algunos casos de acuerdo con el método de cualquier aspecto de la presente invención, los valores de intensidad brutos de todas las características de la sonda que interrogan a una secuencia afin se pueden procesar para determinar la presencia o la ausencia de un segmento funcional de la siguiente manera: en una primera etapa, los valores de ruido de fondo, si están presentes, (B) se restan de los valores de intensidad de la señal (I) para cada una de las, por ejemplo, diez características de replicación con el fin de obtenerse valores de intensidad de la señal
10 corregidos (C).

$$C_n = I_n - B_n$$

15 donde n toma valores entre 1 y el número de características replicadas, por ejemplo, 10. Opcionalmente, en una segunda etapa, algunas, por ejemplo, una, normalmente dos de entre el número de características replicadas, por ejemplo, diez, características de replicación (R_{i1} , R_{i2}) cuyos valores de intensidad de la señal corregidos son los mayores (C_{i1} , C_{i2}) y algunas, por ejemplo, una, normalmente, dos del número de características de replicación, por ejemplo, diez características de replicación (R_{i9} , R_{i10}), cuyos valores de intensidad de la señal corregidos son los menores (C_{i9} , C_{i10}), se eliminan del análisis. Esta etapa también se puede describir como un "recorte" con, por ejemplo, un recorte del veinte por ciento en cada extremo de la distribución del conjunto de datos:

20 Valores de distribución:

$$C_{i1} > C_{i2} > C_{i3} > C_{i4} > C_{i5} > C_{i6} > C_{i7} > C_{i8} > C_{i9} > C_{i10}$$

25 Distribución de valores recortada:

$$C_{i3} > C_{i4} > C_{i5} > C_{i6} > C_{i7} > C_{i8}$$

Características de replicación desechadas:

$$R_{i1}, R_{i2}, R_{i9}, R_{i10}$$

30 En una tercera etapa, se obtiene una medida de tendencia central, por ejemplo, media aritmética, otras medias, mediana de la intensidad para un conjunto de sondas agrupando los valores de intensidad de la señal que corresponden a las características de replicación restantes. A modo de ejemplo, para un conjunto de sondas que consiste en cuatro sondas, i, j, k, l, se promedian veinticuatro valores, que corresponden al producto de seis
35 características de repetición restantes por cuatro sondas.

$$Media I = Media (I_{i3-i8}, I_{j3-j8}, I_{k3-k8}, I_{l3-l8})$$

40 En una cuarta etapa, se obtiene un fondo medio para un conjunto de sondas agrupando los valores de ruido de fondo (o un equivalente, por ejemplo, una medida de tendencia central, tal como media aritmética, otras medias, mediana de intensidades) que corresponden a las características de replicación restantes. Como en el ejemplo anterior:

$$Media B = Media (B_{i3-i8}, B_{j3-j8}, B_{k3-k8}, B_{l3-l8})$$

45 En una quinta etapa, la medida de la intensidad de señal de tendencia central, por ejemplo, la media aritmética, otras medias, mediana, se divide entre la medida de ruido de fondo de tendencia central, por ejemplo, media aritmética, otras medias, mediana, obteniéndose una proporción de intensidad/fondo (I/B) para cada conjunto de sondas, es decir, para cada polimorfismo.

$$I/B = Media I / Media B$$

55 Existen diferencias entre los conjuntos de sondas en la intensidad de la señal y el ruido de fondo que detectan tras la hibridación con sus secuencias afines, es decir, en sus proporciones de I/B. También existen diferencias entre los conjuntos de sondas en el hueco o intervalo entre las proporciones de I/B para un segmento genético cuando está ausente frente a cuando está presente. También existen diferencias entre los conjuntos de sondas en el grado de variación de sus proporciones de I/B, tanto entre ensayos como entre muestras. Como consecuencia, ciertos conjuntos de sondas son mejores indicadores de la presencia o la ausencia de un segmento genético que otros. Por

lo tanto, resulta deseable seleccionar entre los conjuntos de sondas que interrogan a segmentos no variantes y/o SNP en un segmento genético dado aquellos cuyas proporciones de I/B discernen mejor entre la ausencia y la presencia de dicho segmento.

5 La determinación de la presencia o la ausencia de un segmento genético en una muestra dada se realiza asignando dicha muestra a uno de dos grupos de posibles proporciones de I/B, en los que se han determinado previamente los intervalos para las posibles proporciones de I/B cuando el segmento genético en cuestión está presente y cuando el segmento genético en cuestión está ausente en un conjunto de muestras de referencia para las que se conoce la presencia frente a la ausencia de un segmento funcional. Los valores de proporción de I/B para una muestra dada se comparan con un límite de I/B inferior (L_{NC}) y un límite de I/B superior (U_{NC}) de la región de I/B "Sin determinación". La región de I/B "Sin determinación" corresponde a un intervalo de proporciones de I/B que no permiten la determinación de la presencia frente a la ausencia del segmento genético con suficiente confianza. Está situado entre la proporción de I/B mínima de muestras en las que se sabe que el segmento genético está presente [$p(I/B)_i$] y la proporción de I/B mínima de muestras en las que se sabe que el segmento genético está ausente [$a(I/B)_n$]. La determinación de un límite de I/B inferior (L_{NC}) y un límite de I/B superior (U_{NC}) de la región de I/B "Sin determinación" se explica con mayor detalle en el presente documento.

Los presentes inventores han encontrado que es posible derivar los límites L_{NC} y U_{NC} de manera que la región sin determinación sea menor que la diferencia absoluta entre [$p(I/B)_i$] y [$a(I/B)_n$], manteniendo todavía suficiente confianza estadística para clasificar una lectura de proporción de I/B de una muestra de ensayo como "presente" o "ausente" (véase, por ejemplo, la Figura 2, en la que el L_{NC} es superior a [$a(I/B)_n$] y el U_{NC} es inferior a [$p(I/B)_i$]). La derivación de una región sin determinación que es inferior a la diferencia absoluta entre [$p(I/B)_i$] y [$a(I/B)_n$] es ventajosa, porque permite reducir al mínimo el número de lecturas ambiguas "sin determinación", manteniendo la confianza en las lecturas que están fuera de la región sin determinación.

25

Selección de conjuntos de sondas óptimos

La presente invención proporciona un método de selección de conjuntos de sondas óptimos para determinar la presencia frente a la ausencia de un segmento genético entre los conjuntos de sondas que interrogan a segmentos no variantes y/o variantes situados en dicho segmento genético. En ciertos casos, de acuerdo con el método de la presente invención, la selección comprende las siguientes etapas.

30

En primer lugar, para cada conjunto de sondas, se crea una distribución de proporciones de I/B a partir de un conjunto de muestras de tamaño m , en el que se sabe que el segmento genético de interés está presente (p), y una distribución de proporciones de I/B de un conjunto de muestras de tamaño n , en el que se sabe que el segmento genético de interés está ausente (a):

35

$$p(I/B)_1 < p(I/B)_2 < \dots < p(I/B)_{m-1} < p(I/B)_m$$

$$a(I/B)_1 < a(I/B)_2 < \dots < a(I/B)_{n-1} < a(I/B)_n$$

40 En segundo lugar, para cada conjunto de sondas, se calcula la amplitud del intervalo (A) restando la mayor proporción de I/B en el conjunto de muestras en el que se sabe que el segmento genético de interés está ausente de la menor proporción de I/B de la muestra en la que se sabe que el segmento genético de interés está presente.

40

$$A = p(I/B)_i - a(I/B)_n$$

45

A puede adoptar valores negativos, cero o (más a menudo) positivos. El poder de un conjunto de sondas de diferenciar entre la ausencia y la presencia del segmento genético se correlaciona con el tamaño de A . Por lo tanto, el proceso de selección de conjuntos de sondas óptimos favorece aquellos conjuntos para los que A es máximo.

50 Basándose en su mayor A , se puede seleccionar cualquier número de conjuntos de sondas, a menudo, dos conjuntos de sondas óptimos para determinar la presencia frente a la ausencia del segmento genético. Se puede verificar el poder diferenciador óptimo de los conjuntos de sondas seleccionados comparando los conjuntos de sondas seleccionados con pares seleccionados aleatoriamente de conjuntos de sondas y/o con todos los conjuntos de sondas disponibles dentro del segmento genético de interés.

55

Opcionalmente, también se pueden realizar las siguientes etapas:

En tercer lugar, para cada conjunto de sondas, se calcula la medida de tendencia central, por ejemplo, media aritmética, otras medias, mediana, para la proporción de I/B en el conjunto de muestras en el que se sabe que el segmento genético de interés está presente (p Mediana I/B), y se calcula la medida de tendencia central, por ejemplo, media aritmética, otras medias, mediana para la proporción de I/B en el conjunto de muestras en el que

60

se sabe que el segmento genético de interés está ausente (*aMediana I/B*).

$$pMediana I/B = Mediana [p(I/B)_1, \dots, p(I/B)_m]$$

$$aMediana I/B = Mediana [a(I/B)_1, \dots, a(I/B)_n]$$

5 En cuarto lugar, para cada conjunto de sondas, se calcula la suma de la medida de tendencia central, por ejemplo, media aritmética, otras medias, mediana, para las proporciones de I/B (M), es decir, la proporción para el conjunto de muestras en el que se sabe que el segmento genético de interés está presente más la proporción para el conjunto de muestras en el que se sabe que el segmento genético de interés está ausente.

$$M = pMediana I/B + aMediana I/B$$

10 Dado que la *pMediana I/B* > 0 y la *aMediana I/B* > 0, *M* > 0. Para los conjuntos de sondas con un A dado, el poder de diferenciación entre la ausencia y la presencia del segmento genético disminuye a medida que M aumenta. Por lo tanto, el proceso de selección de conjuntos de sondas óptimos favorece aquellos conjuntos para los que M es mínimo.

15 Basándose en su menor M, se puede seleccionar cualquier número de conjuntos de sondas, a menudo, dos conjuntos de sondas óptimos para determinar la presencia frente a la ausencia del segmento genético. Se puede verificar el poder diferenciador óptimo de los conjuntos de sondas seleccionados comparando los conjuntos de sondas seleccionados con pares seleccionados aleatoriamente de conjuntos de sondas y/o con todos los conjuntos de sondas disponibles dentro del segmento genético de interés.

Opcionalmente, se puede realizar la siguiente etapa:

25 En quinto lugar, para cada conjunto de sondas, se calcula una puntuación (S) dividiendo la amplitud del intervalo (A) entre la suma de la medida de tendencia central, por ejemplo, media aritmética, otras medias, mediana para las proporciones de I/B (M):

$$S = \frac{A}{M} = \frac{p(I/B)_1 - a(I/B)_n}{pMediana IB + aMediana IB}$$

30 Dado que $p(I/B)_1 < pMediana I/B$, y dado que $a(I/B)_n > aMediana I/B$, $A < M$. Por lo tanto, S adopta valores absolutos entre 0 y 1. Dado que $M > 0$, cuando $A < 0$, $S < 0$, y cuando $A > 0$, $S > 0$.

35 Basándose en su mayor S, se puede seleccionar cualquier número de conjuntos de sondas, a menudo, dos conjuntos de sondas óptimos para determinar la presencia frente a la ausencia del segmento genético. Se puede verificar el poder diferenciador óptimo de los conjuntos de sondas seleccionados comparando los conjuntos de sondas seleccionados con pares seleccionados aleatoriamente de conjuntos de sondas y/o con todos los conjuntos de sondas del segmento genético de interés.

40 Para los segmentos genéticos para los que ningún par óptimo de conjuntos de sondas supere pares seleccionados aleatoriamente de conjuntos de sondas o todos los conjuntos de sondas, se puede usar uno cualquiera, más de uno o todos los conjuntos de sondas para determinar la presencia frente a la ausencia del segmento genético.

Establecimiento de los límites sin determinación

45 La presente invención proporciona un método de establecimiento de un límite de I/B inferior (L_{NC}) y un límite de I/B superior (U_{NC}) de la región de I/B "Sin determinación".

Cuando la amplitud del intervalo (A) mencionado anteriormente es positiva

$$A = p(I/B)_1 - a(I/B)_n$$

50 el límite de I/B inferior (L_{NC}) de la región de I/B "Sin determinación" puede ser igual a la mayor proporción de I/B del conjunto de muestras de referencia en el que se sabe que el segmento genético de interés está ausente ($L_{NC} = a(I/B)_n$), y el límite de I/B superior (U_{NC}) de la región de I/B "Sin determinación" puede ser igual a la menor proporción de I/B del conjunto de muestras de referencia en el que se sabe que el segmento de interés está presente ($U_{NC} =$

$p(I/B)_i$.

Ventajosamente, se puede realizar la siguiente etapa:

5 Para cada conjunto de sondas óptimo, se puede calcular un umbral de proporción de I/B ponderado ($T_{I/B}$).

$T_{I/B}$ es un punto de referencia situado entre las proporciones de I/B del conjunto de muestras en el que se sabe que el segmento genético de interés está presente y las proporciones de I/B del conjunto de muestras en el que se sabe que el segmento genético de interés es ausente.

10 El valor de $T_{I/B}$ se determina a partir de la medida de tendencia central, por ejemplo, media aritmética, otras medias, mediana y proporciones de I/B mínimas de las muestras en las que se sabe que el segmento genético está presente [$pMediana\ I/B$, $p(I/B)_i$] y la medida de la tendencia central, por ejemplo, media aritmética, otras medias, mediana y proporciones de I/B máximas de las muestras en las que se sabe que el segmento genético está ausente [$aMediana\ I/B$, $a(I/B)_n$].

15 En concreto, se puede establecer el valor de $T_{I/B}$ de manera que las distancias de $T_{I/B}$ a $pMediana\ I/B$ y $aMediana\ I/B$ sean proporcionales a las distancias de $p(I/B)_i$ y $a(I/B)_n$ a la $pMediana\ I/B$.

$$20 \quad \frac{pMediana\ I/B - T_{I/B}}{pMediana\ I/B - p(I/B)_i} = \frac{T_{I/B} - aMediana\ I/B}{a(I/B)_n + aMediana\ I/B}$$

Por lo tanto,

$$25 \quad T_{I/B} = \frac{pMediana\ I/B \cdot a(I/B)_n - aMediana\ I/B \cdot p(I/B)_i}{pMediana\ I/B + a(I/B)_n - aMediana\ I/B - p(I/B)_i}$$

Los límites de I/B inferior (L_{NC}) y superior (U_{NC}) de la región de I/B "Sin determinación" pueden ser iguales al umbral de proporción de I/B ponderado ($T_{I/B}$).

30 Opcionalmente, se puede realizar la siguiente etapa adicional: En primer lugar, para cada conjunto de sondas óptimo, se calcula el límite superior de la región de I/B "Sin determinación" (U_{NC}).

35 La región de I/B "Sin determinación" corresponde a un intervalo de proporciones de I/B que no permiten la determinación de la presencia frente a la ausencia del segmento genético con suficiente confianza. Está situado entre la proporción de I/B mínima de muestras en las que se sabe que el segmento genético está presente [$p(I/B)_i$] y la proporción máxima de I/B de muestras en las que se sabe que el segmento genético está ausente [$a(I/B)_n$].

40 El valor de U_{NC} se calcula a partir de la medida de la tendencia central, por ejemplo, la media aritmética, otras medias, la mediana y las proporciones mínimas de I/B de muestras en las que se sabe que el segmento genético está presente [$pMediana\ I/B$, $p(I/B)_i$]. En concreto, U_{NC} es tal que las distancias de U_{NC} a $T_{I/B}$ y de $p(I/B)_i$ a $T_{I/B}$ son proporcionales a las distancias de $p(I/B)_i$ a $T_{I/B}$ y $pMediana\ I/B$ a $T_{I/B}$.

$$\frac{U_{NC} - T_{I/B}}{p(I/B)_i - T_{I/B}} = \frac{p(I/B)_i - T_{I/B}}{pMediana\ I/B - T_{I/B}}$$

45 Por lo tanto,

$$U_{NC} = \frac{T_{I/B} \cdot [pMediana\ I/B - 2 \cdot p(I/B)_i] + p(I/B)_i^2}{pMediana\ I/B - T_{I/B}}$$

50 En segundo lugar, para cada conjunto de sondas óptimo, se calcula el límite inferior de la región de I/B "Sin determinación" (L_{NC}). El valor de L_{NC} se calcula a partir de la medida de la tendencia central, por ejemplo, la media aritmética, otras medias, la mediana y las proporciones máximas de I/B de muestras en las que se sabe que el segmento genético está ausente [$aMediana\ I/B$, $a(I/B)_n$]. En concreto, L_{NC} es tal que las distancias de L_{NC} a $T_{I/B}$ y de $p(I/B)_n$ a $T_{I/B}$ son proporcionales a las distancias de $a(I/B)_n$ a $T_{I/B}$ y $aMediana\ I/B$ a $T_{I/B}$.

$$\frac{T_{I/B} - L_{NC}}{T_{I/B} - a(I/B)_n} = \frac{T_{I/B} - a(I/B)_n}{T_{I/B} - aMediana I/B}$$

Por lo tanto,

$$L_{NC} = \frac{T_{I/B} \cdot [2 \cdot a(I/B)_n - aMediana I/B] - a(I/B)_n^2}{T_{I/B} - aMediana I/B}$$

5

Determinación de la presencia frente a la ausencia de un segmento genético en una muestra de ensayo

10 La presente invención proporciona, en ciertos aspectos, un método de establecimiento de la presencia frente a la ausencia de un segmento genético en una muestra de ensayo mediante proporciones de I/B a partir de conjuntos de sondas óptimos.

15 La determinación de la presencia frente a la ausencia del segmento genético de interés en una muestra de ensayo puede realizarse de acuerdo con los valores de las proporciones de I/B de uno o más conjuntos de sondas óptimos (PS_1 , PS_2 , etc.) para esa muestra y los valores de U_{NC} , L_{NC} .

20 Por ejemplo, cuando se usa solo un conjunto de sondas (PS_1), se determina que el segmento genético de interés está presente cuando I/B de $PS_1 > U_{NC}$, se determina que está ausente cuando I/B de $PS_1 < L_{NC}$, y no se realiza ninguna determinación cuando $U_{NC} \geq I/B$ de $PS_1 \geq L_{NC}$.

25 Por ejemplo, cuando se usan dos o más conjuntos de sondas, puede determinarse la presencia o la ausencia del segmento genético de interés: a) mediante un consenso estricto de una determinación de presente o de una determinación de ausente para todos los conjuntos de sondas usados (Tabla 1: ejemplo para dos conjuntos de sondas); o b) mediante un consenso mayoritario sin determinaciones contradictorias para todos los conjuntos de sondas usados (Tabla 2: ejemplo para dos conjuntos de sondas).

Tabla 1 - consenso estricto

Valor de I/B de PS_1	Valor de I/B de PS_2	Determinación de segmento genético
I/B de $PS_1 > U_{NC}$	I/B de $PS_2 > U_{NC}$	Presente
I/B de $PS_1 < L_{NC}$	I/B de $PS_2 < L_{NC}$	Ausente
$U_{NC} \geq I/B$ de $PS_1 \geq L_{NC}$	$U_{NC} \geq I/B$ de $PS_2 \geq L_{NC}$	Sin determinación
$U_{NC} \geq I/B$ de $PS_1 \geq L_{NC}$	I/B de $PS_2 > U_{NC}$	Sin determinación
$U_{NC} \geq I/B$ de $PS_1 \geq L_{NC}$	I/B de $PS_2 < L_{NC}$	Sin determinación
I/B de $PS_1 > U_{NC}$	$U_{NC} \geq I/B$ de $PS_2 \geq L_{NC}$	Sin determinación
I/B de $PS_1 < L_{NC}$	$U_{NC} \geq I/B$ de $PS_2 \geq L_{NC}$	Sin determinación

30 La Tabla 1 muestra la determinación de la presencia o la ausencia del segmento genético de interés en una muestra mediante un consenso estricto de bien una determinación de presente o una determinación de ausente cuando se usan dos conjuntos de sondas.

Tabla 2 – Consenso mayoritario

Valor de I/B de PS_1	Valor de I/B de PS_2	Determinación de segmento genético
I/B de $PS_1 > U_{NC}$	I/B de $PS_2 > U_{NC}$	Presente
I/B de $PS_1 < L_{NC}$	I/B de $PS_2 < L_{NC}$	Ausente
$U_{NC} \geq I/B$ de $PS_1 \geq L_{NC}$	$U_{NC} \geq I/B$ de $PS_2 \geq L_{NC}$	Sin determinación
$U_{NC} \geq I/B$ de $PS_1 \geq L_{NC}$	I/B de $PS_2 > U_{NC}$	Presente
$U_{NC} \geq I/B$ de $PS_1 \geq L_{NC}$	I/B de $PS_2 < L_{NC}$	Ausente
I/B de $PS_1 > U_{NC}$	$U_{NC} \geq I/B$ de $PS_2 \geq L_{NC}$	Presente

$I/B \text{ de } PS_1 < LNC$	$UNC \geq I/B \text{ de } PS_2 \geq LNC$	Ausente
------------------------------	--	---------

La Tabla 2 muestra la determinación de la presencia o la ausencia del segmento genético de interés en una muestra mediante un consenso mayoritario sin determinaciones contradictorias de bien una determinación de presente o una determinación de ausente cuando se usan dos conjuntos de sondas.

La siguiente información se presenta a modo de ejemplos y no pretende ser una limitación del alcance de las reivindicaciones.

Ejemplo

Determinación del exón 5 del gen RHD

El siguiente ejemplo se refiere a un método de genotipificación de grupos sanguíneos humanos, tales como Rhesus (RH). El método descrito en el presente documento se ha aplicado a 11 conjuntos de sondas en la genotipificación de grupos sanguíneos que interrogan a variantes del exón 05 de RHD (SNP). El proceso descrito a continuación procede del análisis de dos conjuntos de muestras, uno consistente en 197 muestras en las que se sabe que el exón 05 de RHD está presente, y el otro consiste en 161 muestras en las que se sabe que el exón 05 de RHD está ausente.

Materiales y métodos

De acuerdo con el presente ejemplo, se tomaron muestras que contenían ácido nucleico de la sangre. La extracción del ADN se llevó a cabo de acuerdo con el protocolo proporcionado por el fabricante para la purificación de ADN Genómico ("QIAamp[®] DNA Blood Mini o Midi Kit"). El ADN purificado se cuantificó en un aparato de medición (espectrofotómetro) para verificar la concentración y la proporción de A260/A280 (1,6-1,95).

Se sometieron las regiones de ADN que pueden contener los segmentos genéticos cuya presencia o ausencia ha de determinarse a una reacción de amplificación para obtener productos de amplificación. Puede usarse cualquier técnica o método adecuado para la amplificación. Se llevaron a cabo PCR multiplex, usando pares apropiados de cebadores de PCR de oligonucleótidos. A 38 ciclos de 3 etapas (desnaturalización durante 45 s a 95 °C, hibridación durante 60 s a 60 °C y extensión durante 90 s a 72 °C) le siguió una extensión final a 72 °C durante 10 min.

A continuación, se fragmentaron los productos de PCR usando una mezcla de fragmentación que contenía EDTA 0,32 mM, 0,18 U/μl de fosfatasa alcalina y 04 U/μl de ADNasa I, incubando a 37 °C durante 30 minutos, seguido de la inactivación por calor a 95 °C durante 10 minutos.

Después se sometieron los productos de fragmentación a marcaje indirecto con Biotina-11-ddUTP usando técnicas convencionales mediante incubación a 37 °C durante 60 minutos.

Entonces, se pusieron en contacto los productos de fragmentación marcados con sondas oligonucleotídicas, que son capaces de detectar los segmentos genéticos correspondientes mediante hibridación en condiciones adecuadas, fijadas sobre un soporte sólido. Por lo general, las condiciones de hibridación permiten una hibridación específica entre las sondas y los ácidos nucleicos diana correspondientes para formar complejos de hibridación de sonda/diana específicos, reduciendo al mínimo al mismo tiempo la hibridación entre las sondas portadoras de uno o más desapareamientos al ADN. La hibridación tuvo lugar a temperatura ambiente, bajo una presión de N₂ de 275,79 kPa (40 psi [2,7 bar]), durante 90 minutos. Los conjuntos de sondas usados fueron como se muestran en la siguiente Tabla 3.

Tabla 3

NÚMERO DE CONJUNTO DE SONDAS	SECUENCIA DE LA SONDA	SEQ ID NO:
1	AGAAGTCCAATCGAAAGGAAGAATG	102
1	AGAAGTCCAATCTAAAGGAAGAATG	103
1	GAAGTCCAATCGAAAGGAAGAAT	104
1	GAAGTCCAATCTAAAGGAAGAAT	105
2	CAAGGACTATCAGGCCACGGGGTCA	106
2	CAAGGACTATCAGCCCACGGGGTCA	107
2	TGACCCCGTGGCCTGATAGTCCTTG	108
2	TGACCCCGTGGGCTGATAGTCCTTG	109

ES 2 641 600 T3

3	CTGGCCCCCAGGCGCCCTCTTCT	110
3	CTGGCCCCCAGTCGCCCTCTTCT	111
3	CTGGCCCCCAGGCGCCCTCTTCT	110
3	CTGGCCCCCAGTCGCCCTCTTCT	111
4	CTGGCCAAGTTTCAACTCTGC	112
4	CTGGCCAAGTGTCAACTCTGC	113
4	TGGCCAAGTTTCAACTCTG	114
4	TGGCCAAGTGTCAACTCTG	115
5	AGTTTCAACTCTGCTCTGCTGAGAA	116
5	AGTTTCAACTCTCCTCTGCTGAGAA	117
5	AGTTTCAACTCTGCTCTGCTGAGAA	116
5	AGTTTCAACTCTCCTCTGCTGAGAA	117
6	CTGCTCTGCTGAGAAGTCCAATCGA	118
6	CTGCTCTGCTGAAAAGTCCAATCGA	119
6	TGCTCTGCTGAGAAGTCCAATCG	120
6	TGCTCTGCTGAAAAGTCCAATCG	121
7	AGAAGTCCAATCGAAAGGAAGAATG	102
7	AGAAGTCCAATCCAAAGGAAGAATG	122
7	GAAGTCCAATCGAAAGGAAGAAT	104
7	GAAGTCCAATCCAAAGGAAGAAT	123
8	AGGAAGAATGCCGTGTTCAACAC	26
8	AGGAAGAATGCCGTGTTCAACACC	27
8	AGGAAGAATGCCGTGTTCAACAC	26
8	AGGAAGAATGCCGTGTTCAACACC	27
9	GGAAGAATGCCGTGTTCAACACC	124
9	GGAAGAATGCCATGTTCAACACC	125
9	GAAGAATGCCGTGTTCAACAC	126
9	GAAGAATGCCATGTTCAACAC	127
10	GGCTCACCCCCAAGGGAAGGGAAGA	28
10	GGCTCACCCCCAGGGAAGGGAAGAT	29
10	GGCTCACCCCCAAGGGAAGGGAAGA	28
10	GGCTCACCCCCAGGGAAGGGAAGAT	29
11	AGAAGTCCAATCGAAAGGAAGAATG	102
11	AGAAGTCCAATCAAAGGAAGAATG	128
11	GAAGTCCAATCGAAAGGAAGAAT	104
11	GAAGTCCAATCAAAGGAAGAAT	129

Tras la hibridación, se añadió un conjugado de estreptavidina-Cy3 para completar el marcaje indirecto. La detección y cuantificación de la señal de fluorescencia se realizó usando un escáner Innopsys® 700A/710.

ES 2 641 600 T3

5	<p>18,7; 19,0; 19,0; 19,0; 19,2; 19,3; 19,3; 19,3; 19,4; 19,5; 19,6; 19,8; 19,8; 19,9; 20,0; 20,0; 20,3; 20,3; 20,3; 20,4; 20,5; 20,5; 20,5; 20,5; 20,7; 20,9; 20,9; 20,9; 21,0; 21,0; 21,0; 21,1; 21,1; 21,1; 21,2; 21,3; 21,4; 21,7; 21,9; 21,9; 21,9; 21,9; 22,0; 22,4; 22,4; 22,9; 22,9; 23,0; 23,2; 23,4; 23,5; 23,8; 23,8; 24,1; 24,1; 24,1; 24,4; 24,8; 25,0; 25,3; 25,5; 25,9; 25,9; 26,0; 27,0; 27,4; 27,6; 28,2; 28,3; 29,0; 29,4; 30,5; 30,7; 30,9; 31,2; 31,3; 31,5; 31,9; 31,9; 32,1; 32,1; 32,4; 32,6; 32,6; 32,8; 33,1; 33,3; 33,4; 33,9; 34,4; 34,8; 35,0; 35,0; 35,3; 35,3; 35,4; 35,6; 35,6; 35,8; 35,9; 36,0; 36,1; 36,2; 36,3; 36,4; 36,6; 36,8; 36,9; 37,0; 37,1; 37,1; 37,4; 37,6; 38,0; 38,1; 38,3; 38,6; 39,2; 39,6; 40,1; 40,4; 40,5; 40,5; 40,8; 41,3; 42,1; 42,1; 42,3; 42,7; 43,2; 44,4; 45,8; 45,9; 45,9; 47,6; 54,9; 57,3; 58,1.</p>
10	<p>p_{PS7}: 9,6; 9,7; 10,2; 11,3; 11,6; 12,2; 12,6; 12,6; 12,7; 12,9; 13,3; 13,3; 13,7; 14,0; 14,1; 14,2; 14,2; 14,3; 14,5; 14,5; 14,6; 14,8; 14,9; 14,9; 15,1; 15,4; 15,4; 15,5; 15,5; 16,1; 16,2; 16,2; 16,3; 16,4; 16,4; 16,5; 16,8; 16,9; 17,1; 17,1; 17,1; 17,2; 17,3; 17,5; 17,5; 17,5; 17,6; 17,7; 17,7; 17,8; 17,8; 17,9; 18,0; 18,0; 18,0; 18,3; 18,3; 18,4; 18,4; 18,4; 18,5; 18,5; 18,6; 18,7; 18,7; 19,0; 19,0; 19,1; 19,2; 19,3; 19,5; 19,6; 19,6; 19,7; 19,7; 19,9; 19,9; 20,0; 20,1; 20,1; 20,1; 20,3; 20,3; 20,4; 20,5; 20,6; 20,8; 20,9; 20,9; 21,1; 21,1; 21,2; 21,5; 21,6; 21,7; 21,9; 22,1; 22,4; 22,4; 22,5; 22,5; 22,5; 22,7; 22,9; 23,1; 23,1; 23,1; 23,5; 23,7; 23,7; 23,7; 23,9; 24,2; 24,2; 24,6; 24,6; 25,3; 25,5; 25,6; 26,1; 26,4; 26,8; 28,4; 29,6; 29,7; 30,1; 30,1; 30,1; 30,7; 30,8; 31,1; 31,4; 31,4; 31,6; 32,1; 32,4; 32,6; 32,7; 33,0; 33,0; 33,2; 33,5; 33,5; 33,9; 34,1; 34,3; 34,6; 34,7; 34,7; 34,7; 34,7; 34,8; 35,3; 35,4; 35,7; 35,8; 36,0; 36,0; 36,0; 36,1; 36,2; 36,4; 36,6; 36,8; 36,8; 37,4; 37,6; 38,1; 38,2; 38,4; 39,0; 39,1; 39,4; 40,1; 40,2; 40,4; 40,5; 40,8; 41,3; 41,3; 41,6; 42,3; 42,6; 43,1; 43,5; 43,5; 44,1; 45,1; 46,6; 49,8; 50,3; 51,3; 56,4; 57,1; 59,6.</p>
15	
20	<p>p_{PS8}: 9,7; 10,1; 11,2; 12,2; 12,6; 12,9; 13,2; 13,4; 13,5; 13,6; 13,7; 13,7; 14,1; 14,1; 14,2; 14,4; 14,5; 14,5; 14,6; 14,8; 14,9; 15,0; 15,1; 15,3; 15,3; 15,3; 15,5; 15,7; 15,8; 16,1; 16,2; 16,2; 16,5; 16,9; 17,0; 17,0; 17,1; 17,2; 17,3; 17,6; 17,6; 17,8; 17,8; 17,8; 17,9; 18,0; 18,0; 18,2; 18,2; 18,4; 18,5; 18,6; 18,8; 19,1; 19,3; 19,3; 19,3; 19,3; 19,4; 19,4; 19,5; 19,7; 19,7; 19,8; 20,1; 20,2; 20,3; 20,3; 20,4; 20,4; 20,4; 20,7; 20,8; 20,9; 20,9; 21,0; 21,1; 21,1; 21,1; 21,2; 21,2; 21,2; 21,3; 21,4; 21,4; 21,5; 21,7; 21,9; 22,0; 22,0; 22,2; 22,2; 22,2; 22,5; 22,5; 22,8; 22,9; 23,0; 23,1; 23,2; 23,3; 23,6; 23,7; 23,7; 23,8; 23,8; 24,0; 24,1; 24,3; 24,3; 24,4; 24,8; 25,0; 25,2; 25,5; 25,9; 25,9; 26,0; 26,1; 26,3; 26,4; 26,8; 26,8; 27,6; 27,7; 28,1; 28,6; 29,8; 29,8; 30,1; 30,4; 30,7; 30,9; 31,8; 31,8; 31,9; 32,1; 32,3; 32,8; 32,9; 33,0; 33,4; 33,4; 33,5; 33,5; 33,6; 34,0; 34,3; 34,5; 34,6; 34,6; 34,9; 35,0; 35,4; 35,5; 35,7; 35,8; 35,8; 36,1; 36,2; 36,2; 36,2; 36,4; 36,5; 36,8; 37,4; 37,4; 37,4; 37,6; 38,0; 38,1; 38,1; 38,1; 38,2; 38,2; 38,6; 38,7; 39,0; 39,3; 39,5; 39,9; 40,2; 40,3; 40,4; 40,6; 41,0; 41,2; 42,3; 42,5; 44,6; 44,7; 46,4; 48,4; 49,8; 51,0; 51,2; 53,3.</p>
25	
30	<p>p_{PS9}: 20,7; 21,1; 22,9; 23,4; 23,8; 25,2; 26,1; 27,1; 27,2; 27,5; 28,8; 28,9; 29,5; 29,7; 29,7; 30,1; 30,3; 30,4; 30,6; 30,6; 30,9; 30,9; 31,0; 31,3; 31,6; 31,6; 31,7; 31,8; 31,8; 32,3; 32,4; 32,5; 33,0; 33,2; 33,8; 34,5; 35,1; 35,1; 35,6; 35,6; 36,1; 36,2; 36,2; 36,7; 36,7; 36,8; 37,1; 37,1; 37,1; 38,1; 38,4; 38,5; 38,6; 38,9; 39,1; 39,3; 39,5; 39,6; 39,6; 39,7; 39,9; 39,9; 39,9; 40,0; 40,0; 40,3; 40,4; 40,8; 40,9; 41,0; 41,2; 41,5; 41,5; 41,6; 41,8; 42,0; 42,0; 42,4; 42,5; 42,9; 42,9; 42,9; 42,9; 43,0; 43,1; 43,2; 43,2; 43,3; 43,6; 43,6; 43,8; 44,0; 44,1; 44,2; 44,4; 45,4; 45,6; 45,7; 45,7; 45,8; 46,2; 46,4; 46,8; 46,9; 47,4; 47,5; 47,7; 47,9; 47,9; 48,5; 48,6; 48,8; 48,9; 48,9; 49,2; 49,4; 50,4; 50,7; 50,8; 51,0; 51,3; 51,6; 52,9; 53,3; 53,4; 53,5; 53,8; 54,1; 54,2; 55,2; 55,7; 56,5; 57,0; 57,5; 58,7; 59,0; 59,4; 59,6; 60,6; 61,5; 61,6; 61,8; 62,6; 62,9; 63,0; 63,3; 63,5; 64,0; 64,2; 64,6; 64,9; 65,1; 65,3; 65,3; 65,3; 65,7; 65,7; 65,8; 65,8; 66,4; 67,7; 67,7; 67,8; 67,8; 68,5; 69,8; 69,9; 69,9; 70,2; 71,3; 71,4; 71,7; 73,0; 73,0; 73,3; 73,5; 73,5; 73,9; 74,9; 74,9; 75,8; 75,8; 75,8; 76,6; 76,6; 76,8; 77,7; 78,4; 78,9; 80,2; 86,0; 91,1; 91,3; 94,5; 96,2; 99,8; 108,6.</p>
35	
40	
45	<p>p_{PS10}: 17,7; 16,4; 16,9; 18,6; 20,3; 21,0; 21,2; 21,8; 21,9; 22,5; 22,5; 22,5; 22,6; 23,3; 23,3; 23,9; 24,1; 24,5; 24,6; 24,8; 25,0; 25,4; 25,4; 25,5; 25,5; 25,6; 25,8; 26,2; 26,4; 26,5; 26,8; 26,8; 27,7; 27,8; 28,0; 28,0; 28,2; 28,3; 28,6; 28,7; 29,4; 29,6; 29,6; 29,7; 29,8; 29,9; 30,0; 30,0; 30,2; 30,3; 30,3; 30,3; 30,4; 30,7; 30,9; 30,9; 31,0; 31,0; 31,1; 31,4; 31,5; 31,6; 31,7; 31,9; 32,1; 32,3; 32,4; 32,6; 32,6; 32,8; 32,9; 32,9; 33,1; 33,2; 33,3; 33,4; 33,6; 33,6; 33,7; 34,0; 34,0; 34,2; 34,3; 34,5; 34,7; 34,9; 35,0; 35,1; 35,2; 35,4; 35,5; 35,8; 35,8; 35,9; 36,2; 36,7; 37,1; 37,3; 37,4; 37,4; 37,5; 37,5; 37,7; 38,2; 38,3; 38,3; 38,4; 38,4; 38,8; 38,9; 39,1; 39,2; 39,4; 39,6; 39,7; 39,7; 39,9; 40,1; 40,4; 40,8; 40,8; 41,1; 41,7; 42,3; 43,7; 43,9; 43,9; 44,2; 44,4; 44,6; 45,0; 45,6; 45,7; 46,0; 46,2; 47,2; 47,2; 47,8; 47,8; 47,9; 47,9; 48,0; 48,3; 48,3; 48,5; 48,6; 48,7; 49,0; 49,0; 49,7; 49,9; 50,0; 50,2; 50,8; 50,9; 51,1; 51,3; 51,7; 52,6; 52,9; 53,0; 53,1; 53,2; 53,7; 53,8; 54,1; 54,3; 54,4; 54,6; 55,0; 55,1; 55,1; 55,7; 55,8; 56,2; 56,7; 56,8; 56,9; 57,3; 58,2; 58,7; 59,0; 59,3; 59,7; 60,1; 62,3; 62,3; 63,1; 63,7; 64,9; 65,2; 67,6; 72,8; 76,0; 77,0; 77,7; 80,0.</p>
50	
55	<p>p_{PS11}: 21,0; 17,4; 17,8; 18,0; 18,5; 18,9; 19,3; 20,2; 20,3; 21,0; 21,2; 21,3; 21,6; 22,0; 22,0; 22,6; 22,7; 22,8; 23,0; 23,1; 23,3; 23,5; 23,5; 23,5; 23,6; 23,7; 23,7; 23,8; 24,0; 24,0; 24,0; 24,1; 24,4; 24,7; 24,7; 24,9; 24,9; 25,2; 25,2; 25,7; 25,7; 25,8; 25,8; 25,8; 26,6; 26,6; 26,7; 26,7; 27,4; 27,6; 27,7; 27,8; 27,8; 27,9; 28,1; 28,1; 28,4; 28,5; 28,6; 28,7; 28,7; 29,2; 29,5; 29,6; 29,7; 29,9; 29,9; 30,2; 30,3; 30,8; 30,9; 30,9; 31,1; 31,2; 31,4; 31,4; 31,5; 31,7; 31,9; 32,1; 32,2; 32,3; 32,4; 32,5; 32,7; 32,8; 32,8; 33,4; 33,4; 33,9; 34,0; 34,1; 34,2; 34,4; 34,6; 34,7; 34,8; 34,9; 34,9; 35,0; 35,3; 35,6; 35,6; 35,7; 35,7; 36,4; 36,5; 36,7; 36,7; 36,9; 37,1; 37,1; 37,3; 37,4; 37,5; 38,1; 38,2; 38,2; 38,4; 38,7; 38,9; 38,9; 38,9; 39,5; 39,5; 39,8; 39,9; 40,0; 40,1; 40,2; 40,3; 40,3; 40,4; 40,6; 40,8; 41,0; 41,0; 41,9; 43,0; 43,0; 43,1; 43,2; 43,2; 43,7; 43,8; 44,1; 44,3; 44,4; 45,3; 45,4; 45,6; 45,9; 46,0; 46,0; 46,2; 46,5; 46,5; 46,7; 47,1; 47,7; 47,7; 47,8; 47,8; 48,3; 48,6; 49,6; 49,8; 50,3; 50,5; 50,9; 51,3; 51,5; 51,5; 51,7; 51,8; 52,0; 52,1; 52,4; 52,4; 52,6; 52,7; 53,0; 53,6; 54,3; 55,0; 57,8; 58,1; 58,5; 58,7; 65,6; 65,9; 68,5; 68,8; 70,4; 72,6; 78,7; 81,4.</p>
60	

65 Se derivó la siguiente distribución de las proporciones de I/B para cada uno de los 11 conjuntos de sondas ($PS1-PS11$) en un conjunto de muestras de tamaño $m = 161$, en el que se sabe que está ausente el exón 05 de RHD (a)

- $p(\text{Mediana I/B})_{PS3} = 26,7$
- $p(\text{Mediana I/B})_{PS4} = 56,1$
- $p(\text{Mediana I/B})_{PS5} = 48,8$
- $p(\text{Mediana I/B})_{PS6} = 21,9$
- $p(\text{Mediana I/B})_{PS7} = 22,4$
- $p(\text{Mediana I/B})_{PS8} = 23,1$
- $p(\text{Mediana I/B})_{PS9} = 45,7$
- $p(\text{Mediana I/B})_{PS10} = 37,4$
- $p(\text{Mediana I/B})_{PS11} = 34,9.$

Después, se determina una medida de tendencia central, por ejemplo, la media aritmética, otras medias o la mediana para la proporción de I/B para cada conjunto de sondas del conjunto de muestras en el que se sabe que está ausente el exón 05 de RHD. En este caso, la medida de la tendencia central fue la mediana ($p\text{Mediana I/B}$)

$$a\text{Mediana I/B} = \text{Mediana} [p(I/B)_1, \dots, p(I/B)_m]$$

- $a(\text{Mediana I/B})_{PS1} = 1,4$
- $a(\text{Mediana I/B})_{PS2} = 5,3$
- $a(\text{Mediana I/B})_{PS3} = 1,9$
- $a(\text{Mediana I/B})_{PS4} = 3,7$
- $a(\text{Mediana I/B})_{PS5} = 2,7$
- $a(\text{Mediana I/B})_{PS6} = 1,3$
- $a(\text{Mediana I/B})_{PS7} = 1,2$
- $a(\text{Mediana I/B})_{PS8} = 1,2$
- $a(\text{Mediana I/B})_{PS9} = 1,5$
- $a(\text{Mediana I/B})_{PS10} = 2,0$
- $a(\text{Mediana I/B})_{PS11} = 1,1.$

A continuación, se determinó la suma de la medida de tendencia central, por ejemplo, la media aritmética, otras medias o la mediana para las proporciones de I/B (en este caso, la suma de las medianas (M)), es decir, la proporción para el conjunto de muestras en el que se sabe que el exón 05 de RHD está presente más la proporción para el conjunto de muestras en el que se sabe que el exón 05 de RHD está ausente, mediante el conjunto de sondas.

$$M = p\text{Mediana I/B} + a\text{Mediana I/B}$$

- $M_{PS1} = 5,4+1,4 = 6,8$
- $M_{PS2} = 65,3+5,3 = 70,6$
- $M_{PS3} = 26,7+1,9 = 28,6$
- $M_{PS4} = 56,1+3,7 = 59,8$
- $M_{PS5} = 48,1+2,7 = 51,5$
- $M_{PS6} = 21,9+1,3 = 23,2$
- $M_{PS7} = 22,4+1,6 = 24,1$
- $M_{PS8} = 23,1+1,2 = 24,4$
- $M_{PS9} = 45,7+1,5 = 47,2$
- $M_{PS10} = 37,4+2,0 = 39,4$
- $M_{PS11} = 34,9+1,1 = 36,0.$

A continuación, se determinó la puntuación del conjunto de sondas (S) para cada conjunto de sondas dividiendo la amplitud del intervalo (A) entre la suma de la medida de la tendencia central, por ejemplo, la media aritmética, otras medias o la mediana para las proporciones de I/B (M) (en este caso, la medida de tendencia central fue la mediana):

$$S = A/M = \frac{p(I/B)_1 - a(I/B)_m}{p\text{Mediana I/B} + a\text{Mediana I/B}}$$

- $S_{PS1} = -2,2/6,8 = -0,33$
- $S_{PS2} = -10,3/70,6 = -0,15$
- $S_{PS3} = 0,0/28,6 = 0,00$
- $S_{PS4} = -4,4/59,8 = -0,07$
- $S_{PS5} = 2,1/51,5 = 0,04$
- $S_{PS6} = 4,5/23,2 = 0,19$
- $S_{PS7} = 1,85/24,1 = 0,08$

$$S_{PS8} = 5,15/24,4 = 0,21$$

$$S_{PS9} = 14,81/47,2 = 0,31$$

$$S_{PS10} = 2,99/39,4 = 0,08$$

$$S_{PS11} = 12,94/36,0 = 0,36.$$

5 Basándose en sus puntuaciones más altas, se seleccionaron dos conjuntos de sondas, *PS9* y *PS11* para la determinación de la presencia frente a la ausencia del exón 05 de RHD. Sin embargo, se puede seleccionar un número diferente de conjuntos de sondas (esencialmente cualquier número) igual o superior a uno.

10 Las Figuras 1A y 1B muestran, para conjuntos de sondas óptimos *PS9* y *PS11* (ID de genotipificación de grupo sanguíneo BCV053 y BCV055, respectivamente), la frecuencia de las muestras en el conjunto de muestras de referencia en función de su proporción de I/B.

15 A continuación, se determinó el umbral de la proporción de I/B ponderado ($T_{I/B}$) que separa las proporciones de I/B de las muestras en las que se sabe que el exón 05 de RHD está presente de las proporciones de I/B de las muestras en las que se sabe que el exón 05 de RHD está ausente, mediante el conjunto óptimo de sondas de acuerdo con la siguiente fórmula:

$$T_{I/B} = \frac{p_{Mediana\ I/B} \cdot a(I/B)_n - a_{Mediana\ I/B} \cdot p(I/B)_\lambda}{p_{Mediana\ I/B} + a(I/B)_n - a_{Mediana\ I/B} - p(I/B)_\lambda}$$

$$T_{I/B\ PS9} = \frac{45,7 \cdot 5,9 - 1,5 \cdot 20,7}{45,7 + 5,9 - 1,5 - 20,7} = 8,1$$

$$T_{I/B\ PS11} = \frac{34,9 \cdot 4,5 - 1,1 \cdot 17,4}{34,9 + 4,5 - 1,1 - 17,4} = 6,6$$

20 Se determinó entonces el límite superior (UNC) de la región "Sin determinación", para cada conjunto óptimo de sondas, de acuerdo con la fórmula siguiente:

$$U_{NC} = \frac{T_{I/B} \cdot [p_{Mediana\ I/B} - 2 \cdot p(I/B)_\lambda] + p(I/B)_\lambda^2}{p_{Mediana\ I/B} - T_{I/B}}$$

$$U_{NC\ PS9} = \frac{8,1 \cdot (45,7 - 2 \cdot 20,7) + 20,7^2}{45,7 - 8,1} = 12,3$$

$$U_{NC\ PS11} = \frac{6,6 \cdot (34,9 - 2 \cdot 17,4) + 17,4^2}{34,9 - 6,6} = 10,8$$

25 Se determinó entonces el límite inferior (LNC) de la región "Sin determinación", para cada conjunto óptimo de sondas, de acuerdo con la fórmula siguiente

$$L_{NC} = \frac{T_{I/B} \cdot [2 \cdot a(I/B)_n - a_{Mediana\ I/B}] - a(I/B)_n^2}{T_{I/B} - a_{Mediana\ I/B}}$$

ES 2 641 600 T3

n.º 01	n.º 2	n.º 1	CCTGCCCTCTGGGCCCTAACACTG	1.
n.º 01	n.º 2	n.º 2	CCTGCCCTCTGCGACCTAACACTG	4.
n.º 01	n.º 2	n.º 2	CCTGCCCTCTGCGCACTAACACTG	5.
n.º 02	n.º 1	n.º 1	GGTCATCACACTGTTTCAGGTATT	6.
n.º 02	n.º 1	n.º 1	TGGTCATCACACTGTTTCAGGTATTG	7.
n.º 02	n.º 1	n.º 2	GGTCATCACACCGTTTCAGGTATT	8.
n.º 02	n.º 1	n.º 2	TGGTCATCACACCGTTTCAGGTATTG	9.
n.º 02	n.º 2	n.º 1	GTCATCACACTGTTTCAGGTATTG	10.
n.º 02	n.º 2	n.º 2	GTCATCACACTTCAGGTATTGGG	11.
n.º 03	n.º 1	n.º 1	CCCAGTATTCGGCTGGCCACCATGA	12.
n.º 03	n.º 1	n.º 1	CCAGTATTCGGCTGGCCACCATG	13.
n.º 03	n.º 1	n.º 2	CCCAGTATTCGGTGGCCACCATGAG	14.
n.º 03	n.º 1	n.º 2	CCAGTATTCGGTGGCCACCATGA	15.
n.º 03	n.º 2	n.º 1	AGGTGACAGCTTTAGGCAACCTGAG	16.
n.º 03	n.º 2	n.º 2	AGGTGACAGCTTAGGCAACCTGAGG	17.
n.º 04	n.º 1	n.º 1	ACATGAACATGATGCACATCTACGT	18.
n.º 04	n.º 1	n.º 1	CATGAACATGATGCACATCTACG	19.
n.º 04	n.º 1	n.º 2	ACATGAACATGACGCACATCTACGT	20.
n.º 04	n.º 1	n.º 2	CATGAACATGACGCACATCTACG	21.
n.º 04	n.º 2	n.º 1	AACATGATGCACATCTACGTGTTCCG	22.
n.º 04	n.º 2	n.º 1	ACATGATGCACATCTACGTGTTCC	23.
n.º 04	n.º 2	n.º 2	AACCTGAGGCACTTCTACGTGTTCCG	24.
n.º 04	n.º 2	n.º 2	ACCTGAGGCACTTCTACGTGTTCC	25.
n.º 05	n.º 1	n.º 1	AGGAAGAATGCCGTGTTCAACAC	26.
n.º 05	n.º 1	n.º 2	AGGAAGAATGCGTGTTCAACACC	27.
n.º 05	n.º 2	n.º 1	GGCTCACCCCAAGGGAAGGGAAGA	28.
n.º 05	n.º 2	n.º 2	GGCTCACCCCAAGGGAAGGGAAGAT	29.
n.º 06	n.º 1	n.º 1	ATGGTGTCTCTCTCTACCTTGCTT	30.
n.º 06	n.º 1	n.º 1	TGGTGTCTCTCTCTACCTTGCT	31.
n.º 06	n.º 1	n.º 2	ATGGTGTCTCTACCTTGCTTCCTT	32.
n.º 06	n.º 1	n.º 2	TGGTGTCTCTACCTTGCTTCCT	33.
n.º 06	n.º 2	n.º 1	TTTGCAGACTTATGTGCACAGTGCG	34.
n.º 06	n.º 2	n.º 1	TTGCAGACTTATGTGCACAGTGC	35.
n.º 06	n.º 2	n.º 2	TTTGCAGACTTAGGTGCACAGTGCG	36.
n.º 06	n.º 2	n.º 2	TTGCAGACTTAGGTGCACAGTGC	37.
n.º 06	n.º 3	n.º 1	TTTGCAGACTTATGTGCACAGTGCG	34
n.º 06	n.º 3	n.º 1	TTGCAGACTTATGTGCACAGTGC	35
n.º 06	n.º 3	n.º 2	TTTGCAGACTTAAGTGCACAGTGCG	38.
n.º 06	n.º 3	n.º 2	TTGCAGACTTAAGTGCACAGTGC	39.
n.º 06	n.º 4	n.º 1	GCAGACTTATGTGCACAGTGCGG	40.

ES 2 641 600 T3

n.º 06	n.º 4	n.º 1	CAGACTTATGTGCACAGTGCG	41
n.º 06	n.º 4	n.º 2	GCAGACTTATGGGCACAGTGCGG	42
n.º 06	n.º 4	n.º 2	CAGACTTATGGGCACAGTGCG	43
n.º 06	n.º 5	n.º 1	TTGGCAGGAGGCGTGGCTGTG	44
n.º 06	n.º 5	n.º 1	ACAGCCACGCCTCCTGCCA	45
n.º 06	n.º 5	n.º 2	TTGGCAGGAGACGTGGCTGTG	46
n.º 06	n.º 5	n.º 2	ACAGCCACGTCTCCTGCCA	47
n.º 06	n.º 6	n.º 1	GCGTGGCTGTGGGTACCTCGTGCA	48
n.º 06	n.º 6	n.º 1	CGTGGCTGTGGGTACCTCGTGTC	49
n.º 06	n.º 6	n.º 2	GCGTGGCTGTGGATACCTCGTGCA	50
n.º 06	n.º 6	n.º 2	CGTGGCTGTGGATACCTCGTGTC	51
n.º 06	n.º 7	n.º 1	GGCTGTGGGTACCTCGTGTCACC	52
n.º 06	n.º 7	n.º 1	GCTGTGGGTACCTCGTGTCAC	53
n.º 06	n.º 7	n.º 2	GGCTGTGGGTATCTCGTGTCACC	54
n.º 06	n.º 7	n.º 2	GCTGTGGGTATCTCGTGTCAC	55
n.º 06	n.º 8	n.º 1	TGGGTACCTCGTGTCACCTGATCCC	56
n.º 06	n.º 8	n.º 1	GGGTACCTCGTGTCACCTGATCC	57
n.º 06	n.º 8	n.º 2	TGGGTACCTCGTATCACCTGATCCC	58
n.º 06	n.º 8	n.º 2	GGGTACCTCGTATCACCTGATCC	59
n.º 06	n.º 9	n.º 1	GGCTTGCCATGGTGCTGGGTC	60
n.º 06	n.º 9	n.º 1	GCTTGCCATGGTGCTGGGT	61
n.º 06	n.º 9	n.º 2	GGCTTGCCATTGTGCTGGGTC	62
n.º 06	n.º 9	n.º 2	GCTTGCCATTGTGCTGGGT	63
n.º 06	n.º 10	n.º 1	GTGGCTGGGCTGATCTCCGTCGG	64
n.º 06	n.º 10	n.º 1	TGGCTGGGCTGATCTCCGTCG	65
n.º 06	n.º 10	n.º 2	GTGGCTGGGCTTGGCTGATCTCC	66
n.º 06	n.º 10	n.º 2	TGGCTGGGCTTGGCTGATCTC	67
n.º 06	n.º 11	n.º 1	GGCTGATCTCCGTCGGGGGAGCC	68
n.º 06	n.º 11	n.º 1	GCTGATCTCCGTCGGGGGAGC	69
n.º 06	n.º 11	n.º 2	GGCTGATCTCCATCGGGGGAGCC	70
n.º 06	n.º 11	n.º 2	GCTGATCTCCATCGGGGGAGC	71
n.º 06	n.º 12	n.º 1	GGAGCCAAGTACCTGCCGGTAAG	72
n.º 06	n.º 12	n.º 1	GAGCCAAGTACCTGCCGGTAA	73
n.º 06	n.º 12	n.º 2	GGAGCCAAGTAACTGCCGGTAAG	74
n.º 06	n.º 12	n.º 2	GAGCCAAGTAACTGCCGGTAA	75
n.º 06	n.º 13	n.º 1	AGTACCTGCCGGTAAGAACTAGAC	76
n.º 06	n.º 13	n.º 1	GTACCTGCCGGTAAGAACTAGA	77
n.º 06	n.º 13	n.º 2	AGTACCTGCCGGAACTAGACAACT	78
n.º 06	n.º 13	n.º 2	GTACCTGCCGGAACTAGACAAC	79
n.º 07	n.º 1	n.º 1	TGGGTCTGCTTGGAGAGATCATCTA	80

n.º 07	n.º 1	n.º 1	GGGTCTGCTTGGAGAGATCATCT	81
n.º 07	n.º 1	n.º 2	TGGGTCTGCTTGAAGAGATCATCTA	82
n.º 07	n.º 1	n.º 2	GGGTCTGCTTGAAGAGATCATCT	83
n.º 07	n.º 2	n.º 1	CTGCTGGTGCCTTGATACCGTCGGAG	84
n.º 07	n.º 2	n.º 1	TGCTGGTGCTTGATACCGTCGGGA	85
n.º 07	n.º 2	n.º 2	CTGCTGGTGCCTTCATACTGTCTGGA	86
n.º 07	n.º 2	n.º 2	TGCTGGTGCTTCATACTGTCTGG	87
n.º 08	n.º 1	n.º 1	GTCTCCTGACAGGTCAGTGTGAGGC	88
n.º 08	n.º 1	n.º 1	TCTCCTGACAGGTCAGTGTGAGG	89
n.º 08	n.º 1	n.º 2	GTCTCCTGACAGATCAGTGTGAGGC	90
n.º 08	n.º 1	n.º 2	TCTCCTGACAGATCAGTGTGAGG	91
n.º 09	n.º 1	n.º 1	GCATTTAAACAGGTTTGCTCCTAAA	92
n.º 09	n.º 1	n.º 1	TGCATTTAAACAGGTTTGCTCCTAAA	93
n.º 09	n.º 1	n.º 2	GCATTTAAACAGCTTTTGCTCCTAAA	94
n.º 09	n.º 1	n.º 2	TGCATTTAAACAGCTTTTGCTCCTAAA	95
n.º 09	n.º 2	n.º 1	TATTTTGATGACCAAGTTTTCTGGA	96
n.º 09	n.º 2	n.º 1	ATTTTGATGACCAAGTTTTCTGG	97
n.º 09	n.º 2	n.º 2	TATTTTGATGACTAAGTTTTCTGGA	98
n.º 09	n.º 2	n.º 2	ATTTTGATGACTAAGTTTTCTGG	99
n.º 10	n.º 1	n.º 1	GATGCTTTTGCTTAAAATCCAACAG	100
n.º 10	n.º 1	n.º 2	ATGCTTTTGCTTAAAATCCAACAG	101

Referencias

5 Botstein *et al.* (1980). *Am. J. Hum. Genet.* 32:314-331.

Chang *et al.* (1998). *Blood* 92:2602-2604.

Dutrillaux *et al.* (1990). *Cancer Genet. Cytocenet.* 49:203-217.

10 Hensel *et al.* (1990). *Cancer Res.* 50:3067.

Jin *et al.* (1993). *Blood* 82:2281-2288.

Kitano *et al.* (2000). *American Genetic Association* 91:211-214.

15 Mullis *et al.* (1986). *Cold Spring Harbor Symp. Quant. Biol.* 51:263-273.

Mullis K. B. y Faloona F. A. (1987). *Methods Enzymol.* 155:335-350.

20 Olerup O y Zetterquist H (1992). *Tissue Antigens* 39:225-235.

Sanger F y Coulson A. R. (1975). *J. Mol. Biol.* 94:441-448.

Smith *et al.* (1986). *Nature* 321:674-679.

25 Southern E. M. (1975). *J Mol Biol.* 98:503-517.

Patente de EE.UU. n.º 6599701. "Identifying organisms by detecting intronic nucleic acids".
 Solicitud de patente PCT n.º WO 2010/008071. "Method for screening of gene mutation by utilizing quantification
 30 technique".

Solicitud de patente PCT n.º WO 2001/018245. "Detection of alterations in a gene by long range PCR using human mobile elements".

LISTADO DE SECUENCIAS

5 <110> Progenika Biopharma S.A.
 <120> Sondas y métodos de determinación de la presencia o la ausencia de segmentos genéticos

10 <130> CSC/FP6697312
 <160> 129
 <170> PatentIn versión 3.3

15 <210> 1
 <211> 25
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

25 <400> 1
 cctgcccctc tgggcctaa cactg 25

30 <210> 2
 <211> 25
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

40 <400> 2
 cctgcccctc tgagcctaa cactg 25

45 <210> 3
 <211> 25
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

55 <400> 3
 cctgcccctc tgagcactaa cactg 25

60 <210> 4
 <211> 25
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 <400> 4
 cctgcccctc tgcgacctaa cactg 25

ES 2 641 600 T3

5
 <210> 5
 <211> 25
 <212> ADN
 <213> Secuencia artificial

10
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

15
 <400> 5
 cctgcccctc tgcgcactaa cactg 25

20
 <210> 6
 <211> 23
 <212> ADN
 <213> Secuencia artificial

25
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

30
 <400> 6
 ggatcatcaca ctgttcaggt att 23

35
 <210> 7
 <211> 25
 <212> ADN
 <213> Secuencia artificial

40
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

45
 <400> 7
 tggatcatcac actgttcagg tattg 25

50
 <210> 8
 <211> 23
 <212> ADN
 <213> Secuencia artificial

55
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

60
 <400> 8
 ggatcatcaca ccgttcaggt att 23

65
 <210> 9
 <211> 25
 <212> ADN
 <213> Secuencia artificial

70
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

75
 <400> 9
 tggatcatcac accgttcagg tattg 25

80
 <210> 10

<211> 23
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

10

<400> 10
 gtcacacac tggtcagga ttg 23

<210> 11
 <211> 23
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

20

<400> 11
 gtcacacac ttcaggtatt ggg 23

25

<210> 12
 <211> 25
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

35

<400> 12
 cccagttattc ggctggccac catga 25

40

<210> 13
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

45

<400> 13
 ccagttattcg gctggccacc atg 23

<210> 14
 <211> 25
 <212> ADN
 <213> Secuencia artificial

50

<220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

55

<400> 14
 cccagttattc ggtggccacc atgag 25

60

<210> 15
 <211> 23
 <212> ADN

65

ES 2 641 600 T3

<213> Secuencia artificial
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 5
 <400> 15
 ccagtattcg gtggccacca tga 23
 10
 <210> 16
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 15
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 20
 <400> 16
 aggtgacagc ttaggcaac ctgag 25
 25
 <210> 17
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 35
 <400> 17
 aggtgacagc ttaggcaacc tgagg 25
 40
 <210> 18
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 <400> 18
 acatgaacat gatgcacatc tacgt 25
 50
 <210> 19
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 60
 <400> 19
 catgaacatg atgcacatct acg 23
 65
 <210> 20
 <211> 25
 <212> ADN
 <213> Secuencia artificial

ES 2 641 600 T3

<220>
 <221> fuente
 <222> 1..25
 5 <223> /nota="Descripción de la secuencia artificial: sonda"

 <400> 20
 acatgaacat gacgcacatc tacgt 25

 10 <210> 21
 <211> 23
 <212> ADN
 <213> Secuencia artificial

 15 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

 20 <400> 21
 catgaacatg acgcacatct acg 23

 25 <210> 22
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 30 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

 35 <400> 22
 aacatgatgc acatctacgt gttcg 25

 40 <210> 23
 <211> 23
 <212> ADN
 <213> Secuencia artificial

 45 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

 50 <400> 23
 acatgatgca catctacgtg ttc 23

 55 <210> 24
 <211> 25
 <212> ADN
 <213> Secuencia artificial

 60 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

 65 <400> 24
 aacctgaggc acttctacgt gttcg 25

 <210> 25
 <211> 23
 <212> ADN
 <213> Secuencia artificial

ES 2 641 600 T3

<220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 5
 <400> 25
 acctgaggca cttctacgtg ttc 23
 <210> 26
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 10
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 15
 <400> 26
 aggaagaatg ccgtgttcaa cac 23
 <210> 27
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 20
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 25
 <400> 27
 aggaagaatg cgtgttcaac acc 23
 <210> 28
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 35
 <400> 28
 ggctcacccc caaggaagg gaaga 25
 <210> 29
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 40
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 45
 <400> 29
 ggctcacccc caggaaggg aagat 25
 <210> 30
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 50
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 55
 <400> 29
 ggctcacccc caggaaggg aagat 25
 <210> 30
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 60
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 65
 <400> 29
 ggctcacccc caggaaggg aagat 25
 <210> 30
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 70
 <220>

ES 2 641 600 T3

<222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

5 <400> 35
 ttgcagactt atgtgcacag tgc 23

10 <210> 36
 <211> 25
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

20 <400> 36
 ttgcagact taggtgcaca gtgcg 25

25 <210> 37
 <211> 23
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

35 <400> 37
 ttgcagactt aggtgcacag tgc 23

40 <210> 38
 <211> 25
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

50 <400> 38
 ttgcagact taagtcaca gtgcg 25

55 <210> 39
 <211> 23
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

65 <400> 39
 ttgcagactt aagtcacag tgc 23

<210> 40
 <211> 23
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <222> 1..23

ES 2 641 600 T3

<223> /nota="Descripción de la secuencia artificial: sonda"

5 <400> 40
gcagacttat gtgcacagtg cgg 23

<210> 41
<211> 21
<212> ADN
<213> Secuencia artificial

10 <220>
<221> fuente
<222> 1..21
<223> /nota="Descripción de la secuencia artificial: sonda"

15 <400> 41
cagacttatg tgcacagtgc g 21

20 <210> 42
<211> 23
<212> ADN
<213> Secuencia artificial

25 <220>
<221> fuente
<222> 1..23
<223> /nota="Descripción de la secuencia artificial: sonda"

30 <400> 42
gcagacttat gggcacagtg cgg 23

35 <210> 43
<211> 21
<212> ADN
<213> Secuencia artificial

40 <220>
<221> fuente
<222> 1..21
<223> /nota="Descripción de la secuencia artificial: sonda"

45 <400> 43
cagacttatg ggcacagtgc g 21

50 <210> 44
<211> 21
<212> ADN
<213> Secuencia artificial

55 <220>
<221> fuente
<222> 1..21
<223> /nota="Descripción de la secuencia artificial: sonda"

60 <400> 44
ttggcaggag gcgtggctgt g 21

65 <210> 45
<211> 19
<212> ADN
<213> Secuencia artificial

<220>
<221> fuente
<222> 1..19
<223> /nota="Descripción de la secuencia artificial: sonda"

ES 2 641 600 T3

<400> 45
 acagccacgc ctctgcca 19

5 <210> 46
 <211> 21
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

15 <400> 46
 ttggcaggag acgtggctgt g 21

20 <210> 47
 <211> 19
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <221> fuente
 <222> 1..19
 <223> /nota="Descripción de la secuencia artificial: sonda"

30 <400> 47
 acagccacgt ctctgcca 19

35 <210> 48
 <211> 25
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

45 <400> 48
 gcgtggctgt ggtacctcg tgtca 25

50 <210> 49
 <211> 23
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

60 <400> 49
 cgtggctgtg ggtacctcg gtc 23

65 <210> 50
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

ES 2 641 600 T3

	<code><400> 50</code>		
	<code>gcgtggctgt ggatacctcg tgtca</code>		25
5	<code><210> 51</code> <code><211> 23</code> <code><212> ADN</code> <code><213> Secuencia artificial</code>		
10	<code><220></code> <code><221> fuente</code> <code><222> 1..23</code> <code><223> /nota="Descripción de la secuencia artificial: sonda"</code>		
15	<code><400> 51</code>		
	<code>cgtaggctgtg gatacctcgt gtc</code>		23
20	<code><210> 52</code> <code><211> 23</code> <code><212> ADN</code> <code><213> Secuencia artificial</code>		
25	<code><220></code> <code><221> fuente</code> <code><222> 1..23</code> <code><223> /nota="Descripción de la secuencia artificial: sonda"</code>		
30	<code><400> 52</code>		
	<code>ggctgtgggt acctcgtgtc acc</code>		23
35	<code><210> 53</code> <code><211> 21</code> <code><212> ADN</code> <code><213> Secuencia artificial</code>		
40	<code><220></code> <code><221> fuente</code> <code><222> 1..21</code> <code><223> /nota="Descripción de la secuencia artificial: sonda"</code>		
45	<code><400> 53</code>		
	<code>gctgtgggta cctcgtgtca c</code>		21
50	<code><210> 54</code> <code><211> 23</code> <code><212> ADN</code> <code><213> Secuencia artificial</code>		
55	<code><220></code> <code><221> fuente</code> <code><222> 1..23</code> <code><223> /nota="Descripción de la secuencia artificial: sonda"</code>		
60	<code><400> 54</code>		
	<code>ggctgtgggt atctcgtgtc acc</code>		23
65	<code><210> 55</code> <code><211> 21</code> <code><212> ADN</code> <code><213> Secuencia artificial</code>		
70	<code><220></code> <code><221> fuente</code> <code><222> 1..21</code> <code><223> /nota="Descripción de la secuencia artificial: sonda"</code>		
75	<code><400> 55</code>		

ES 2 641 600 T3

	gctgtgggta tctcgtgtca c	21
5	<210> 56 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <222> 1..25 <223> /nota="Descripción de la secuencia artificial: sonda"	
15	<400> 56 tgggtacctc gtgtcacctg atccc	25
20	<210> 57 <211> 23 <212> ADN <213> Secuencia artificial	
25	<220> <221> fuente <222> 1..23 <223> /nota="Descripción de la secuencia artificial: sonda"	
30	<400> 57 gggtagctcg tgtcacctga tcc	23
35	<210> 58 <211> 25 <212> ADN <213> Secuencia artificial	
40	<220> <221> fuente <222> 1..25 <223> /nota="Descripción de la secuencia artificial: sonda"	
45	<400> 58 tgggtacctc gtatcacctg atccc	25
50	<210> 59 <211> 23 <212> ADN <213> Secuencia artificial	
55	<220> <221> fuente <222> 1..23 <223> /nota="Descripción de la secuencia artificial: sonda"	
60	<400> 59 gggtagctcg tatcacctga tcc	23
65	<210> 60 <211> 21 <212> ADN <213> Secuencia artificial	
70	<220> <221> fuente <222> 1..21 <223> /nota="Descripción de la secuencia artificial: sonda"	
75	<400> 60 ggcttgccat ggtgctgggt c	21

ES 2 641 600 T3

5 <210> 61
 <211> 19
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <221> fuente
 <222> 1..19
 <223> /nota="Descripción de la secuencia artificial: sonda"

15 <400> 61
 gcttgccatg gtgctgggt 19

20 <210> 62
 <211> 21
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

30 <400> 62
 ggcttgccat tggctgggt c 21

35 <210> 63
 <211> 19
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <221> fuente
 <222> 1..19
 <223> /nota="Descripción de la secuencia artificial: sonda"

45 <400> 63
 gcttgccatt gtgctgggt 19

50 <210> 64
 <211> 23
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

60 <400> 64
 gtggctgggc tgatctccgt cgg 23

65 <210> 65
 <211> 21
 <212> ADN
 <213> Secuencia artificial

70 <220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

75 <400> 65
 tggctgggct gatctccgtc g 21

ES 2 641 600 T3

5
 <210> 66
 <211> 23
 <212> ADN
 <213> Secuencia artificial

10
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

15
 <400> 66
 gtggctgggc ttgctgatc tcc 23

20
 <210> 67
 <211> 21
 <212> ADN
 <213> Secuencia artificial

25
 <220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

30
 <400> 67
 tggctgggct tggctgatc t c 21

35
 <210> 68
 <211> 23
 <212> ADN
 <213> Secuencia artificial

40
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

45
 <400> 68
 ggctgatctc cgctggggga gcc 23

50
 <210> 69
 <211> 21
 <212> ADN
 <213> Secuencia artificial

55
 <220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

60
 <400> 69
 gctgatctcc gtcggggag c 21

65
 <210> 70
 <211> 23
 <212> ADN
 <213> Secuencia artificial

70
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

75
 <400> 70
 ggctgatctc catcggggga gcc 23

80
 <210> 71

ES 2 641 600 T3

<211> 21
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

10

<400> 71
 gctgatctcc atcgggggag c 21

<210> 72
 <211> 23
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

20

<400> 72
 ggagccaagt acctgccggt aag 23

25

<210> 73
 <211> 21
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

35

<400> 73
 gagccaagta cctgccggt a 21

40

<210> 74
 <211> 23
 <212> ADN
 <213> Secuencia artificial

45

<220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

50

<400> 74
 ggagccaagt aactgccggt aag 23

<210> 75
 <211> 21
 <212> ADN
 <213> Secuencia artificial

55

<220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

60

<400> 75
 gagccaagta actgccggt a 21

65

<210> 76
 <211> 25

ES 2 641 600 T3

<212> ADN
 <213> Secuencia artificial

5 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

10 <400> 76
 agtacctgcc ggtaagaaac tagac 25

15 <210> 77
 <211> 23
 <212> ADN
 <213> Secuencia artificial

20 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

25 <400> 77
 gtacctgccg gtaagaaact aga 23

30 <210> 78
 <211> 25
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

40 <400> 78
 agtacctgcc gaaactaga caact 25

45 <210> 79
 <211> 23
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

55 <400> 79
 gtacctgccg gaaactagac aac 23

60 <210> 80
 <211> 25
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

<400> 80
 tgggtctgct tggagagatc atcta 25

<210> 81
 <211> 23
 <212> ADN

<213> Secuencia artificial
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 5
 <400> 81
 gggctgctt ggagatca tct 23
 10
 <210> 82
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 15
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 20
 <400> 82
 tggctgct tgaagatc atcta 25
 25
 <210> 83
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 30
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 35
 <400> 83
 gggctgctt gaagatca tct 23
 40
 <210> 84
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 45
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 <400> 84
 ctgctggtgc tgataccgt cggag 25
 50
 <210> 85
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 55
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 60
 <400> 85
 tgctggtgct tgataccgtc gga 23
 65
 <210> 86
 <211> 25
 <212> ADN
 <213> Secuencia artificial

ES 2 641 600 T3

<220>
<221> fuente
<222> 1..25
5 <223> /nota="Descripción de la secuencia artificial: sonda"

<400> 86
ctgctggtgc ttcatactgt ctgga 25

10 <210> 87
<211> 23
<212> ADN
<213> Secuencia artificial

15 <220>
<221> fuente
<222> 1..23
<223> /nota="Descripción de la secuencia artificial: sonda"

20 <400> 87
tgctggtgct tcatactgtc tgg 23

25 <210> 88
<211> 25
<212> ADN
<213> Secuencia artificial

30 <220>
<221> fuente
<222> 1..25
<223> /nota="Descripción de la secuencia artificial: sonda"

35 <400> 88
gtctcctgac aggtcagtgt gaggc 25

40 <210> 89
<211> 23
<212> ADN
<213> Secuencia artificial

45 <220>
<221> fuente
<222> 1..23
<223> /nota="Descripción de la secuencia artificial: sonda"

50 <400> 89
tctcctgaca ggtcagtgtg agg 23

55 <210> 90
<211> 25
<212> ADN
<213> Secuencia artificial

60 <220>
<221> fuente
<222> 1..25
<223> /nota="Descripción de la secuencia artificial: sonda"

65 <400> 90
gtctcctgac agatcagtgt gaggc 25

<210> 91
<211> 23
<212> ADN
<213> Secuencia artificial

ES 2 641 600 T3

<222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

5 <400> 101
 atgctttgc ttaaaaatcc aacag 25

10 <210> 102
 <211> 25
 <212> ADN
 <213> Secuencia artificial

15 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

20 <400> 102
 agaagtcaa tcgaaaggaa gaatg 25

25 <210> 103
 <211> 25
 <212> ADN
 <213> Secuencia artificial

30 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

35 <400> 103
 agaagtcaa tctaaaggaa gaatg 25

40 <210> 104
 <211> 23
 <212> ADN
 <213> Secuencia artificial

45 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

50 <400> 104
 gaagtccaat cgaaggaag aat 23

55 <210> 105
 <211> 23
 <212> ADN
 <213> Secuencia artificial

60 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"

65 <400> 105
 gaagtccaat ctaaaggaag aat 23

70 <210> 106
 <211> 25
 <212> ADN
 <213> Secuencia artificial

75 <220>
 <221> fuente
 <222> 1..25

<223> /nota="Descripción de la secuencia artificial: sonda"
 <400> 106
 caaggactat caggccacgg ggtca 25
 5
 <210> 107
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 10
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 15
 <400> 107
 caaggactat cagcccacgg ggtca 25
 <210> 108
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 20
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 25
 <400> 108
 tgaccccggtg gcctgatagt ccttg 25
 30
 <210> 109
 <211> 25
 <212> ADN
 <213> Secuencia artificial
 35
 <220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"
 40
 <400> 109
 tgaccccggtg ggctgatagt ccttg 25
 45
 <210> 110
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 50
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 55
 <400> 110
 ctggccccca ggcgcctct tct 23
 <210> 111
 <211> 23
 <212> ADN
 <213> Secuencia artificial
 60
 <220>
 <221> fuente
 <222> 1..23
 <223> /nota="Descripción de la secuencia artificial: sonda"
 65

<400> 111
 ctggcccca gtcgccctct tct 23

5 <210> 112
 <211> 21
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

15 <400> 112
 ctggccaagt ttcaactctg c 21

20 <210> 113
 <211> 21
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <221> fuente
 <222> 1..21
 <223> /nota="Descripción de la secuencia artificial: sonda"

30 <400> 113
 ctggccaagt gtcaactctg c 21

35 <210> 114
 <211> 19
 <212> ADN
 <213> Secuencia artificial

40 <220>
 <221> fuente
 <222> 1..19
 <223> /nota="Descripción de la secuencia artificial: sonda"

45 <400> 114
 tggccaagtt tcaactctg 19

50 <210> 115
 <211> 19
 <212> ADN
 <213> Secuencia artificial

55 <220>
 <221> fuente
 <222> 1..19
 <223> /nota="Descripción de la secuencia artificial: sonda"

60 <400> 115
 tggccaagtg tcaactctg 19

65 <210> 116
 <211> 25
 <212> ADN
 <213> Secuencia artificial

<220>
 <221> fuente
 <222> 1..25
 <223> /nota="Descripción de la secuencia artificial: sonda"

ES 2 641 600 T3

	<400> 116 agtttcaact ctgctctgct gagaa	25
5	<210> 117 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <222> 1..25 <223> /nota="Descripción de la secuencia artificial: sonda"	
15	<400> 117 agtttcaact ctctctgct gagaa	25
20	<210> 118 <211> 25 <212> ADN <213> Secuencia artificial	
25	<220> <221> fuente <222> 1..25 <223> /nota="Descripción de la secuencia artificial: sonda"	
30	<400> 118 ctgctctgct gagaagtcca atcga	25
35	<210> 119 <211> 25 <212> ADN <213> Secuencia artificial	
40	<220> <221> fuente <222> 1..25 <223> /nota="Descripción de la secuencia artificial: sonda"	
45	<400> 119 ctgctctgct gaaaagtcca atcga	25
50	<210> 120 <211> 23 <212> ADN <213> Secuencia artificial	
55	<220> <221> fuente <222> 1..23 <223> /nota="Descripción de la secuencia artificial: sonda"	
60	<400> 120 tgctctgctg agaagtcca tcg	23
65	<210> 121 <211> 23 <212> ADN <213> Secuencia artificial	
	<220> <221> fuente <222> 1..23 <223> /nota="Descripción de la secuencia artificial: sonda"	
	<400> 121	

ES 2 641 600 T3

	tgctctgctg aaaagtccaa tcg	23
5	<210> 122 <211> 25 <212> ADN <213> Secuencia artificial	
10	<220> <221> fuente <222> 1..25 <223> /nota="Descripción de la secuencia artificial: sonda"	
15	agaagtccaa tccaaaggaa gaatg	25
20	<210> 123 <211> 23 <212> ADN <213> Secuencia artificial	
25	<220> <221> fuente <222> 1..23 <223> /nota="Descripción de la secuencia artificial: sonda"	
30	gaagtccaat ccaaggaag aat	23
35	<210> 124 <211> 23 <212> ADN <213> Secuencia artificial	
40	<220> <221> fuente <222> 1..23 <223> /nota="Descripción de la secuencia artificial: sonda"	
45	ggaagaatgc cgtgttcaac acc	23
50	<210> 125 <211> 23 <212> ADN <213> Secuencia artificial	
55	<220> <221> fuente <222> 1..23 <223> /nota="Descripción de la secuencia artificial: sonda"	
60	ggaagaatgc catgttcaac acc	23
65	<210> 126 <211> 21 <212> ADN <213> Secuencia artificial	
70	<220> <221> fuente <222> 1..21 <223> /nota="Descripción de la secuencia artificial: sonda"	
75	gaagaatgcc gtgttcaaca c	21

ES 2 641 600 T3

5 <210> 127
<211> 21
<212> ADN
<213> Secuencia artificial

10 <220>
<221> fuente
<222> 1..21
<223> /nota="Descripción de la secuencia artificial: sonda"

15 <400> 127
gaagaatgcc atgtcaaca c 21

20 <210> 128
<211> 25
<212> ADN
<213> Secuencia artificial

25 <220>
<221> fuente
<222> 1..25
<223> /nota="Descripción de la secuencia artificial: sonda"

30 <400> 128
agaagtccaa tcaaaggaa gaatg 25

35 <210> 129
<211> 23
<212> ADN
<213> Secuencia artificial

40 <220>
<221> fuente
<222> 1..23
<223> /nota="Descripción de la secuencia artificial: sonda"

<400> 129
gaagtcaat caaaaggaag aat 23

REIVINDICACIONES

1. Un método de determinación de la presencia o la ausencia de un segmento genético de interés, tal como un exón, un intrón o un promotor, en una muestra que contiene ADN, comprendiendo el método:

(i) poner en contacto un primer conjunto de sondas, que comprende una pluralidad de réplicas de al menos una sonda oligonucleotídica que interroga a una primera secuencia afín dentro de dicho segmento de interés, con:

- (a) una pluralidad de muestras de referencia en las que el segmento genético de interés esté ausente, y
- (b) una pluralidad de muestras de referencia en las que el segmento genético de interés esté presente,

en condiciones que permitan que se produzca la hibridación de la sonda y la secuencia afín;

(ii) medir la intensidad de hibridación del conjunto de sondas de cada una de las muestras de referencia, obteniéndose de este modo, de la etapa (i)(a), un primer grupo de valores de intensidad de hibridación para las muestras de referencia en las que está ausente el segmento genético de interés y, de la etapa (i)(b), un segundo grupo de valores de intensidad de hibridación para las muestras de referencia en las que está presente el segmento genético de interés; y

(iii) establecer una región "sin determinación" de los valores de intensidad de hibridación que se encuentra en la región entre dichos primer y segundo grupos y que está limitada por un límite sin determinación inferior, L_{NC} , y un límite sin determinación superior, U_{NC} , en donde el L_{NC} y el U_{NC} representan límites de confianza estadística para asignar un valor de intensidad de la hibridación a dichos primer y segundo grupos, respectivamente, y se calculan de la siguiente manera:

$$L_{NC} = \frac{T_{I/B} \cdot [2 \cdot a(I/B)_n - aMediana I/B] - a(I/B)_n^2}{T_{I/B} - aMediana I/B}$$

$$U_{NC} = \frac{T_{I/B} \cdot [pMediana I/B - 2 \cdot p(I/B)_1] + p(I/B)_1^2}{pMediana I/B - T_{I/B}}$$

donde:

$$T_{I/B} = \frac{pMediana IB \cdot a(I/B)_n - aMediana IB \cdot p(I/B)_1}{pMediana IB + a(I/B)_n - aMediana IB - p(I/B)_1}$$

I/B es la intensidad de hibridación del conjunto de sondas;

aMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que está ausente el segmento genético de interés;

pMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que está presente el segmento genético de interés;

$a(I/B)_n$ es el mayor valor de I/B medido de las muestras de referencia en las que está ausente el segmento genético de interés; y

$p(I/B)_1$ es el menor valor de I/B medido de las muestras de referencia en las que está presente el segmento genético de interés;

iv) poner en contacto el primer conjunto de sondas con al menos una muestra de ensayo que contenga ADN en condiciones que permitan que se produzca la hibridación de la sonda y la secuencia afín;

v) medir la intensidad de la hibridación del conjunto de sondas para la al menos una muestra de ensayo; y

vi) comparar la intensidad de hibridación medida con el L_{NC} y el U_{NC} ;

en el que una intensidad de hibridación medida por debajo del L_{NC} indica que dicho segmento genético de interés está ausente en la muestra de ensayo y una intensidad de hibridación medida por encima del U_{NC} indica que dicho segmento genético de interés está presente en la muestra de ensayo;

y en el que el ADN de las muestras de referencia y/o la muestra de ensayo se ha marcado con un marcador fluorescente;

y en el que la medición de la intensidad de hibridación comprende medir la señal de fluorescencia del marcador fluorescente en cada ubicación de la sonda oligonucleotídica;

y en el que la intensidad de hibridación para cada conjunto de sondas se mide como I/B, donde:

I se determina como la media o la mediana de la señal de fluorescencia medida combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente tras el recorte de las lecturas extremas; y

B se determina como la media o la mediana de la señal de fluorescencia de fondo combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente tras el recorte de las lecturas extremas.

2. Un método de acuerdo con la reivindicación 1, en el que dicha secuencia afín comprende un sitio polimórfico y en el que dicho primer conjunto de sondas comprende una pluralidad de réplicas de al menos dos, al menos tres o al menos cuatro sondas oligonucleotídicas específicas de alelos que distinguen los alelos de dicho sitio polimórfico.

3. Un método de acuerdo con la reivindicación 2, en el que dicha muestra de ensayo comprende dicho segmento genético de interés, comprendiendo además el método la genotipificación de dicha muestra de ensayo para identificar un alelo en dicho sitio polimórfico.

4. Un método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que se emplea al menos un segundo conjunto de sondas además de dicho primer conjunto de sondas y en el que el segundo conjunto de sondas comprende una pluralidad de réplicas de al menos una sonda oligonucleotídica que interroga a una segunda secuencia afín dentro de dicho segmento de interés, comprendiendo el método además establecer una segunda región sin determinación, limitada por los límites L_{NC} y U_{NC} para el segundo conjunto de sondas.

5. Un método de acuerdo con la reivindicación 4, en el que la muestra de ensayo se clasifica como aquella que tiene o que no tiene dicho segmento de interés basado en un consenso estricto entre las determinaciones del primer, PS_1 , y el segundo, PS_2 , conjuntos de sondas, de acuerdo con un consenso estricto como el mostrado en la Tabla 1:

(Tabla 1)

Valor de I/B de PS_1	Valor de I/B de PS_2	Determinación de segmento genético
$I/B \text{ de } PS_1 > U_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Presente
$I/B \text{ de } PS_1 < L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Ausente
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Sin determinación
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Sin determinación
$I/B \text{ de } PS_1 > U_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación
$I/B \text{ de } PS_1 < L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación

6. Un método de acuerdo con la reivindicación 4, en el que la muestra de ensayo se clasifica como aquella que tiene o que no tiene dicho segmento de interés basado en un consenso mayoritario entre las determinaciones del primer, PS_1 , y el segundo, PS_2 , conjuntos de sondas en ausencia de determinaciones contradictorias, de acuerdo con un consenso mayoritario como el mostrado en la Tabla 2:

(Tabla 2)

Valor de I/B de PS_1	Valor de I/B de PS_2	Determinación de segmento genético
$I/B \text{ de } PS_1 > U_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Presente
$I/B \text{ de } PS_1 < L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Ausente
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Sin determinación
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 > U_{NC}$	Presente
$U_{NC} \geq I/B \text{ de } PS_1 \geq L_{NC}$	$I/B \text{ de } PS_2 < L_{NC}$	Ausente
$I/B \text{ de } PS_1 > U_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Presente
$I/B \text{ de } PS_1 < L_{NC}$	$U_{NC} \geq I/B \text{ de } PS_2 \geq L_{NC}$	Ausente

7. Un método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que las sondas oligonucleotídicas están unidas a un soporte sólido, opcionalmente en el que:

- (a) las sondas oligonucleotídicas están unidas a un soporte sólido esencialmente plano en forma de una matriz; o
- (b) las sondas oligonucleotídicas están unidas a partículas, tales como perlas, cilindros o cualquier otra partícula de tamaño micrométrico o nanométrico.

8. Un método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que:

- (a) la muestra de ensayo comprende ADN amplificado a partir de ADN genómico de un sujeto de ensayo, ADN que opcionalmente se ha fragmentado y/o marcado con un marcador detectable; y/o
 (b) cada una de las muestras de referencia comprende ADN amplificado a partir de ADN genómico de un sujeto de referencia, ADN que opcionalmente se ha fragmentado y/o marcado con un marcador detectable.

9. Un método de acuerdo con una cualquiera de las reivindicaciones anteriores, en el que el segmento genético de interés comprende un exón del gen RHD humano.

10. Un método de acuerdo la reivindicación 9, en el que las sondas oligonucleotídicas se seleccionan de las sondas mostradas en la Tabla 4.

11. Un método de acuerdo con la reivindicación 10, en donde el método comprende además la genotipificación de la muestra de ensayo para identificar al menos un alelo en un sitio de polimorfismo de un solo nucleótido, SNP, en el gen RHD humano.

12. Un método según lo definido en una cualquiera de las reivindicaciones 1 a 11, en el que dicho primer conjunto de sondas se selecciona:

(i) proporcionando una pluralidad de conjuntos de sondas candidato, comprendiendo cada conjunto de sondas candidato una pluralidad de réplicas de al menos una sonda oligonucleotídica que interroga a una secuencia afín dentro de un segmento genético de interés;

(ii) poniendo en contacto cada uno de los conjuntos de sondas candidato con:

(a) una pluralidad de muestras de referencia en las que esté ausente el segmento genético de interés; y

(b) una pluralidad de muestras de referencia en las que esté presente el segmento genético de interés;

en condiciones que permitan que se produzca la hibridación de la sonda y la secuencia afín;

(ii) midiendo la intensidad de hibridación del conjunto de sondas de cada una de las muestras de referencia, obteniéndose de este modo, de la etapa (ii)(a), para cada conjunto de sondas candidato un primer grupo de valores de intensidad de hibridación para las muestras de referencia en las que está ausente el segmento genético de interés y, de la etapa (ii)(b), un segundo grupo de valores de intensidad de hibridación para las muestras de referencia en las que está presente el segmento genético de interés; y

(iv) puntuando cada uno de los conjuntos de sondas candidato dividiendo la amplitud del intervalo entre el mayor valor de intensidad de hibridación medido del primer grupo y el menor valor de intensidad de hibridación medido del segundo grupo entre la suma de la mediana o de la media de los valores de intensidad de hibridación del primer y del segundo grupos, obteniéndose así una puntuación del conjunto de sondas para cada uno de los conjuntos de sondas candidato;

en el que una puntuación del conjunto de sondas superior indica que es más probable que el conjunto de sondas candidato sea óptimo para determinar la presencia o la ausencia del segmento genético de interés, y en el que las muestras de referencia se marcan con un marcador fluorescente y en el que la intensidad de hibridación se mide como I/B, donde:

I se determina como la media o la mediana de la señal de fluorescencia medida combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente tras el recorte de las lecturas extremas; y

B se determina como la media o la mediana de la señal de fluorescencia de fondo combinada de las réplicas de cada sonda oligonucleotídica, opcionalmente tras el recorte de las lecturas extremas.

13. Un método de acuerdo con la reivindicación 12, en el que la puntuación de conjunto de sondas (S) para cada conjunto de sondas candidato se calcula de la siguiente manera:

$$S = \frac{A}{M} = \frac{p(I/B)_1 - a(I/B)_n}{pMediana \ IB + aMediana \ IB}$$

donde:

aMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que está ausente el segmento genético de interés;

pMediana I/B es la mediana o la media de los valores de I/B medidos de las muestras de referencia en las que está presente el segmento genético de interés;

a(I/B)_n es el mayor valor de I/B medido de las muestras de referencia en las que está ausente el segmento genético de interés; y

p(I/B)₁ es el menor valor de I/B medido de las muestras de referencia en las que está presente el segmento genético de interés.

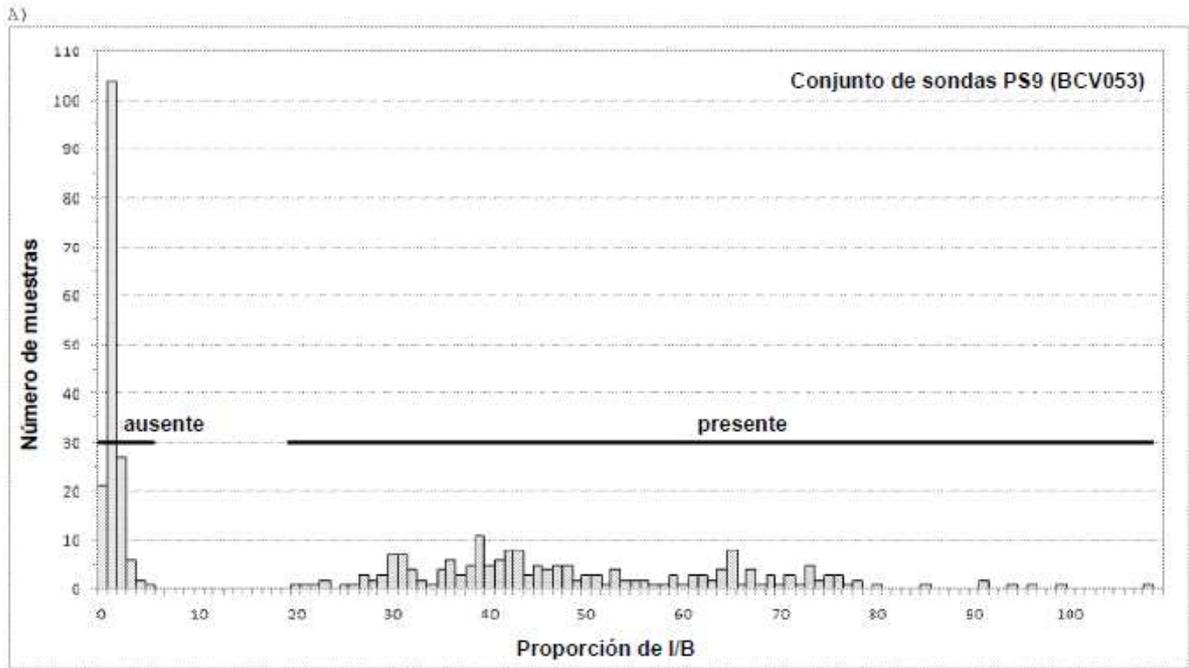


Figura 1

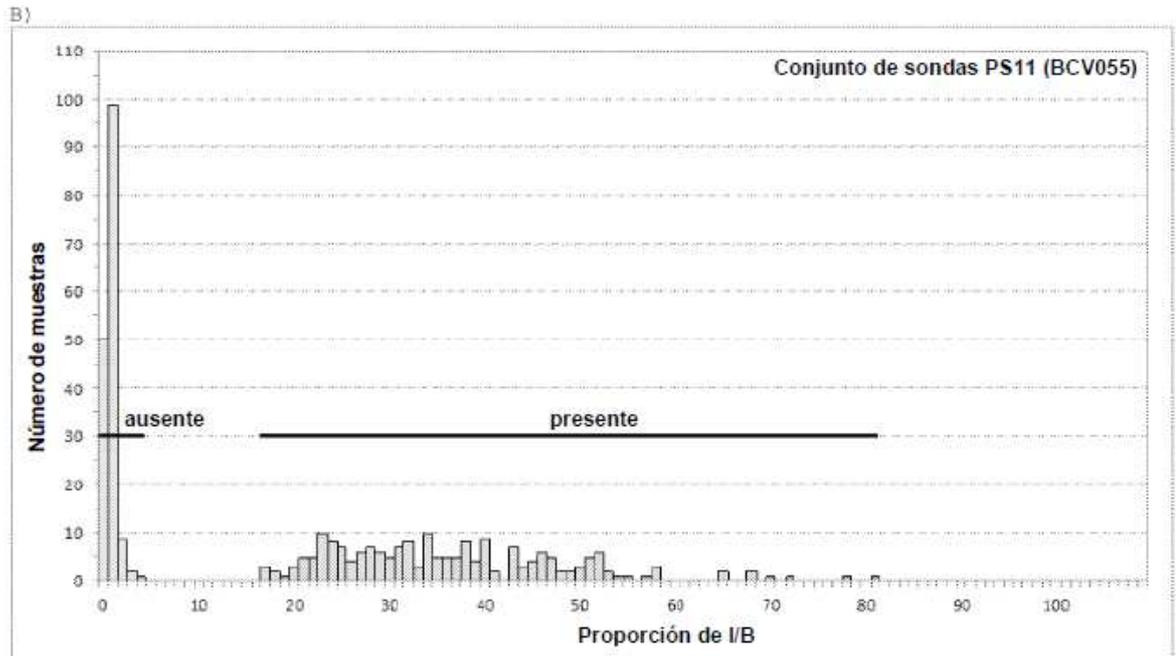


Figura 1

A)

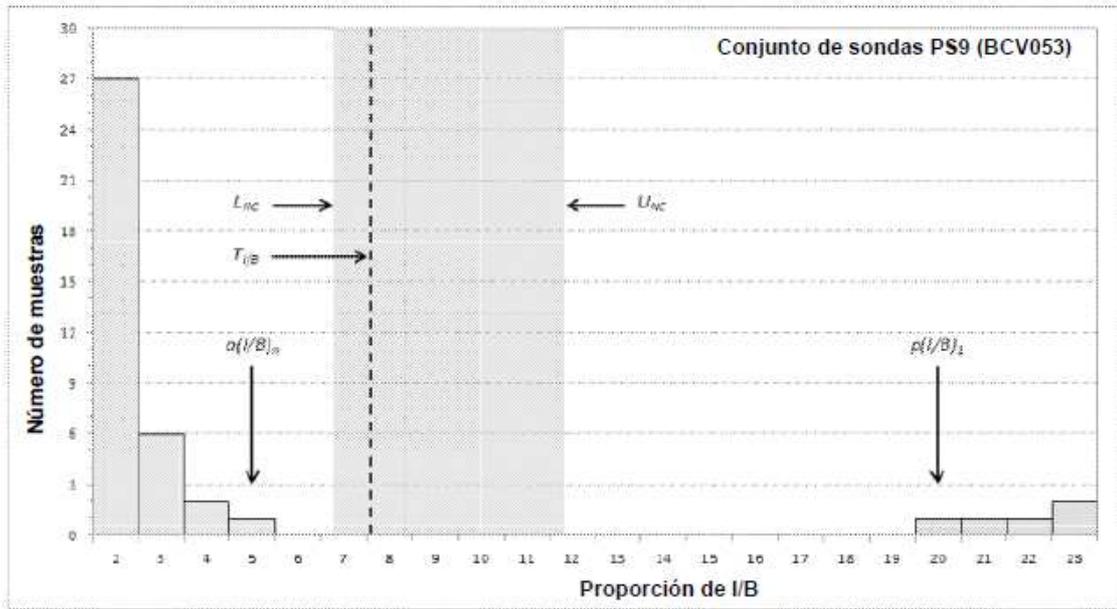


Figura 2

