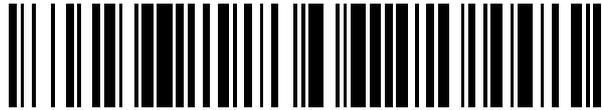


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 472 965**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **04.03.2011 E 11001840 (5)**

97 Fecha y número de publicación de la concesión europea: **16.04.2014 EP 2495336**

54 Título: **Nuevo tipo de sondas universales para la detección de variantes genómicas**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
03.07.2014

73 Titular/es:

F. HOFFMANN-LA ROCHE AG (100.0%)
Grenzacherstrasse 124
4070 Basel, CH

72 Inventor/es:

BECK, REINHARD;
BERGMANN, FRANK;
HÄRTEIS, RITA;
HEINDL, DIETER;
MAURITZ, RALF y
WALCH, HEIKO

74 Agente/Representante:

ISERN JARA, Jorge

ES 2 472 965 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

Nuevo tipo de sondas universales para la detección de variantes genómicas

5 La presente invención se refiere a una composición que comprende un primer juego de sondas y un segundo juego de sondas, compuesto de uno o más nucleótidos de ADN y cinco o más nucleótidos de ANB (ácido nucleico bloqueado) (o análogos de ANB como, por ejemplo, ANE (ácido nucleico con puente de 2'-O,4'-C-etileno) o derivados 2'-amino-ANB), en la que la sonda difiere en una, dos o tres posiciones aleatorias del ANB y la base en una posición discriminante. Además, la presente invención se refiere a un método para detectar variantes genómicas
10 mediante las sondas anteriormente indicadas. Además, la invención se refiere a una biblioteca de una pluralidad de sondas.

Es ampliamente conocido que los polimorfismos genéticos desempeñan un papel crucial para la salud y la predisposición a diversas enfermedades en prácticamente todos los organismos. Los polimorfismos genéticos y las mutaciones pueden presentar un gran impacto sobre el metabolismo del organismo afectado en el caso de que
15 resulte afectado un gen, o pueden permanecer inadvertidas (mutación silenciosa) en el caso de que resulte afectada una parte no codificante del genoma. En general, los polimorfismos genéticos están causados por la mutación en un gen. La mutación puede ser una mutación de desplazamiento de marco de lectura, una delección de un gen o de una parte del gen, la repetición de un gen, la inserción de un gen o un intercambio de un solo nucleótido.

La delección de un gen o de una parte del mismo y la alteración del número de copia de un gen conducen a una variante de número de copia (VNC). Las VNC pueden estar provocadas por reorganizaciones genómicas, tales como
20 delecciones, duplicaciones y traslocaciones. Las VNC se encuentran asociadas a diferentes tipos de cáncer, tales como, por ejemplo, el cáncer pulmonar de células no pequeñas, aunque también asociadas a autismo y a esquizofrenia.

Un intercambio de un solo nucleótido conduce a un polimorfismo de nucleótido simple (PNS). Se asocian numerosas enfermedades a los PNS, tales como, por ejemplo, la anemia falciforme, el trastorno de hipercoagulabilidad asociado a la variante factor V Leiden, la predisposición al asma y el cáncer (por ejemplo los oncogenes).
30

Por lo tanto, en particular, los polimorfismos de nucleótidos simples (PNS) y las variantes de número de copia (VNC) desempeñan un papel crucial.

En consecuencia, las herramientas analíticas para detectar una mutación, en particular una VNC o un PNS, en muestras biológicas presentan un gran impacto en la investigación actual, más particularmente una VNC o un PNS en muestras de pacientes presentan un gran impacto sobre la medicina actual. En particular, en vista del creciente envejecimiento de la población y en un contexto en el que la atención sanitaria y la protección de la salud presentan una creciente importancia, la detección de una predisposición genética o de una enfermedad genética resulta cada vez más importante. Además, el campo todavía bastante nuevo de la medicina personalizada se basa
40 esencialmente en la detección de los polimorfismos genéticos.

De acuerdo con lo anteriormente expuesto, en los últimos años se han desarrollado varios métodos para detectar polimorfismos de nucleótidos simples (PNS) y variantes de número de copia (VNC). Por ejemplo, los polimorfismos genéticos pueden detectarse mediante métodos tales como, por ejemplo, la secuenciación del gen de interés, la extensión de bases individuales (EBI), las micromatrices de ácido desoxirribonucleico (ADN), tales como, por ejemplo, una matriz de PNS o un chip Affymetrix™, la PCR basada en sondas TaqMan®, la hibridación genómica comparativa basada en matrices o la hibridación in situ comparativa.
45

El documento nº WO 2004/1135 da a conocer bibliotecas de sondas de ANB en las que se encuentran cubiertas todas las posibilidades de secuencia. Las sondas con secuencias definidas se utilizan en ensayos de detección de polimorfismos, entre otros ensayos. El documento nº WO 2004/113563 no proporciona composiciones con sondas que comprenden nucleótidos de ANB aleatorizados según las reivindicaciones.
50

Sin embargo, para la totalidad de los métodos anteriormente indicados, se requiere la síntesis de sondas altamente específicas. De esta manera, para cada locus potencial de un polimorfismo o de una mutación, deben sintetizarse por lo menos dos sondas altamente específicas representativas de dos genotipos diferentes y compararse entre sí. En consecuencia, para cada alelo en cuestión debe sintetizarse individualmente una sonda. Por lo tanto, el análisis de numerosos loci polimórficos potenciales en un experimento requiere mucho tiempo, y es laborioso y caro. Para algunos de los métodos anteriormente indicados, además debe marcarse cada sonda. Este procedimiento de marcaje nuevamente requiere mucho tiempo, y es laborioso y caro. Otros métodos, tales como los métodos basados en matrices, proporcionan miles de diferentes variantes pero resultan difíciles de analizar y no proporcionan resultados cuantitativos. Además, estos métodos no consiguen detectar mutaciones desconocidas. Con el fin de conseguir una especificidad elevada, las sondas deben presentar una determinada longitud mínima. En
60

consecuencia, las sondas con una especificidad elevada son comparablemente largas y, por lo tanto, resultan difíciles de sintetizar. Además, dichas sondas largas con frecuencia no consiguen detectar los polimorfismos de nucleótidos simples (PNS) debido a que las diferencias racionales entre un apareamiento total y un único desapareamiento son excesivamente pequeñas.

Por otra parte, existen métodos para la cuantificación de los niveles de expresión. Por ejemplo, puede llevarse a cabo una reacción en cadena de la polimerasa (PCR) en un ciclador de PCR en tiempo real, tal como, por ejemplo, un LightCycler®. La PCR también puede ser una PCR de transcriptasa inversa. Además, la PCR que se lleva a cabo en un ciclador de PCR en tiempo real puede combinarse con pigmentos de ADN intercalantes, tales como, por ejemplo, SYBR Green I, o con una sonda marcada, tal como una sonda específica o semiespecífica. Además, dicha sonda semiespecífica puede ser, por ejemplo, una sonda Universal Probe Library™ (UPL). Una UPL consiste de sondas que comprenden nucleótidos de ANB (ácido nucleico bloqueado) en lugar de nucleótidos de ADN. Las sondas están diseñadas de manera que se unen a numerosas dianas. Algunas de las sondas se unen a más de 7.000 transcritos en un único transcriptoma. Sin embargo, los métodos basados en dichas sondas no consiguen regularmente detectar o cuantificar mutaciones definidas. Por lo tanto, la utilización de una sonda semiespecífica, tal como, por ejemplo, una sonda UPL, no consigue unirse a un locus definido específicamente.

Por lo tanto, todavía existe una necesidad no satisfecha de una sonda que consista de una cadena de ácidos nucleicos corta que pueda, por una parte, utilizarse universalmente para la detección de diversas secuencias diana y que sea, por otra parte, específica de alelo y permita la detección de una mutación específica, en particular de un polimorfismo de nucleótido simple (PNS).

En el contexto de la presente invención, se proporciona un juego de sondas que comprende nucleótidos de ADN y de ANB. Preferentemente, en el extremo 5', las nucleobases se encuentran determinadas, mientras que en el extremo 3', existen uno o más, preferentemente dos o tres, nucleótidos aleatorios (posiciones tambaleantes). Este diseño de sonda permite que un número controlable de sondas resulte suficiente para la aplicación a una multitud de problemas.

Inesperadamente, al utilizar el juego de sondas definido en la presente memoria, las sondas de unión a dianas específicas resultaron suficientes para discriminar diferentes alelos de mutaciones específicas, en particular polimorfismos de nucleótidos simples (PNS). Concomitantemente, cada juego de sondas puede utilizarse inesperadamente para muchos genes diana diferentes para detectar los PNS.

Un primer aspecto de la presente invención se refiere a una composición que comprende un primer juego de sondas y un segundo juego de sondas, presentando cada una de las sondas ocho nucleótidos, en el que los ocho nucleótidos están compuestos de uno a tres nucleótidos de ADN y cinco a siete nucleótidos de ANB (ácido nucleico bloqueado), en el que todas las sondas del primer y segundo juegos de sondas presentan secuencias de nucleótidos idénticas con la excepción de:

- (i) la base o bases en una, dos o tres posiciones aleatorias de ANB, y
- (ii) la base en una posición discriminante,

en el que la posición o dos o tres posiciones aleatorias de ANB y la posición discriminante se encuentran situadas en posiciones idénticas en todas las sondas del primer y segundo juegos,

en el que en cada posición aleatoria de ANB, la base se selecciona independientemente de entre adenina, citosina, guanina y timina y cualquier secuencia posible resultante de la variación o variaciones de bases en una, dos o tres posiciones aleatorias de ANB se encuentra representada por como mínimo una sonda en cada juego de sondas, y en el que la base en la posición discriminante es idéntica dentro de cada juego de sondas pero difiere entre el primer y el segundo juegos de sondas.

En el contexto de la presente invención, el término "composición" puede entenderse en el sentido más amplio como cualquier mezcla que comprende un primer juego de sondas y un segundo juego de sondas. La composición puede comprender además un solvente adecuado para las sondas de la presente invención. Este solvente puede ser, por ejemplo, agua, un tampón acuoso (que comprende, por ejemplo, TAPS (ácido 3-[[tris(hidroximetil)metil]amino]propanosulfónico) o bicina (N,N-bis(2-hidroxietil)glicina), Tris (tris(hidroximetil)metilamina), tricina (tris(hidroximetil)metilglicina), HEPES (ácido 2-hidroxietil-1-piperazín-etanosulfónico), TES (ácido 2-[[tris(hidroximetil)metil]amino]etanosulfónico), MOPS (ácido 3-(N-morfolino)propanosulfónico), PIPES (ácido piperazín-N,N'-bis(2-etanosulfónico)), cacodilato (ácido dimetil-arsínico), CSS (citrato sódico salino), MES (ácido 2-(N-morfolino)etanosulfónico), fosfato, hidrógenofosfato, dihidrógenofosfato, citrato, acetato, y/o borax), un solvente orgánico (por ejemplo dimetilsulfóxido (DMSO), dimetilformamida (DMF)) o una combinación de los mismos. Además, la composición puede contener una o más sales inorgánicas y/o orgánicas, en particular la composición puede contener sales de magnesio, sodio, potasio, calcio, cloro, fosfato, hidrogenofosfato, dihidrógenofosfato, citrato, acetato y/o bórax. La composición puede comprender además otras sustancias, tales

como, por ejemplo, sustancias biológicas (por ejemplo proteínas, péptidos, aminoácidos, sacáridos, lípidos, etc.), polímeros sintéticos (por ejemplo polietilén-imina (PEI), hidroxipropilmetacrilamida (HPMA), polietilenglicol (PEG)), pigmentos fluorescentes de tinción del ADN (por ejemplo verde SYBR, bromuro de etidio), uno o más detergentes, surfactantes y/o emulsionantes (por ejemplo dodecilsulfato sódico (SDS)), uno o más quelantes (por ejemplo etilén-diamina-tetraacetato (EDTA)), agentes terapéuticos o combinaciones de dos o más de los mismos. Alternativamente, la composición puede secarse o liofilizarse. Alternativamente, las sondas también pueden inmovilizarse sobre un soporte sólido, tal como una superficie de matriz y/o una superficie de perla.

Tal como se utiliza en el contexto de la presente invención, el término "nucleótido" puede entenderse en el sentido más amplio como un monómero de un ácido desoxirribonucleico (ADN) o una cadena de ácido nucleico bloqueado (ANB) (o un análogo de ANB). Tal como entenderá el experto en la materia, cada nucleótido de la sonda comprende una nucleobase. Los términos "base y "nucleobase" tal como se utilizan en la presente memoria pueden entenderse intercambiamente en el sentido más amplio entendido por el experto en la materia. La nucleobase puede ser, por ejemplo, adenina (A), timina (T), citosina (C), guanina (G), uracilo (U) o metilcitosina (mC).

Los nucleótidos en la presente invención se conjugan mediante formación de éster de los grupos 3'- y 5'-hidroxilo de la ribosa, desoxirribosa y/o derivados de ribosa (por ejemplo ribosa bloqueada) con aniones fosfato, tal como es ampliamente conocido de la técnica. Puede no haber ninguna otra fracción molecular unida covalentemente que una los nucleótidos de la cadena con otros aparte de las fracciones nucleótidos y anión fosfato. En particular, puede no haber conectores o espaciadores situados entre los nucleótidos. En el contexto de la presente invención, la longitud de la cadena de ácidos nucleicos normal es de ocho nucleótidos. Se entenderá que esta longitud se refiere a la longitud de la cadena de ácidos nucleicos, aunque puede no excluir que otras fracciones moleculares (tales como, por ejemplo, uno o más fluoróforos, uno o más inhibidores, una o más fracciones de unión o similares) puedan ser añadidas a la sonda, en particular pueden ser añadidas al extremo de la sonda. El término "extremo" de la sonda puede entenderse como el "extremo 3'" o el "extremo 5'" de la sonda. En la presente memoria, los términos "extremo" y "final" pueden entenderse intercambiamente. En particular, una fracción molecular puede conjugarse con el grupo 3'- y/o 5'-hidroxilo de la sonda. El experto en la materia observará que las expresiones "extremo 5'" tal como se utiliza en la presente memoria puede referirse al extremo 5' de la cadena de nucleótidos, pero puede no excluir que en el extremo 5' se añada otra fracción molecular (tal como, por ejemplo, un fluoróforo, un inhibidor, una fracción de unión o similar) al extremo 5' de la sonda. El experto en la materia observará que las expresiones "extremo 3'" tal como se utiliza en la presente memoria puede referirse al extremo 3' de la cadena de nucleótidos, pero puede no excluir que en el extremo 3' se añada otra fracción molecular (tal como, por ejemplo, un fluoróforo, un inhibidor, una fracción de unión o similar) al extremo 3' de la sonda.

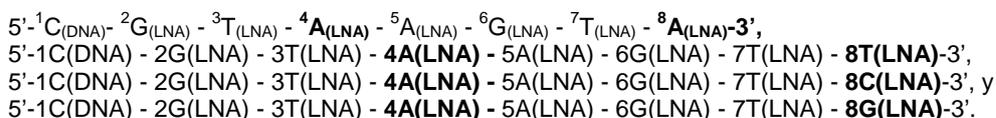
Tal como se utiliza en el contexto de la presente invención, las expresiones "ácido nucleico bloqueado" y "ANB", respectivamente, pueden entenderse en el sentido más amplio como un nucleótido, en el que el anillo de ribosa se encuentra "bloqueado" por un puente adicional que conecta el átomo de oxígeno 2' con el átomo de carbono 4' del nucleótido (por ejemplo un puente metileno) (documento nº WO 99/14226). Por lo tanto, el ANB puede entenderse como ARN modificado inaccesible, en el que el puente "bloquea" la ribosa en la conformación de extremo 3'. Los nucleótidos de ANB, así como los oligómeros de ANB se encuentran disponibles comercialmente. La conformación de ribosa bloqueada es conocido que incrementa el apilamiento de las bases y la preorganización del esqueleto, incrementando significativamente las propiedades de hibridación con una cadena diana de ADN o ARN y, por lo tanto, incrementando la fuerza de unión por cada par de bases (estabilidad térmica/temperatura de fusión incrementadas). Los ANB han sido utilizados para micromatrices de ADN, como sondas de FISH y como sondas de PCR en tiempo real. Los ANB son ampliamente resistentes a la actividad de endo- y exo-nucleasas. Son nucleótidos de ANB alternativos, por ejemplo, ANE (ácido nucleico con puentes 2'-O,4'-C-etileno) o derivados 2'-amino-ANB (ver K. Morita et al., Bioorg. Med. Chem. Lett. 12:73-76, 2002, y S.K. Singh et al., J. Org. Chem. 63:10035, 1998) o derivados 5'-metilo de ANB (ver el documento nº WO 2010/077578). Son alternativas adicionales, los ácidos nucleicos inmovilizados ANI (ver las patentes US nº 7.427.672 y nº 7.217.805) o el ácido nucleico bicíclico hexitol (ver el documento nº WO 2009/100320). El ANB puede combinarse con otros nucleótidos, tales como, por ejemplo, nucleótidos de ADN. Además, estos oligómeros se encuentran disponibles comercialmente. Tal como se utiliza en la presente memoria, la cadena de ácidos nucleicos tal como se utiliza en el contexto de la presente invención es una molécula que comprende n nucleótidos de ADN (n=1, 2 ó 3) y 8-n nucleótidos de ácido nucleico bloqueado (ANB).

Tal como se utiliza en la presente memoria, las expresiones "primer juego de sondas" y "segundo juego de sondas" se refieren a dos juegos de sondas, en los que ambos juegos de sondas presentan secuencias de nucleótidos idénticas excepto por la base o bases en la posición o posiciones aleatorias del ANB y la base en una posición discriminante.

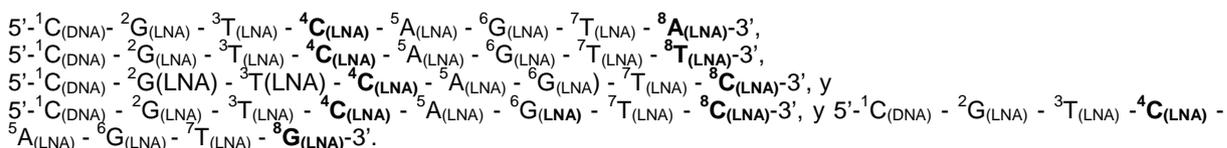
En el contexto de la presente invención, la expresión "posición aleatoria de ANB" se refiere a una posición en la secuencia de nucleótidos en la que la nucleobase es cualquier nucleótido del grupo que consiste de adenina (A), timina (T), citosina (C) o guanina (G). Alternativamente, la nucleobase también puede ser uracilo (U) o metil-citosina (mC) u otra nucleobase que pueda formar un par de bases con una nucleobase complementaria. En el contexto de

la presente invención, las expresiones "posición aleatoria" y "posición de tambaleo" pueden entenderse intercambiamente. La nucleobase se selecciona independientemente de entre las nucleobases A, T, C y G. En el contexto de la presente invención, existen una, dos o tres posiciones aleatorias de ANB en cada sonda del primer y segundo juegos de sondas. Se entenderá que un juego de sondas de la presente invención puede contener 4ⁿ sondas de diferente secuencia, en las que n se refiere al número de posiciones aleatorias del ANB. De esta manera, en el caso de que exista una posición aleatoria, tal como se ejemplifica posteriormente, n=4 y 4¹=4, el juego de sondas contendría 4 sondas diferentes. De acuerdo con lo anterior, en el caso de que existan dos posiciones aleatorias de ANB en cada sonda, n=2 y 4²=16, el juego de sondas contendría 16 sondas diferentes y en el caso de que existan tres posiciones aleatorias de ANB en cada sonda, n=3 y 4³=64, el juego de sondas contendría 64 sondas diferentes. La posición aleatoria de ANB o dos o tres posiciones aleatorias de ANB se encuentran situadas en la misma posición en todas las sondas del primer y segundo juegos.

Con el fin de ejemplificar el diseño de las sondas en un ejemplo no limitativo, una composición que comprende un primer juego de sondas y un segundo juego de sondas con una posición aleatoria de ANB puede presentar, por ejemplo, las secuencias siguientes. El primer juego de sondas puede ser una mezcla de las sondas siguientes:



El correspondiente segundo juego de sondas puede ser una mezcla de las sondas siguientes:



En la presente memoria, los superíndices 1 a 8 caracterizan la posición del nucleótido en la sonda desde el extremo 5'. El código de tres letras de subíndice entre paréntesis indica si el nucleótido es un nucleótido de ácido desoxirribonucleico (ADN) o un nucleótido de ácido nucleico bloqueado (ANB).

Tal como puede observarse a partir del ejemplo anteriormente indicado, las sondas presentan una longitud de ocho nucleótidos. El nucleótido en la posición 1 desde el extremo 5' es un nucleótido de ADN. Los nucleótidos en las posiciones 2 a 8 desde el extremo 5' son nucleótidos de ANB. La posición aleatoria del ANB se encuentra situada en la posición 8 desde el extremo 5'. Por lo tanto, cada juego de sondas contiene cuatro sondas diferentes. En el ejemplo anterior, la posición discriminante se encuentra situada en la posición 4 desde el extremo 5' de las sondas. Las cuatro sondas del primer juego de sondas portan una nucleobase adenina (A) en la posición 4 desde el extremo 5', por lo tanto en la posición discriminante, mientras que las cuatro sondas del segundo juego de sondas portan una nucleobase citosina (C) en la posición discriminante.

Se entenderá que el ejemplo anterior pretende explicar el diseño de sonda de una sonda típica de la presente invención, pero no pretender limitar el alcance de la presente invención.

Tal como se utiliza en la presente memoria, la expresión "posición discriminante" se refiere a una posición en la sonda en la que un nucleótido difiere en los dos juegos de sondas. La nucleobase en la posición discriminante es idéntica dentro de cada juego de sondas pero difiere entre el primer y el segundo juegos de sondas. Puede sustituirse una fracción de adenina por timina, citosina o guanina. Puede sustituirse una fracción de timina por guanina, citosina o adenina. Puede sustituirse una fracción de citosina por timina, guanina o adenina. Puede sustituirse una fracción de guanina por timina, citosina o adenina. En el contexto de la presente invención, la posición discriminante se encuentra situada en la misma posición en todas las sondas del primer y segundo juegos.

La expresión "posición idéntica" tal como se utiliza en la presente memoria se refiere a la posición en la secuencia de la sonda. En el contexto de la presente invención, el término "posición" se refiere a la posición de nucleótido desde el extremo 5' tal como se utiliza de manera habitual para las secuencias de ácidos nucleicos. La expresión "secuencia idéntica" se refiere a dos secuencias cada una de las cuales presenta el mismo tipo de nucleótido, y de esta manera también el mismo tipo de nucleobase, en una posición idéntica de la sonda.

La composición puede comprender dos, tres, cuatro, cinco o más juegos diferentes de sondas, cada uno de los cuales puede detectar un genotipo particular de un locus particular. En el contexto de la presente invención, el término "locus" puede entenderse en el sentido más amplio como una posición en la secuencia de nucleótidos de un gen, en particular un gen diana al que puede unirse la sonda de la invención. El locus también puede designarse

como la parte de la secuencia diana en cuestión o como la secuencia diana. En el locus puede encontrarse situada una mutación potencial. Alternativamente, el locus no comprende una mutación. El locus puede ser un gen en el que se produce una mutación de desplazamiento de marco de lectura, una parte de un gen en el que se produce una mutación de desplazamiento de marco de lectura, un grupo de dos, tres, cuatro, cinco o más nucleótidos o un único nucleótido.

El término "genotipo" puede entenderse en el sentido más amplio como la constitución genética de un organismo o un virus (es decir, la constitución alélica específica del organismo o virus). Tal como se utiliza en la presente memoria, el término "organismo" se refiere a todos los seres vivos, tales como, por ejemplo, bacterias, arqueobacterias, animales, plantas y hongos. Además, el término organismo puede referirse al cuerpo muerto de un ser vivo que estaba vivo anteriormente. El término "virus" puede incluir partículas de tipo vírico.

El término "mutación" tal como se utiliza en la presente memoria puede entenderse en el sentido más amplio como una alteración de la secuencia de nucleótidos. Puede producirse una mutación en una sección codificante del genoma o puede producirse en una sección no codificante del genoma. Una mutación producida en una sección codificante del genoma puede resultar en una secuencia de aminoácidos alterada de un polipéptido codificado por dicho gen al expresarse el mismo. Alternativamente, una mutación puede ser una mutación silenciosa, en la que la secuencia de aminoácidos del polipéptido expresado no resulta afectada o en la que se produce la mutación en una región no codificante del genoma. Una mutación puede producirse de manera natural en una población. Alternativamente, una mutación puede estar provocada por xenobióticos o por radiación. Dichos xenobióticos pueden ser un agente mutagénico tal como es conocido de la técnica (por ejemplo un agente alquilante (por ejemplo mostazas nitrogenadas (por ejemplo ciclofosfamida, mecloretamina o mustina (HN2), uramustina o mostaza uracilo, melfalán, clorambucilo e ifosfamida), nitrosoureas (por ejemplo carmustina, lomustina y estreptozocina), alquilsulfonatos (por ejemplo busulfán), tiotepa y sus análogos, derivados del platino (por ejemplo cisplatino, carboplatino, nedaplatino, oxaliplatino, satraplatino, tetranitrato de triplatino), procarbazona, altretamina, aflatoxina y aflatoxina y productos metabólicos y derivados de los mismos, nitrito, anilina y productos metabólicos y derivados de los mismos, benceno y productos metabólicos y derivados de los mismos, aromáticos policíclicos y productos metabólicos y derivados de los mismos), nitrosaminas, arsénico, asbesto, berilio y sus compuestos, óxido de etileno, compuestos de cromo hexavalente (VI), radón, cloruro de vinilo, tabaquismo, etc.). Dicha radiación puede ser, por ejemplo, radiación ultravioleta (UV), radiación de rayos X, radiación radioactiva/nuclear (por ejemplo radiación alfa, beta o gamma) o radiación cósmica.

En la presente memoria, el término "gen" puede entenderse en el sentido más amplio conocido de la técnica, como una unidad de la secuencia de nucleótidos de un genoma tal como se conoce en la técnica. El genoma es la entidad de los genes de un organismo.

La mutación puede ser, por ejemplo, una mutación de desplazamiento de marco de lectura, una deleción de un gen o de una parte del gen, la repetición de un gen, la inserción de un gen o un intercambio de un solo nucleótido. Adicionalmente, también la conjugación de las fracciones moleculares con una o más nucleobases puede entenderse como una mutación, en particular en el caso de que dicha conjugación pueda conducir a alteraciones en el producto de transcripción y/o traducción. Dicha conjugación puede ser cualquier conjugación conocida de la técnica, tal como, por ejemplo, la metilación de una nucleobase, la pérdida de un grupo metileno de una nucleobase, la conjugación de un agente alquilante (por ejemplo las mostazas nitrogenadas (por ejemplo la ciclofosfamida, la mecloretamina o la mustina (HN2), la uramustina o la mostaza uracilo, el melfalán, el clorambucilo y la ifosfamida), las nitrosoureas (por ejemplo la carmustina, la lomustina y la estreptozocina), los alquilsulfonatos (por ejemplo el busulfán), la tiotepa y sus análogos, los derivados de platino (por ejemplo el cisplatino, el carboplatino, el nedaplatino y el oxaliplatino, el satraplatino y el tetranitrato de triplatino), la procarbazona, la altretamina, la aflatoxina y productos metabólicos y derivados de los mismos, el nitrito, la nitrosamina, la anilina y los productos metabólicos y derivados de los mismos, el benceno y los productos metabólicos y los derivados de los mismos, los aromáticos policíclicos y los derivados de los mismos) con una nucleobase.

En el contexto de la presente invención, la mutación preferentemente es una deleción de un gen o de una parte del gen, la repetición de un gen o el intercambio de un solo nucleótido. Más preferentemente, la mutación es el intercambio de nucleótido simple (PNS).

Una mutación puede resultar en una secuencia de nucleótidos alterada, tal como una secuencia de ADN alterada. Por lo tanto, una mutación puede resultar en diferentes alelos de un gen. Alternativamente, una mutación puede resultar en diferentes alelos en una parte no codificante del genoma. La existencia de por lo menos dos alelos diferentes puede entenderse como un polimorfismo. Sin embargo, también puede producirse un polimorfismo de manera natural en toda la población. Por lo tanto, el término "polimorfismo", tal como se utiliza en la presente memoria, puede entenderse en el sentido más amplio como la incidencia de diferentes genotipos de un gen específico. El término "genotipo" puede entenderse en el sentido más amplio como las secuencias de nucleótidos en un gen. Las diferentes formas de un gen producidas debida a un polimorfismo, es decir las formas polimórficas,

también pueden denominarse "alelos". En la población global pueden existir dos, tres, cuatro, cinco, seis o más alelos diferentes de un gen. Tal como se utiliza en el contexto de la presente invención, el término "polimorfismo" no se refiere a determinada incidencia de un gen mutado en toda la población y puede comprender además un único individuo que muestre un genotipo específico.

5 El polimorfismo puede resultar en diferentes fenotipos o puede ser silencioso. En la presente memoria, el término "silencioso" se refiere a que, aunque existen genotipos diferentes, los fenotipos no son distinguibles mediante los métodos conocidos de la técnica.

10 Preferentemente, el polimorfismo conduce a diferentes fenotipos. Algunos fenotipos pueden conducir a determinadas enfermedades o condiciones patológicas que se producen como resultado directo de las secuencias de nucleótidos alteradas (por ejemplo cáncer, anemia de células falciformes, hipercoagulabilidad). Alternativamente, diferentes fenotipos pueden resultar en diferencias en la predisposición a determinadas enfermedades o condiciones patológicas (por ejemplo cáncer, autismo, esquizofrenia y diabetes mellitus) y/o diferentes fenotipos pueden resultar en diferencias en el metabolismo, tales como, por ejemplo, polimorfismos de enzimas metabolizadores xenobióticos en la alcohol deshidrogenasa, aldehído deshidrogenasa, glutatión-S-transferasa, glucuronosil-transferasa o un elemento de la familia del citocromo P450 (CYP).

20 Un polimorfismo que resulte de un intercambio de nucleótido simple puede denominarse polimorfismo de nucleótido simple (PNS). Un polimorfismo que resulte de una delección de un gen o de una parte del mismo o de la alteración del número de copia de un gen puede denominarse variante de número de copia (VNC).

25 Puede utilizarse una única posición discriminante para detectar una mutación puntual en un determinado locus del ADN diana. Tal como se utiliza en la presente memoria, las expresiones "mutación puntual", "mutación de un único par de bases", "sustitución de una única base" u otras expresiones conocidas por el experto en la materia pueden entenderse intercambiamente. La posición en el gen diana puede entenderse como el locus de interés.

30 En una realización preferente de la presente invención, la posición discriminante se encuentra en la posición 2, 3, 4, 5, 6 ó 7 desde el extremo 5' en cada sonda, preferentemente en la posición 3, 4 ó 5, más preferentemente en la posición 4.

El nucleótido en la posición 1 desde el extremo 5' puede ser un nucleótido de ANB o un nucleótido de ADN. En la realización más preferente, el nucleótido en la posición 1 desde el extremo 5' es un nucleótido de ADN.

35 Cada sonda de la presente invención comprende entre uno y tres nucleótidos de ADN y entre cinco y siete ANB. El experto en la materia podrá entender que la sonda puede comprender además uno o más fracciones no nucleotídicas, tales como, por ejemplo, uno o más fluoróforos, uno o más inhibidores, uno o más conectores (por ejemplo un conector alquilo, un conector PEG, un conector peptídico, un conector sacárido), uno o más pigmentos no fluorescentes (por ejemplo una fracción dinitrofenilo o verde malaquita), uno o más fracciones de unión que pueden unirse a otras moléculas (por ejemplo una maleimida, un isotiocianato o un éster activo (por ejemplo éster de succinimidilo, éster de p-nitrofenilo)) y/o una o más fracciones selectivamente ligantes a moléculas de alto peso molecular (por ejemplo biotina, metotrexato o glucocorticoides). Una sonda conjugada con biotina puede detectarse mediante la utilización de estreptavidina marcada. Una sonda conjugada con metotrexato puede detectarse mediante la utilización de dihidrofolato reductasa (DHFR). Una sonda conjugada con un glucocorticoide puede detectarse con un anticuerpo o un derivado de anticuerpo (por ejemplo un fragmento Fab, un anticuerpo de cadena sencilla, un diacuerpo, un triacuerpo o un tandab).

En una realización preferente de la presente invención, cada sonda consiste de:

- 50
- un nucleótido de ADN y siete de ANB,
 - dos nucleótidos de ADN y seis de ANB, o
 - tres nucleótidos de ADN y cinco de ANB.

55 El experto en la materia entenderá que, en la presente memoria, la expresión "consiste de" se refiere meramente al contenido de nucleótidos de las sondas, pero que no excluye que la sonda pueda comprender además fracciones no nucleotídicas, tales como, por ejemplo, fluoróforos, inhibidores, moléculas de unión y conectores tal como se especifica en la presente memoria.

60 Más preferentemente, el nucleótido de ADN o los dos o tres nucleótidos de ADN de la sonda pueden encontrarse situados próximos al extremo 5' de dicha sonda. Todavía más preferentemente, por lo menos uno de los cinco nucleótidos 5'-terminales es un nucleótido de ADN. Todavía más preferentemente, la sonda presenta un nucleótido de ADN en la posición 1 desde el extremo 5', la sonda presenta un nucleótido de ADN en la posición 2 desde el extremo 5',

Todavía más preferentemente, únicamente el nucleótido en la posición 1 desde el extremo 5' es un nucleótido de ADN.

5 De acuerdo con lo anteriormente expuesto, los 5 a 7 nucleótidos en las posiciones 1, 2, 3, 4, 5, 6, 7 y/o 8 desde el extremo 5' pueden ser nucleótidos de ANB. Preferentemente, por lo menos los nucleótidos en las posiciones 4, 5 y 6 desde el extremo 5' son nucleótidos de ANB, más preferentemente por lo menos los nucleótidos en las posiciones 4, 5, 6 y 7 desde el extremo 5' son nucleótidos de ANB, todavía más preferentemente por lo menos los nucleótidos en las posiciones 4, 5, 6, 7 y 8 desde el extremo 5' son nucleótidos de ANB, todavía más preferentemente por lo menos los nucleótidos en las posiciones 3, 4, 5, 6, 7 y 8 desde el extremo 5' son nucleótidos de ANB, todavía más preferentemente los nucleótidos en las posiciones 2, 3, 4, 5, 6, 7 y 8 desde el extremo 5' son nucleótidos de ANB.

Más preferentemente, el nucleótido en la posición 1 desde el extremo 5' es un nucleótido de ADN y los nucleótidos en las posiciones 2, 3, 4, 5, 6, 7 y 8 desde el extremo 5' son nucleótidos de ANB.

15 En cada posición, la sonda puede presentar un nucleótido determinado o un nucleótido aleatorio. Las posiciones pueden ser iguales para todas las sondas de un juego de sondas. Preferentemente, la composición comprende por lo menos dos juegos de sondas. La composición puede comprender uno, dos, tres, cuatro o más juegos de sondas.

20 Tal como se utiliza en la presente memoria, la expresión "nucleótido determinado" se refiere a una posición de un determinado nucleótido en la sonda, en la que el tipo del nucleótido (por ejemplo adenina (A), timina (T), citosina (C), guanina (G), uracilo (U), 5-metilcitosina (mC)) es conocido. Preferentemente, la nucleobase en una posición determinada es A, T, C, G o U, más preferentemente, la nucleobase en una posición determinada es A, T, C o G.

25 Preferentemente, en el extremo 5', los nucleótidos están determinados. Más preferentemente, por lo menos el nucleótido en la posición 1 desde el extremo 5' está determinado, por lo menos el nucleótido en la posición 2 desde el extremo 5' está determinado, por lo menos el nucleótido en la posición 3 desde el extremo 5' está determinado, por lo menos el nucleótido en la posición 4 desde el extremo 5' está determinado, por lo menos el nucleótido en la posición 5 desde el extremo 5' está determinado,

30 por lo menos el nucleótido en la posición 6 desde el extremo 5' está determinado, por lo menos el nucleótido en las posiciones 1 y 2 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1 y 3 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1 y 4 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1 y 5 desde el extremo 5' están determinados,

35 por lo menos el nucleótido en las posiciones 1 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2 y 3 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2 y 4 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2 y 6 desde el extremo 5' están determinados,

40 por lo menos el nucleótido en las posiciones 3 y 4 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 3 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 3 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 4 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 4 y 6 desde el extremo 5' están determinados,

45 por lo menos el nucleótido en las posiciones 5 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 2 y 3 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 2 y 4 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 2 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 2 y 6 desde el extremo 5' están determinados,

50 por lo menos el nucleótido en las posiciones 1, 3 y 4 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 3 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 3 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 4 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 1, 4 y 6 desde el extremo 5' están determinados,

55 por lo menos el nucleótido en las posiciones 1, 5 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2, 3 y 4 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2, 3 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2, 3 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2, 4 y 5 desde el extremo 5' están determinados,

60 por lo menos el nucleótido en las posiciones 2, 4 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 2, 5 y 6 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 3, 4 y 5 desde el extremo 5' están determinados, por lo menos el nucleótido en las posiciones 3, 4 y 6 desde el extremo 5' están determinados,

por lo menos el nucleótido en las posiciones 3, 5 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 4, 5 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 3 y 4 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 3 y 5 desde el extremo 5' están determinados,
 5 por lo menos el nucleótido en las posiciones 1, 2, 3 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 4 y 5 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 4 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 5 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 3, 4 y 5 desde el extremo 5' están determinados,
 10 por lo menos el nucleótido en las posiciones 1, 3, 4 y 6 desde el extremo 5' están determinados,
 por lo menos los nucleótidos en las posiciones 2, 3, 4 y 5 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 2, 3, 4 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 2, 3, 5 y 6 desde el extremo 5' están determinados,
 15 por lo menos el nucleótido en las posiciones 1, 2, 3, 4 y 5 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 3, 4 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 3, 5 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 2, 4, 5 y 6 desde el extremo 5' están determinados,
 por lo menos el nucleótido en las posiciones 1, 3, 4, 5 y 6 desde el extremo 5' están determinados,
 20 por lo menos el nucleótido en las posiciones 2, 3, 4, 5 y 6 desde el extremo 5' están determinados, o
 por lo menos los nucleótidos en las posiciones 1, 2, 3, 4, 5 y 6 desde el extremo 5' están determinados.

Todavía más preferentemente, por lo menos el nucleótido en la posición 1 desde el extremo 5' está determinado,
 todavía más preferentemente por lo menos los nucleótidos en las posiciones 1 y 2 desde el extremo 5' están
 25 determinados, todavía más preferentemente por lo menos los nucleótidos en las posiciones 1, 2 y 3 desde el
 extremo 5' están determinados, todavía más preferentemente por lo menos los nucleótidos en las posiciones 1, 2, 3
 y 4 desde el extremo 5' están determinados, todavía más preferentemente por lo menos los nucleótidos en las
 posiciones 1, 2, 3, 4 y 5 ó los nucleótidos en las posiciones 1, 2, 3, 4, 5 y 6 desde el extremo 5' están determinados.

Preferentemente, la posición o posiciones aleatorias se encuentran situadas en las posiciones 5, 6, 7 ó 8 desde el
 30 extremo 5'. Más preferentemente, dos posiciones aleatorias se encuentran situadas en las posiciones 5 y 6, 5 y 7, 5
 y 8, 6 y 7, 6 y 8 ó 7 y 8 desde el extremo 5'; tres posiciones aleatorias se encuentran situadas en las posiciones 5, 6
 y 7 desde el extremo 5', en las posiciones 5, 6 y 8 desde el extremo 5', en las posiciones 5, 7 y 8 desde el extremo
 5', en las posiciones 6, 7 y 8 desde el extremo 5', o cuatro posiciones aleatorias en las posiciones 5, 6, 7 y 8 desde el
 35 extremo 5'. Todavía más preferentemente, dos posiciones aleatorias se encuentran situadas en las posiciones 7 y 8
 desde el extremo 5' o tres posiciones aleatorias se encuentran situadas en las posiciones 6, 7 y 8 desde el extremo
 5'.

En una realización preferente, la composición de la presente invención se caracteriza porque:

- 40 a) el juego de sondas presenta una posición aleatoria de ANB en la posición 5, 6, 7 ó 8 desde el extremo 5', o
- b) el juego de sondas presenta dos posiciones aleatorias de ANB en las posiciones 5 y 6, 5 y 7, 5 y 8, 6 y 7, 6 y
 8, ó 7 y 8 desde el extremo 5', preferentemente en las posiciones 7 y 8, o
- c) el juego de sondas presenta tres posiciones aleatorias de ANB en las posiciones 5, 6 y 7, 5, 6 y 8, ó 6, 7 y 8
 desde el extremo 5', preferentemente en las posiciones 6, 7 y 8.

Las sondas preferentemente pueden presentar la estructura general siguiente: 5'-D-L-L-L-L-L-L-X-3',
 5'-D-L-L-L-L-L-X-L-3',
 5'-D-L-L-L-L-X-L-L-3',
 5'-D-L-L-L-L-L-X-X-3',
 50 5'-D-L-L-L-L-X-L-X-3',
 5'-D-L-L-L-L-X-X-L-3',
 5'-D-L-L-L-L-X-X-L-3',
 5'-D-L-L-L-L-X-L-X-3',
 55 5'-D-L-L-L-L-X-X-L-3',
 5'-D-L-L-L-L-X-X-X-3',
 5'-D-L-L-L-L-X-L-X-X-3',
 5'-D-L-L-L-L-X-X-L-X-3',
 5'-D-L-L-L-L-X-X-X-3',
 60 5'-L-D-L-L-L-L-L-X-3',
 5'-L-D-L-L-L-L-X-L-3',
 5'-L-D-L-L-L-X-L-L-3',
 5'-L-D-L-L-X-L-L-L-3',

5' - L - D - L - L - L - X - X - 3',
 5' - L - D - L - L - X - L - X - 3',
 5' - L - D - L - L - X - L - L - X - 3',
 5' - L - D - L - L - X - X - L - 3',
 5' - L - D - L - L - X - L - X - L - 3',
 5' - L - D - L - L - X - X - L - 3',
 5' - L - D - L - L - L - X - X - X - 3',
 5' - L - D - L - L - X - L - X - X - 3',
 5' - L - D - L - L - X - X - L - X - 3',
 10 5' - L - D - L - L - X - X - X - L - 3',
 5' - L - L - D - L - L - L - X - 3',
 5' - L - L - D - L - L - L - X - L - 3',
 5' - L - L - D - L - L - X - L - L - 3',
 5' - L - L - D - L - X - L - L - L - 3',
 15 5' - L - L - D - L - L - L - X - X - 3',
 5' - L - L - D - L - L - X - L - X - 3',
 5' - L - L - D - L - X - L - L - X - 3',
 5' - L - L - D - L - L - X - X - L - 3',
 20 5' - L - L - D - L - X - L - X - L - 3',
 5' - L - L - D - L - X - X - L - L - 3',
 5' - L - L - D - L - L - X - X - X - 3',
 5' - L - L - D - L - X - L - X - X - 3',
 5' - L - L - D - L - X - X - L - X - 3',
 25 5' - L - L - L - D - L - L - L - X - 3',
 5' - L - L - L - D - L - L - X - L - 3',
 5' - L - L - L - D - L - X - L - L - 3',
 5' - L - L - L - D - X - L - L - L - 3',
 5' - L - L - L - D - L - L - X - X - 3',
 30 5' - L - L - L - D - L - X - L - X - 3',
 5' - L - L - L - D - X - L - L - X - 3',
 5' - L - L - L - D - L - X - X - L - 3',
 5' - L - L - L - D - X - L - X - L - 3',
 35 5' - L - L - L - D - X - X - L - L - 3',
 5' - L - L - L - D - L - X - X - X - 3',
 5' - L - L - L - D - X - L - X - X - 3',
 5' - L - L - L - D - X - X - L - X - 3',
 5' - L - L - L - D - X - X - X - L - 3',
 40 5' - L - L - L - L - D - L - L - X - 3',
 5' - L - L - L - L - D - L - X - L - 3',
 5' - L - L - L - L - D - X - L - L - 3',
 5' - L - L - L - L - X - D - L - L - 3',
 5' - L - L - L - L - D - L - X - X - 3',
 45 5' - L - L - L - L - D - X - L - X - 3',
 5' - L - L - L - L - X - D - L - X - 3',
 5' - L - L - L - L - D - X - X - L - 3',
 5' - L - L - L - L - X - D - X - L - 3',
 5' - L - L - L - L - X - X - D - L - 3',
 5' - L - L - L - L - D - X - X - X - 3',
 50 5' - L - L - L - L - X - D - X - X - 3',
 5' - L - L - L - L - X - X - D - X - 3',
 5' - L - L - L - L - X - X - X - D - 3',
 5' - L - L - L - L - L - D - L - X - 3',
 5' - L - L - L - L - L - D - X - L - 3',
 55 5' - L - L - L - L - L - X - D - L - 3',
 5' - L - L - L - L - X - L - D - L - 3',
 5' - L - L - L - L - L - D - X - X - 3',
 5' - L - L - L - L - L - X - D - X - 3',
 5' - L - L - L - L - X - L - D - X - 3',
 60 5' - L - L - L - L - L - X - X - D - 3',
 5' - L - L - L - L - X - D - X - L - 3',
 5' - L - L - L - L - X - X - L - D - 3',
 5' - L - L - L - L - L - D - X - 3',

5'-L-L-L-L-L-L-X-D-3',
 5'-L-L-L-L-L-X-L-D-3',
 5'-L-L-L-L-X-L-L-D-3',
 5'-L-L-L-L-L-D-X-X-3', ó
 5'-L-L-L-L-X-L-X-D-3',

en los que D es un nucleótido de ADN, cada L es un nucleótido de ANB, cada L es un nucleótido de ANB y cada X es una posición aleatoria de ANB.

Más preferentemente, la sonda puede presentar la estructura siguiente:

5'-D-L-L-L-L-L-L-X-3',
 5'-D-L-L-L-L-L-X-L-3',
 5'-D-L-L-L-L-X-L-L-3',
 5'-D-L-L-L-X-L-L-L-3',
 5'-D-L-L-L-L-L-X-X-3',
 5'-D-L-L-L-L-X-L-X-3',
 5'-D-L-L-L-L-X-X-L-3',
 5'-D-L-L-L-X-L-L-X-3',
 5'-D-L-L-L-L-X-X-L-3',
 5'-D-L-L-L-X-L-X-L-3',
 5'-D-L-L-L-X-X-L-L-3',
 5'-D-L-L-L-L-X-X-X-3',
 5'-D-L-L-L-X-L-X-X-3',
 5'-D-L-L-L-X-X-L-X-3', ó
 5'-D-L-L-L-X-X-X-L-3',

en las que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es una posición aleatoria de ANB.

Todavía más preferentemente, la sonda puede presentar la estructura siguiente:

5'-D-L-L-L-L-L-L-X-3',
 5'-D-L-L-L-L-L-X-X-3', ó
 5'-D-L-L-L-L-X-X-X-3',

en las que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es una posición aleatoria de ANB.

En la realización más preferente de la presente invención, las sondas presentan la estructura general siguiente:

5'-D-L-L-L-L-L-X-X-X-3',

en las que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es una posición aleatoria de ANB.

En otra realización altamente preferente de la presente invención, las sondas presentan la estructura general siguiente:

5'-D-L-L-L-L-L-X-X-3',

en las que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es una posición aleatoria de ANB.

Las sondas pueden encontrarse marcadas. Preferentemente, las sondas de diferentes juegos de sondas se marcan de manera diferente.

En una realización preferente de la presente invención, las sondas del primer juego de sondas se marcan con un primer marcador y las sondas del segundo juego de sondas se marcan con un segundo marcador, en los que el primer marcador es diferente del segundo marcador.

En el contexto de la presente invención, el término "marcador" puede interpretarse intercambiamente con "marcaje" o "fracción detectable", como cualquier molécula o fracción que permita la discriminación de la sonda de otras moléculas por cualesquiera medios conocidos de la técnica. Tal como se utiliza en la presente memoria, el primer y/o el segundo marcadores pueden seleccionarse de entre el grupo que consiste, aunque sin limitación, de marcadores fluorescentes, inhibidores, pigmentos no fluorescentes, fracciones de unión, átomos radioactivos (por ejemplo ³H, ³²P, ³⁵S, ¹⁴C o lantánidos) o átomos pesados (por ejemplo ²H o ¹³C).

En una realización más preferente, el primer y el segundo marcadores son marcadores fluorescentes.

Tal como se utiliza en la presente memoria, las expresiones "marcador fluorescente", "marcador de fluorescencia", "marcaje fluorescente", "marcaje de fluorescencia", "pigmento fluorescente", "pigmento de fluorescencia", "fluoróforo" y "fracción fluorescente" pueden entenderse intercambiamente. Un marcador fluorescente puede entenderse en el sentido más amplio como una fracción molecular que emite luz cuando es excitado por luz de otra longitud de onda. Típicamente la longitud de onda que es emitida por el fluoróforo se encuentra desplazada a una longitud de onda más larga en comparación con la luz de excitación. Este desplazamiento es conocido como "desplazamiento Stokes" o "desplazamiento de Stokes" por el experto en la materia. El desplazamiento Stokes puede ser inferior a 5

5 nm, superior a 5 nm, superior a 10 nm, superior a 20 nm, superior a 30 nm, superior a 50 nm, superior a 75 nm, superior a 100 nm, superior a 150 nm, superior a 200 nm, superior a 250 nm, superior a 300 nm o incluso superior a 400 nm. El máximo o máximos de absorbancia del fluoróforo adecuado para la detección fluorescente puede encontrarse a una longitud de onda de entre 100 y 280 nm (luz UV-C), de entre 280 y 315 nm (luz UV-B), de entre 315 y 400 nm (luz UV-A), de entre 400 y 750 nm (luz visible), de entre 750 y 1.400 nm (luz IR-A) y de entre 1.400 y 3.000 nm (luz IR-B). El máximo o máximos de excitación del fluoróforo adecuado para la detección fluorescente puede ser de entre 100 y 280 nm (luz UV-C), de entre 280 y 315 nm (luz UV-B), de entre 315 y 400 nm (luz UV-A), de entre 400 y 750 nm (luz visible), de entre 750 y 1.400 nm (luz IR-A) y de entre 1.400 y 3.000 nm (luz IR-B). Preferentemente, los máximos de absorbancia y de excitación de la fluorescencia se encuentran entre 400 y 750 nm.

15 Los marcadores fluorescentes pueden ser, por ejemplo, fluoresceína, isotiocianato de fluoresceína (FITC), carboxifluoresceína, derivados de fluoresceína, pigmentos de rodamina (rodamina, rodamina B, rodamina 6G, tetrametil-rodamina (TAMRA), isotiocianato de rodamina y otros derivados de rodamina), pigmentos de cianina (por ejemplo Cy3, Cy3.5, Cy5, Cy5.5 y Cy7) y derivados de los mismos, pigmentos LC (por ejemplo LC-amarillo 555, LC-rojo 610, LC-rojo 640, LC-rojo 670, LC-rojo 705) y derivados de los mismos, pigmentos Alexa (Alexa 488, Alexa 546 y Alexa 647) y derivados de los mismos, S0387, pigmento HOECHST y derivados de los mismos, isotiocianato de eritrosina y derivados de los mismos, verde Oregon y derivados de los mismos, amarillo Lucifer y derivados de los mismos, ficoeritrina y derivados de los mismos, FAM y derivados de los mismos, LightCycler® amarillo 555 y derivados de los mismos, VIC y derivados del mismo, HEX y derivados del mismo o puntos cuánticos y derivados de los mismos.

25 En el contexto de los marcadores fluorescentes, la expresión "derivado de los mismos" se refiere a sales del fluoróforo y/o a fluoróforos que se conjugan con fracciones no fluorescentes, tales como, por ejemplo, conectores (por ejemplo conectores alquilo, conector PEG) o fracciones de unión (por ejemplo maleimidas, isotiocianato o un éster activo (por ejemplo éster de succinimidilo, éster de p-nitrofenilo, halogenuros de ácido) o una combinación de los mismos.

30 Los marcadores fluorescentes pueden unirse a las sondas por cualesquiera medios. Pueden conjugarse directamente con las sondas o conjugarse mediante un conector. Dicho conector preferentemente puede presentar una longitud inferior a 5 Å, inferior a 10 Å, inferior a 15 Å, inferior a 20 Å, inferior a 25 Å o inferior a 50 Å. Son conocidos diversos conectores por el experto en la materia (ver, por ejemplo, el documento nº WO 84/03285). Más preferentemente, la longitud del conector es de entre 15 Å y 35 Å.

35 Preferentemente, una condición previa es que el enlace de los marcadores con la sonda no influya negativamente (lo que significa que la reducción es poco o nada significativa (<2°C) de la temperatura de fusión) sobre la temperatura de fusión de una sonda dada con el ADN diana. Por lo tanto, los marcadores se unen a la nucleobase en el caso de que el marcador se encuentre unido a los nucleótidos de ADN o ANB o el marcador se encuentre unido a un análogo de fosfato internucleosídico (documento nº WO 2007/059816). Preferentemente, los marcadores se unen al extremo 3' ó 5' de una sonda. Dichos métodos de marcaje son bien conocidos de la técnica y se encuentra disponible una multitud de bloques constructivos comerciales (ver, por ejemplo, Fluorescent oligonucleotides. Versatile tools as probes and primers for DNA and RNA analysis. Wojczewski, Christian; Stolze, Karen; Engels, Joachim W. Synlett (10), 1667-1678, 1999).

45 La sonda puede marcarse con uno, dos, tres, cuatro, cinco o más marcadores fluorescentes, preferentemente la sonda se marca con uno o dos marcadores fluorescentes.

50 En el caso de que la sonda se marque con dos marcadores fluorescentes, estos marcadores fluorescentes pueden ser del mismo tipo o marcadores fluorescentes diferentes. Preferentemente, los marcadores fluorescentes son marcadores fluorescentes diferentes. Más preferentemente, los marcadores fluorescentes emiten y absorben luz de longitudes de onda diferentes, todavía más preferentemente, los dos o más pigmentos fluorescentes forman una pareja de transferencia de energía por resonancia fluorescente (FRET), todavía más preferentemente, la longitud de onda de emisión del fluoróforo donante difiere del del fluoróforo aceptor en por lo menos 25 nm, todavía más preferentemente en por lo menos 50 nm, en por lo menos 100 nm, en por lo menos 150 nm, en por lo menos 200 nm o en por lo menos 250 nm. Todavía más preferentemente, el espectro de emisión de un fluoróforo (donante) se solapa con el espectro de absorbancia del otro fluoróforo (aceptor).

60 La sonda también puede marcarse con uno o más fluoróforos y uno o más inhibidores. Preferentemente, la sonda puede marcarse con un fluoróforo y un inhibidor.

En una realización todavía más preferente de la presente invención, las sondas son sondas de hidrólisis marcadas adicionalmente con un inhibidor.

Un inhibidor tal como se utiliza en la presente memoria es una estructura molecular que puede inhibir la luz emitida por un fluoróforo. Por lo tanto, un inhibidor que se encuentra a una distancia espacial comparativamente reducida a un fluoróforo puede reducir la intensidad de la luz emitida por el fluoróforo al ser excitado. Este efecto es bien conocido por el experto en la materia. Además, un fluoróforo puede, bajo determinadas circunstancias, actuar como inhibidor.

El inhibidor puede inhibir la luz a una distancia espacial inferior a 10, inferior a 9, inferior a 8, inferior a 7, inferior a 6, inferior a 5, inferior a 4, inferior a 3 ó inferior a 2 nucleótidos entre el fluoróforo y el inhibidor.

Tal como se utiliza en el contexto de la presente invención, las expresiones "inhibidor" e "inhibidor oscuro" pueden entenderse intercambiamente. Un inhibidor puede ser cualquier estructura molecular que pueda reducir eficientemente la intensidad de la fluorescencia emitida por el fluoróforo. Un inhibidor puede ser un fluoróforo o una estructura molecular que no emite luz visible, tal como, por ejemplo, dabsilo (ácido dimetilaminoazosulfónico), un inhibidor Black Hole que inhibe en todo el espectro visible, un pigmento de IR QC-1, un inhibidor QxI, un inhibidor Iowa black FQ que inhibe en la parte verde-amarilla del espectro o un inhibidor Iowa black RQ que inhibe en la parte naranja-roja del espectro. El inhibidor puede emitir radiación térmica.

En juegos alternativos de sondas, las sondas pueden presentar una longitud de 7 nucleótidos (en lugar de 8). La sonda y su aplicación son las detalladas anteriormente, en el contexto de las composiciones, métodos y bibliotecas de la invención, en las que las dos o tres posiciones aleatorias de ANB de las sondas de 8 nucleótidos de longitud se reducen a una o dos posiciones aleatorias de ANB, y en las sondas de 7 nucleótidos de longitud, mediante la delección de una posición aleatoria, respectivamente. El juego de sondas puede diseñarse tal como se ha indicado anteriormente, considerando la delección de una posición aleatoria. Tal como se ha indicado anteriormente, dichos nucleótidos aleatorios preferentemente se sitúan en el extremo 3'.

Sin embargo, en juegos alternativos de sondas, las sondas pueden presentar una longitud de 9, 10, 11, 12, 13, 14 ó 15 nucleótidos (en lugar de 8). La sonda y su aplicación son las detalladas anteriormente, en el contexto de las composiciones, métodos y bibliotecas de la invención, en las que el número de posiciones aleatorias de ANB de las sondas se incrementa en comparación con las sondas de 8 nucleótidos de longitud mediante la introducción de una o más posiciones aleatorias adicionales, es decir, una posición aleatoria adicional para cada nucleótido superior al octavo (por ejemplo dos posiciones aleatorias adicionales para una sonda de 10 nucleótidos). El juego de sondas puede diseñarse tal como se ha indicado anteriormente, considerando la adición de una o más posiciones aleatorias. Tal como se ha indicado anteriormente, dichos nucleótidos aleatorios preferentemente se sitúan en el extremo 3'.

Alternativamente, el término "nucleótido" puede comprender además otra nucleobase aparte de A, T, C o G, tal como es conocida de la técnica (tal como, por ejemplo, uracilo (U) o metil-citosina (mC)) conjugada con una fracción molecular que puede polimerizarse. Por lo tanto, el término "nucleótido" puede referirse alterativamente a un nucleótido ácido ribonucleico (ARN), a un nucleótido análogo de ácido nucleico, tal como es conocido por el experto en la materia (por ejemplo un nucleótido ácido péptido-nucleico (APN), un nucleótido morfolino, un nucleótido ácido glicol-nucleico (AGN), un nucleótido ácido treosa-nucleico (ATN) o un nucleótido de ADN metilado. Un nucleótido análogo de ácido nucleico se selecciona de manera que sea capaz de formar una pareja de bases estable y específica con una base natural. Alternativamente, el ácido nucleico bloqueado (ANB) también puede formar una parte de una cadena que comprende por lo menos un nucleótido de ANB y por lo menos un nucleótido de ADN y que comprende además uno de los siguientes: un nucleótido ácido ribonucleico (ARN), un ácido péptido-nucleico (APN), un morfolino, un ácido glicol-nucleico (AGN), un ácido treosa-nucleico (ATN), ANE (ácido nucleico con puentes 2'-O,4'-C-etileno) y derivados 2'-amino-ANB. Alternativamente, dicho nucleótido o nucleótidos de ADN también pueden sustituirse por uno de los nucleótidos anteriormente indicados. Alternativamente, dicho nucleótido o nucleótidos de ANB también pueden sustituirse por uno de los nucleótidos anteriormente indicados.

En un segundo aspecto, la presente invención se refiere a un método para determinar el genotipo en un locus de interés en una muestra obtenida de un sujeto, comprendiendo el método:

- a) poner en contacto la muestra que comprende material genético con la composición de la presente invención, y
- b) detectar la unión de una sonda del primer o del segundo juego de sondas al material genético, determinando de esta manera el genotipo en el locus.

Tal como se utiliza en la presente memoria, la expresión "determinar el genotipo" se refiere al análisis del genotipo de un organismo o de un virus. La determinación del genotipo puede incluir, aunque sin limitación, el análisis de un alelo (por ejemplo el análisis de si el organismo presenta una mutación puntual en un determinado locus o no, y qué nucleótido o nucleobase ha sido sustituido por qué otro nucleótido o nucleobase), la búsqueda de una o más mutaciones puntuales nuevas en el genoma de un organismo o virus, o la determinación o evaluación del número de copia de un gen o de una parte de un gen.

En el contexto de la presente invención, la expresión "locus de interés" puede entenderse intercambiabilmente con el término "locus" en el sentido más amplio como una posición en la secuencia de nucleótidos de un gen, en particular un gen diana al que puede unirse la sonda de la invención. El locus de interés puede comprender uno, dos, tres, cuatro, cinco, seis, siete u ocho nucleótidos. Puede haber una mutación localizada en el locus de interés, o puede no haber ninguna mutación localizada en el locus de interés. Puede haber un polimorfismo de nucleótido simple (PNS) localizado en el locus de interés, o puede no haber ningún polimorfismo de nucleótido simple (PNS) localizado en el locus de interés.

El término "sujeto" tal como se utiliza en la presente memoria puede entenderse en el sentido más amplio como una fuente de material genético. El material genético puede obtenerse de cualquier material orgánico, en particular de una muestra biológica. La muestra biológica puede ser un organismo vivo o muerto, tal como, por ejemplo, una bacteria viva o muerta, un animal vivo o muerto, un ser humano vivo o muerto, un hongo vivo o muerto o una planta viva o muerta, una parte de un animal, de una planta, de un hongo, o de un orgánulo celular o virus, una partícula de tipo vírico, tal como mitocondria, un leucoplasto o un cloroplasto. Alternativamente, el ADN o ARN puede obtenerse de una expulsión, un raspado o un producto de degradación de los organismos anteriormente indicados. Puede obtenerse de, por ejemplo, una muestra de sangre, una muestra sanguínea embrionaria, una muestra sanguínea fetal, una muestra de linfa, una muestra de sangre de cordón, una muestra de líquido cerebrospinal, una muestra de líquido/fluido amniótico, un frotis bucal, un frotis vaginal, una prueba citológica, caspa, un folículo piloso, una o más células extraídas, una muestra espermática, una célula óvulo, una muestra de saliva, una muestra de orina, una muestra de heces, una muestra de linfa, una muestra de líquido inflamatorio, una muestra de cordón umbilical, una muestra de piel, una muestra de médula ósea, una muestra de mucosa, una muestra de tejido, una muestra de agua, una muestra de suelo, una muestra de sedimento o una escena de un crimen.

La expresión "material genético" puede referirse a cualquier tipo de ácido nucleico natural o sintético que transmite o codifica información genética mediante una secuencia de nucleobases. El material genético puede ser ADN o ARN. El ADN puede ser ADN de doble cadena o ADN de cadena sencilla. El ARN puede ser cualquier tipo de ARN conocido de la técnica, tal como, por ejemplo, ARNm, ARNt, ARN ribosómico, ARN vírico, miARN, ARNip, ARNi, ARNnp. El material genético puede obtenerse de una muestra biológica o puede sintetizarse mediante síntesis orgánica. Tal como se utiliza en la presente memoria, el término "muestra" se refiere a cualquier tipo de material que se analiza mediante la sonda de la presente invención. La muestra puede comprender el material genético de interés.

La muestra puede ser una muestra biológica seleccionada de entre el grupo que consiste de un líquido corporal, sangre, orina, suero, mucosa, esputo, heces, muestra epidérmica, piel, raspado bucal, esperma, líquido amniótico, células en cultivo y médula ósea. Además, el material genético alternativamente puede obtenerse mediante síntesis química tal como es conocido de la técnica. Puede ser parte de un genoma natural o de un genoma de un organismo modificado genéticamente. Puede ser material genético lineal o circular, en particular ADN genómico o un plásmido. El material de ADN o ARN también puede haber sido purificado.

La muestra puede obtenerse de un animal. El animal puede ser cualquier tipo de animal, incluyendo animales unicelulares (protozoos) y animales multicelulares (metazoos). Preferentemente, el animal es un mamífero, incluyendo seres humanos. El término "animal" también puede incluir esporas o quistes de un protozoo y quistes y gametos (células germinales) de un metazoo.

Alternativamente, la muestra puede obtenerse de una planta. La planta puede ser cualquier tipo de planta, incluyendo plantas unicelulares y multicelulares. Preferentemente, la planta es una planta útil, tal como un cultivo agrícola. El término "animal" también puede incluir esporas de un protozoo y semillas, frutos, hojas caídas, polen, savia, y gametos (células germinales) de una planta metazoaria. Alternativamente, la muestra puede obtenerse de una bacteria. La bacteria puede ser cualquier tipo de protozoo, incluyendo eubacterias y arqueobacterias. La bacteria puede ser patogénica o no patogénica. Preferentemente, la bacteria es una bacteria patogénica. El término "bacteria" también puede incluir las esporas de una bacteria.

Alternativamente, la muestra puede obtenerse de cualquier tipo de virus. Los virus también pueden incluir virus y fagos animales y vegetales. El virus puede ser una partícula vírica (virión). Alternativamente, el ADN y/o ARN vírico puede obtenerse del genoma de un virus que se encuentra presente en una célula eucariótica o procariótica (célula huésped). El genoma vírico puede integrarse en el genoma de la célula huésped. Preferentemente, el virus es un virus patogénico.

Alternativamente, la muestra puede obtenerse de un hongo. El hongo puede ser cualquier tipo de hongo. Preferentemente, el hongo es un virus patogénico, que es parásito en el interior o sobre la superficie de un organismo huésped, en particular en el interior o sobre la superficie de un animal, incluyendo seres humanos, o una planta. El término "hongo" también puede incluir las esporas de un hongo.

El ADN o ARN obtenido de una muestra biológica puede estar comprendido de una mezcla biológica en bruto o puede haber sido aislado. La expresión "mezcla biológica en bruto" puede incluir, aunque sin limitación, un lisado celular que puede obtenerse mediante cualquier medio conocido de la técnica (tal como lisis mediante, por ejemplo, un tampón hipotónico, uno o más detergentes, sonicación, alcohol, fuerzas de cizalla (con una prensa francesa, un homogeneizador Potter o Downs o pipeteado brusco), enzimas, raspado o uno o más ciclos de congelación-descongelación ("freeze-and-squeeze" [congelar y apretar]) y un raspado que contiene ADN o ARN. El término "aislado" puede referirse a una muestra seca o liofilizada que, aparte de sales, contiene más de 25%, más de 50%, más de 60%, más de 70%, más de 80%, más de 90%, más de 95% o más de 97% en peso de ADN y/o ARN, o una solución acuosa u orgánica, en la que por lo menos 25%, por lo menos 50%, por lo menos 60%, por lo menos 70%, por lo menos 80%, por lo menos 90%, por lo menos 95% o por lo menos 97% en peso del contenido orgánico, aparte del solvente y las sales, se refiere a ADN y/o ARN.

El ADN o ARN puede aislarse mediante cualquier medio conocido de la técnica, tal como, por ejemplo, precipitación en alcohol (por ejemplo etanol o isopropanol) o una mezcla de diferentes alcoholes, métodos basados en cromatografía (por ejemplo cromatografía de intercambio aniónico, cromatografía de exclusión por tamaño), métodos basados en gel (por ejemplo electroforesis en gel (electroforesis en gel de agarosa, electroforesis en gel de poliacrilamida (PAGE)), electroforesis capilar (EC) o combinaciones de dos o más de los mismos. Pueden combinarse varios métodos de purificación en línea o pueden combinarse mediante dos o más etapas de purificación posteriores. En la presente memoria, los términos "aislamiento" y "purificación" pueden entenderse intercambiamente.

El ADN o ARN también puede ser de origen sintético. En particular, una muestra de control puede ser de origen sintético. La muestra de control puede ser un control positivo o negativo, o puede utilizarse para someter a ensayo la selectividad de las sondas. El ADN sintético puede ser lineal o circular. El ADN diana o ARN diana de origen sintético puede marcarse por cualquier medio, en particular por medios descritos para las sondas en la presente memoria. El ADN sintético puede ser de doble cadena o de cadena sencilla. También puede conjugarse con una superficie, en particular con la superficie de una perla o la superficie de un portaobjetos de plástico, vidrio o metal. El ADN o ARN puede ser parte de una matriz, en particular una micromatriz, más particularmente una micromatriz de ADN (por ejemplo una matriz de PNS o un chip Affymetrix, una hibridación genómica comparativa con matrices), tal como es conocido de la técnica.

La secuencia diana puede ser parte de un ADN diana o de un ARN diana. La secuencia diana puede seleccionarse libremente. Por ejemplo, la secuencia de los nucleótidos determinados puede ser conocida por el experto en la materia o puede obtenerse de una base de datos de secuencias conocida por el experto en la materia, tal como, por ejemplo, la Ensembl Sequence Identifier, la RefSeq, la EMBL Sequence Identifier, la EMBL y/o la base de datos genómica NCBI Entrez. Además, el diseño de la sonda puede ser asistido por ordenador, tal como, por ejemplo, el software de diseño de ensayos ProbeFinder basado en Internet o el software de diseño de cebadores LightCycler (ambos de Roche).

La expresión "ADN diana" tal como se utiliza en la presente memoria puede interpretarse como el ADN que es por lo menos parcialmente complementario a la sonda de la invención. Por lo tanto, bajo determinadas condiciones, la sonda puede unirse a su ADN diana. Puede interpretarse que la sonda se une preferentemente al ADN de cadena sencilla complementario. Puede obtenerse un ADN de cadena sencilla por medios conocidos de la técnica, tales como la fusión de una hebra de ADN de doble cadena, tal como se describe en mayor detalle en la presente memoria.

El ADN diana puede ser de cualquier longitud. Preferentemente, el ADN diana presenta una longitud de por lo menos 8 nucleótidos. El ADN diana puede presentar una longitud de más de 8, más de 9, más de 10, más de 20, más de 50, más de 10^2 , más de 10^3 , más de 10^4 , más de 10^5 , más de 10^6 , más de 10^7 , más de 10^8 , más de 10^9 , más de 10^{10} o más nucleótidos. El experto en la materia podrá entender, en el contexto del ADN de doble cadena, que la expresión "nucleótidos de longitud" también puede referirse a la longitud en pares de bases (pb).

El ARN diana puede ser de cualquier longitud. Preferentemente, el ARN diana presenta una longitud de por lo menos 8 nucleótidos. El ARN diana puede presentar una longitud de más de 8, más de 9, más de 10, más de 15, más de 20, más de 30, más de 50, más de 75, más de 100, más de 150, más de 200, más de 300, más de 500, más de 1.000, más de 2.000 o más nucleótidos. El experto en la materia podrá entender, en el contexto del ARN de doble cadena, que la expresión "nucleótidos de longitud" también puede referirse a la longitud en pares de bases (pb).

El ADN diana puede amplificarse por cualquier medio de la técnica. Por ejemplo mediante reacción en cadena de la polimerasa (PCR), amplificación en bacterias, en levaduras, en células de mamífero o en células de insecto. Además, el ADN puede obtenerse a partir de una reacción de transcriptasa inversa, tal como, por ejemplo, reacción en cadena de polimerasa-transcriptasa inversa. La PCR-transcriptasa inversa también puede utilizarse para analizar el transcriptoma o un ARN mensajero (ARNm) particular.

El juego de sondas puede diseñarse para caracterizar el genotipo del gen diana. Un juego de sondas puede presentar una afinidad de unión más fuerte a un locus de un genotipo de tipo salvaje que el genotipo mutado, mientras que un juego de sondas correspondiente puede presentar una afinidad de unión más fuerte al locus del genotipo mutado. Por lo tanto, ambos juegos de sondas se unen preferentemente a diferentes alelos del mismo locus.

En la presente memoria, la expresión "juego de sondas correspondiente" se refiere a una sonda que presenta una secuencia de nucleótidos idéntica, con la excepción de la base o bases en una, dos o tres posiciones aleatorias de ANB y la base en la posición discriminante, a la sonda a la que corresponde.

Tal como se utiliza en el contexto de la presente invención, la expresión "poner en contacto" puede entenderse en el sentido más amplio como la exposición de uno o más juegos de sondas a la muestra y viceversa. La puesta en contacto puede conseguirse, por ejemplo, mediante la mezcla de una solución que comprende el juego o juegos de sondas con una muestra. Opcionalmente, la puesta en contacto puede estar acompañada de una etapa de calentamiento:

tal como se utiliza en toda la invención, el término "detectar" puede entenderse en el sentido más amplio como la medición de una señal producida a partir de una sonda. La sonda puede estar marcada con un marcador y puede detectarse una señal producida por el marcador. Por lo tanto, por ejemplo puede detectarse la señal de fluorescencia producida por una sonda marcada fluorescentemente excitada. Alternativamente, puede detectarse la señal de transferencia de energía de resonancia fluorescente (FRET) producida por una sonda doblemente marcada, la radiación radioactiva producida por una sonda marcada radioactivamente, la temperatura de fusión de la sonda, la despolarización de la fluorescencia de una sonda marcada fluorescentemente y/o la velocidad de difusión de una sonda marcada fluorescentemente. Alternativamente, puede detectarse una señal FRET y/o la correlación cruzada de señales de fluorescencia (FCCS) de dos sondas marcadas fluorescentemente. Estas sondas pueden interactuar entre sí o con una molécula nucleotídica diana o pueden interactuar con dos cadenas de ADN complementarias. Alternativamente, la inhibición de fluorescencia puede detectarse mediante la pérdida, reducción o ausencia de fluorescencia.

En el contexto de la presente invención, la unión del primer juego de sondas puede compararse con la unión del segundo juego de sondas. En la presente memoria, tras la puesta en contacto de los juegos de sondas con una muestra, puede detectarse la intensidad de la señal producida por el primer juego de sondas. De manera similar, puede detectarse la señal producida por el segundo juego de sondas. Las intensidades de las señales pueden compararse entre sí.

El primer y el segundo juegos de sondas pueden marcarse de manera diferente. A continuación, puede detectarse concurrentemente la señal producida por ambos juegos de sondas. La señal producida por cada marcador puede detectarse separadamente. Por ejemplo, ambos juegos de sondas pueden marcarse con dos fluoróforos diferentes que emiten luz de diferentes longitudes de onda. La luz emitida puede detectarse independientemente de otra mediante cualquier medio conocido de la técnica, por ejemplo por detectores de fluorescencia separados (por ejemplo tubos fotomultiplicadores (TFM), fotodiodos de avalancha (FDA)), o puede detectarse con un único detector, aunque la luz de diferente longitud de onda es separada de otra, por ejemplo por uno o más filtros, uno o más espejos dicróicos, uno o más prismas o un detector Meta.

Alternativamente, ambos juegos de sondas pueden marcarse con el mismo marcador, por ejemplo el mismo fluoróforo. A continuación, puede detectarse a continuación la señal producida por ambos juegos de sondas.

Tal como se indica en la presente memoria, el primer y el segundo juegos de sondas pueden diferir en la posición discriminante. Preferentemente, el primer y el segundo juegos de sondas puede diferir en un único nucleótido. Sin embargo, debido a la diferencia en su secuencia de nucleótidos, los dos juegos de sondas pueden presentar una afinidad de unión diferente a su secuencia diana. La sonda que presenta una complementariedad completa con la cadena de nucleótidos complementaria puede presentar una afinidad más elevada que la sonda que presenta un desapareamiento. Preferentemente, una de las sondas de un juego de sondas muestra una complementariedad completa con la cadena diana; de esta manera, todas las nucleobases forman pares de bases con las nucleobases de la cadena diana, mientras que las sondas del segundo juego de sondas portan por lo menos un único desapareamiento.

Un desapareamiento conducirá a una afinidad de unión comparablemente inferior a la de un apareamiento completo. De esta manera, en el equilibrio, se unirá una fracción mayor del juego de sondas que portan la secuencia que se ajusta mejor. Por lo tanto, la unión de la sonda que muestra un apareamiento completo puede identificarse mediante la obtención de una señal de detección alterada de la sonda que muestra unión.

Sin embargo, el experto en la materia observará que el nucleótido diana también puede presentar una secuencia diferente en un locus determinado.

5 Por lo tanto, tal como se indica en la presente memoria, uno de los juegos de sondas puede diseñarse que presente una preferencia para la unión a un genotipo específico en el locus, mientras que otro juego de sondas puede presentar una preferencia por otro genotipo en el locus. Por lo tanto, al comparar la señal de detección de ambas sondas, puede determinarse el genotipo en el locus.

10 Tal como se utiliza en la presente memoria, la expresión "determinar el genotipo en el locus" puede referirse a la discriminación de diferentes genotipos en un locus de interés en dos o más muestras diferentes. Alternativamente, la expresión "determinar el genotipo en el locus" puede referirse a la identificación del genotipo de una muestra. Además, la expresión "determinar el genotipo en el locus" puede referirse además a la identificación de nuevas variantes de genotipo en el genoma de un sujeto.

15 Por ejemplo, el genotipo de un sujeto, en particular un paciente, puede determinarse mediante el método de la presente invención a fin de seleccionar una determinada terapia. La determinación de genotipos mediante el método de la presente invención también puede utilizarse para cribados epidemiológicos de una población. Además, la determinación de un genotipo mediante el método de la presente invención también puede utilizarse para el tratamiento de una enfermedad en un sujeto, para detectar una determinada especie o variante vegetal, para detectar una determinada especie o cepa bacteriana, para detectar una determinada especie o cepa vírica, para detectar una determinada especie o cepa fúngica, para detectar una determinada especie, cepa, raza o variedad animal, o para investigaciones criminales. Además, la determinación de un genotipo mediante el método de la presente invención puede utilizarse además para la detección de características genotípicas de animales, plantas, virus, bacterias y/o hongos. Tal como es conocido de la técnica, las características genotípicas pueden presentar influencias sobre el fenotipo. Por lo tanto, el método puede utilizarse para predecir las características fenotípicas. Además, el método puede utilizarse para detectar una predisposición de una enfermedad en un animal, incluyendo el ser humano, o una planta o un hongo. Además, el método de la presente invención puede utilizarse para el diagnóstico prenatal. En la presente memoria, la muestra puede obtenerse preferentemente de la sangre o la linfa del embrión o feto, la sangre del cordón, la placenta o del líquido/fluido amniótico.

30 En una realización preferente de la presente invención, el locus es un único nucleótido.

Puede resultar de interés un único nucleótido en una posición particular del ADN diana o del ARN diana. Por lo tanto, puede existir un polimorfismo de nucleótido simple (PNS) en el locus.

35 En una realización preferente adicional, el método comprende:

(a) llevar a cabo una etapa de amplificación que comprende la puesta en contacto de la muestra con un juego de cebadores con el fin de producir un producto de amplificación que incluye el locus de interés,

(b) llevar a cabo una etapa de hibridación que comprende poner en contacto el producto de amplificación de la etapa a) con la composición de la presente invención, y

40 (c) detectar la hibridación de una sonda del primer o segundo juego de sondas con el material genético, determinando de esta manera el genotipo en el locus.

45 La expresión "etapa de amplificación" tal como se utiliza en la presente memoria se refiere a cualquier método conocido de la técnica para amplificar el material genético. El ADN puede amplificarse mediante la utilización de una reacción en cadena de la polimerasa, puede amplificarse en células (por ejemplo en células bacterianas, en células de mamífero o en células de insecto). El ARN puede amplificarse mediante la utilización de PCR con transcriptasa inversa.

50 La expresión "puesta en contacto de la muestra con un juego de cebadores" se refiere a la adición de cebadores a la muestra que comprende el ADN diana. Pueden añadirse uno o más enzimas adicionales, tales como ADN polimerasa, una o más sales de magnesio, nucleótidos trifosfato (dATP, dCTP, dTTP, dGTP) y/o un tampón adecuado. Estos ingredientes son bien conocidos por el experto en la materia y se encuentran disponibles comercialmente.

55 La expresión "etapa de hibridación" puede referirse a cualquier método conocido de la técnica para la hibridación de la sonda con su ADN diana o su ARN diana. Las sondas pueden hibridarse con su ADN diana o ARN diana bajo condiciones en las que un ácido nucleico corto se hibrida habitualmente con su secuencia diana. Dichas condiciones son bien conocidas por el experto en la materia.

60 Para la hibridación, el ADN de doble cadena puede en primer lugar desnaturalizarse; de esta manera, ambas cadenas de ADN pueden separarse una de otra. El ADN diana puede desnaturalizarse, por ejemplo mediante calentamiento de la muestra. La cadena de ADN diana puede desnaturalizarse a más de 40°C, a más de 50°C, a más de 60°C, a más de 70°C, a más de 80°C, a más de 90°C, a más de 95°C, a más de 96°C, a más de 97°C, a más

de 98°C o a más de 99°C. Preferentemente, la cadena de ADN diana puede desnaturalizarse a más de 60°C, a más de 70°C, a más de 80°C, a más de 90°C, a más de 95°C, a más de 96°C, a más de 97°C, a más de 98°C o a más de 99°C. Mediante desnaturalización de la cadena de ADN se separan las dos cadenas de la doble hélice. Este procedimiento también puede denominarse "fusión" del ADN.

5 En una solución que comprende el ADN diana y la sonda correspondiente utilizada en la presente invención, la sonda puede aparearse con la cadena de ADN al enfriar dicha solución a menos de 75°C, a menos de 70°C, a menos de 65°C, a menos de 60°C, a menos de 55°C, a menos de 50°C, a menos de 45°C, a menos de 40°C, a menos de 39°C, a menos de 38°C o a menos de 37°C.

10 La expresión "producto de amplificación" puede interpretarse como el producto de la etapa de amplificación descrita en la presente memoria. Típicamente, el producto de amplificación es más corto que el ADN diana, aunque más largo que la sonda. El locus de interés puede encontrarse incluido en el producto de amplificación.

15 La sonda puede utilizarse además para una reacción en cadena de la polimerasa (PCR) tal como es conocida de la técnica. En la presente memoria, las sondas pueden encontrarse marcadas o no marcadas. Preferentemente ambas sondas se encuentran marcadas. Las dos sondas pueden marcarse con dos fluoróforos diferentes. Alternativamente, una sonda puede marcarse con un fluoróforo y la otra sonda puede marcarse con un inhibidor. Una sonda puede marcarse además con dos fluoróforos diferentes o con un fluoróforo y un inhibidor. La sonda o sondas y/o cebador o cebadores están comprendidos en una solución que comprende además el ADN diana, un tampón (que comprende agua y que opcionalmente comprende un tampón del pH, una sal de magnesio y cofactores de la ADN polimerasa), ADN polimerasa, una mezcla de nucleótidos (que comprende dATP, dGTP, dCTP y dTTP).

20 En primer lugar, se desnaturaliza la hebra de ADN diana. Por ejemplo, según su longitud, una hebra de ADN diana puede desnaturalizarse tal como se ha indicado anteriormente. Mediante desnaturalización de la hebra de ADN se separan las dos cadenas de la doble hélice.

25 A continuación, la solución se enfría a menos de 75°C, a menos de 70°C, a menos de 65°C, a menos de 60°C, a menos de 55°C, a menos de 50°C, a menos de 45°C, a menos de 40°C, a menos de 39°C, a menos de 38°C o a menos de 37°C.

30 En el óptimo de temperatura de la ADN polimerasa utilizada, que preferentemente puede encontrarse comprendida en el intervalo de entre 50°C y 80°C, más preferentemente en el intervalo de entre 50°C y 65°C, se alarga la cadena de ADN. A continuación, el siguiente ciclo puede iniciarse mediante desnaturalización de las cadenas de ADN. La detección puede realizarse en tiempo real.

35 La reacción de PCR puede llevarse a cabo por cualquier medio conocido de la técnica. Puede llevarse a cabo manualmente o automatizarse. La PCR automatizada puede llevarse a cabo en un aparato de PCR estándar o en un aparato de PCR en tiempo real (por ejemplo un LightCycler®). La PCR puede ser PCR cuantitativa (PCRc). La PCR puede ser PCR con transcriptasa inversa (RT-PCR). También pueden combinarse entre sí diferentes métodos de PCR. La reacción de PCR puede combinarse mediante la detección de un efecto de FRET. La detección puede realizarse en tiempo real.

40 En una realización preferente adicional, el método se caracteriza porque:

45 (i) la detección se realiza mediante la medición de la presencia o ausencia de fluorescencia,
(ii) la detección se realiza en tiempo real,
(iii) el marcador se selecciona de entre el grupo que consiste de fluoresceína, LC-amarillo 555, FAM, VIC, HEX, rodamina-B, rodamina-6G, LC-rojo 610, LC-rojo 640, LC-rojo 670, LC-rojo 705, Cy3, Cy3.5, Cy5, Cy5.5 y un inhibidor,
50 (iv) la etapa de amplificación utiliza un enzima polimerasa que presenta actividad de exonucleasa 5' a 3', y/o
(v) la muestra es una muestra biológica, preferentemente una muestra seleccionada de entre el grupo que consiste de un líquido corporal, una muestra de sangre, una muestra de orina, suero, mucosa, esputo, heces, una muestra epidérmica, una muestra de piel, raspado bucal, esperma, líquido amniótico, células en cultivo y una muestra de médula ósea.

55 La sonda puede marcarse con un marcador, en particular un marcador fluorescente, tal como, por ejemplo, pigmentos de fluoresceína (fluoresceína, isotiocianato de fluoresceína (FITC), pigmentos de rodamina (rodamina, rodamina-B, rodamina-6G, tetrametil-rodamina (TAMRA), isotiocianato de rodamina), pigmentos de cianina (por ejemplo Cy3, Cy3.5, Cy5, Cy5.5, Cy7), pigmentos LC (por ejemplo LC-amarillo 555, LC-rojo 610, LC-rojo 640, LC-rojo 670, LC-rojo 705), pigmentos Alexa (Alexa 488, Alexa 546, Alexa 647), S0387, pigmento HOECHST, isotiocianato de eritrosina, verde Oregon, amarillo Lucifer, VS, ficoeritrina, FAM, amarillo 555 LightCycler®, HEX o puntos cuánticos.

El fluoróforo puede excitarse con luz de una longitud de onda próxima a un de sus máximos de absorbancia. Alternativamente, el fluoróforo puede excitarse con luz de una intensidad elevada del doble de longitud de onda de uno de sus máximos de absorbancia (efecto de dos fotones).

- 5 La sonda puede marcarse con uno, dos, tres, cuatro, cinco o más marcadores fluorescentes, preferentemente la sonda se marca con uno o dos marcadores fluorescentes.

10 En el caso de que la sonda se marque con dos marcadores fluorescentes, estos marcadores fluorescentes pueden ser del mismo tipo o marcadores fluorescentes diferentes. Preferentemente, la sonda se marca con dos marcadores fluorescentes, en la que el espectro de emisión del fluoróforo donante se solapa con el espectro de absorbancia del fluoróforo aceptor. Todavía más preferentemente, la sonda se marca con dos fluoróforos que permiten FRET (transferencia de energía por resonancia fluorescente).

15 El fluoróforo donante puede absorber luz a una longitud de onda más corta que el fluoróforo aceptor. Además, el fluoróforo donante preferentemente emite luz a una longitud de onda más corta que el fluoróforo aceptor. Todavía más preferentemente, el espectro de emisión de un fluoróforo donante se solapa en gran medida con el espectro de absorbancia del otro fluoróforo aceptor. Alternativamente, el máximo del espectro de excitación del fluoróforo donante puede encontrarse en el doble de la longitud de onda del máximo de absorbancia del fluoróforo aceptor. A continuación, para el FRET se utiliza un sistema de dos fotones.

20 La tecnología FRET es bien conocida por el experto en la materia (ver, por ejemplo, las patentes US nº 4.996.143, nº 5.565.322, nº 5.849.489 y nº 6.162.603). Se basa en el concepto de que la energía se transfiere de un fluoróforo donante a un fluoróforo aceptor que por su parte emite luz. Tras la irradiación del fluoróforo donante con una longitud de onda determinada que resulta absorbida por el fluoróforo donante, se excita el fluoróforo donante. En el caso de que ningún fluoróforo aceptor se encuentre dentro de su distancia espacial próxima, el fluoróforo donante emitirá luz de una determinada longitud de onda desplazada hacia el rojo (efecto batocrómico). Sin embargo, en el caso de que se encuentre un fluoróforo aceptor dentro de su distancia espacial próxima, se producirá FRET. La energía se transfiere del fluoróforo donante al fluoróforo aceptor, producida por transferencia de energía resonante (transferencia energética de Förster); de esta manera, preferentemente sin emisión de luz. De esta manera, el fluoróforo aceptor resulta excitado y puede emitir luz. La luz emitida por el fluoróforo aceptor se encuentra adicionalmente desplazada hacia el rojo (desplazamiento batocrómico, desplazamiento de Stokes). El desplazamiento de Stokes de los fluoróforos utilizados preferentemente puede ser superior a 20 nm, superior a 30 nm, superior a 40 nm, superior a 50 nm, superior a 75 nm, superior a 100 nm, superior a 125 nm, superior a 150 nm o superior a 200 nm.

35 La producción de un efecto FRET puede determinarse y cuantificarse de diferentes maneras. La reducción de la intensidad lumínica emitida por el fluoróforo donante al producirse FRET puede ser cuantificada. Alternativamente o adicionalmente, puede cuantificarse el incremento de la intensidad lumínica emitida por el fluoróforo aceptor al producirse FRET.

40 La intensidad de FRET depende de la distancia espacial entre los fluoróforos donante y aceptor. La distancia espacial en la sonda preferentemente puede ser inferior a 10, inferior a 9, inferior a 8, inferior a 7, inferior a 6, inferior a 5, inferior a 4, inferior a 3 ó inferior a 2 nucleótidos. Preferentemente, la distancia espacial entre los fluoróforos donante y aceptor es menor al radio de Förster.

45 El fluoróforo donante puede excitarse con cualquier tipo de fuente lumínica. Preferentemente, la fuente lumínica emite luz de una longitud de onda definida. La fuente lumínica puede ser, por ejemplo, un láser iónico de argón, una lámpara de argo de mercurio (Hg) de alta intensidad, un diodo LED, un láser de HeNe o un láser de HeCd. La luz de excitación puede dirigirse por uno o más filtros y/o uno o más espejos dicroicos seleccionando la longitud de onda deseada. De manera similar, la luz de emisión puede dirigirse por uno o más filtros y/o uno o más espejos dicroicos seleccionando la longitud de onda deseada.

50 El experimento de FRET puede llevarse a cabo en cualquier diseño experimental conocido de la técnica para la cuantificación de la intensidad de FRET. El diseño experimental puede comprender, por ejemplo, un microscopio de epifluorescencia de recuento de fotones (que contenga el espejo o espejos dicroicos y filtro o filtros apropiados para el seguimiento de la emisión fluorescente en el intervalo particular), un sistema fotomultiplicador de recuento de fotones o fotómetro.

60 La detección de FRET puede llevarse a cabo mediante la medición de la presencia o ausencia de fluorescencia, Tal como se utiliza en el contexto de la presente invención, la expresión "ausencia de fluorescencia" se refiere a una tasa de fluorescencia inferior al 20%, inferior al 15%, inferior al 10%, inferior al 5%, inferior al 4%, inferior al 3%, inferior al 2%, inferior al 1% o inferior al 0,5% de la intensidad de fluorescencia observada como máximo para el

mismo fluoróforo bajo las mismas condiciones (tal como en el mismo tampón, a la misma temperatura, con la misma intensidad de excitación en el mismo aparato con los mismos parámetros).

5 Para una sonda intacta, puede ser detectable un efecto de FRET fuerte. Para una sonda cortada o una sonda de la que se han escindido uno o dos fluoróforos, puede ser detectable un efecto de FRET nulo o mucho menor. Por lo tanto, una sonda de FRET puede ser una sonda de hidrólisis, tal como, por ejemplo, una sonda TaqMan®, tal como se indica en la presente memoria.

10 Alternativamente, la presencia de dos fluoróforos en una estructura molecular también puede determinarse mediante espectroscopía de correlación cruzada de fluorescencia (FCCS). En la presente memoria puede determinarse si las señales de fluorescencia producidas por dos fluoróforos diferentes de una sonda doblemente marcada se difunden conjuntamente o independientemente. Alternativamente, puede determinarse si la fluorescencia producida por dos fluoróforos diferentes de dos sondas marcadas diferentemente de unión al mismo ADN diana se difunden conjuntamente o independientemente.

15 Además, puede utilizarse un ensayo homogéneo de proximidad de luminiscencia amplificada (ALPHA, AlphaScreen).

20 Las sondas también pueden marcarse con un solo fluoróforo. A continuación, la unión a la secuencia diana puede detectarse mediante la medición de la despolarización de la fluorescencia o la velocidad de difusión. La expresión "velocidad de difusión" puede incluir la velocidad de difusión lateral y tangencial, la velocidad rotacional de la molécula, la velocidad rotacional intramolecular, cualquier tipo de oscilación molecular y las combinaciones de los mismos.

25 La despolarización de la fluorescencia se basa en la anisotropía de la fluorescencia. En los ensayos de despolarización de la fluorescencia, la difusión rotacional de una molécula se determina a partir de la descorrelación de la polarización en la fluorescencia, es decir, entre los fotones excitatorios y emitidos (fluorescentes). Esta descorrelación puede medirse como el "tiempo de tambaleo" de la molécula globalmente, o de una parte de la molécula respecto al global. A partir de las constantes de difusión rotacional, el experimentador puede determinar si una sonda de bajo peso molecular de rápido tambaleo se ha unido a su secuencia diana de alto peso molecular de tambaleo lento.

35 La velocidad de difusión puede medirse mediante espectroscopía de correlación de fluorescencia (ECF), tal como es conocida por el experto en la materia. En la presente memoria, se mide el tiempo de permanencia de las moléculas en difusión libre. A partir del tiempo de permanencia medio (constante de difusión), el experimentador podrá determinar si una sonda de bajo peso molecular de rápido tambaleo se ha unido a su secuencia diana de alto peso molecular de tambaleo lento o si la sonda de bajo peso molecular se difunde libremente.

40 La sonda también puede marcarse con uno o más fluoróforos y uno o más inhibidores. Preferentemente, la sonda puede marcarse con un fluoróforo y un inhibidor.

En una realización todavía más preferente de la presente invención, las sondas son sondas de hidrólisis marcadas adicionalmente con un inhibidor.

45 Un inhibidor tal como se utiliza en la presente memoria es una estructura molecular que puede inhibir la luz emitida por un fluoróforo. Por lo tanto, un inhibidor que se encuentra a una distancia especial comparativamente reducida a un fluoróforo puede reducir la intensidad de la luz emitida por el fluoróforo al ser excitado.

50 El inhibidor puede inhibir la luz a una distancia espacial inferior a 10, inferior a 9, inferior a 8, inferior a 7, inferior a 6, inferior a 5, inferior a 4, inferior a 3 ó inferior a 2 nucleótidos entre el fluoróforo y el inhibidor.

La sonda también puede ser una sonda TaqMan®, una baliza molecular, un cebador Scorpion o un cebador lux, tal como son conocidos de la técnica. Preferentemente, la sonda es una sonda TaqMan®.

55 Alternativamente, las sondas pueden marcarse con un pigmento no fluorescente, tal como, por ejemplo, una fracción p-nitrofenilo o verde malaquita, o con moléculas pequeñas reactivas que pueden unirse a otras moléculas, tales como, por ejemplo, maleimidias, isotiocianatos o ésteres activos (por ejemplo ésteres de succinimidilo o ésteres de p-nitrofenilo). Además, las sondas pueden marcarse con moléculas pequeñas de unión selectiva a moléculas de alto peso molecular tales como, por ejemplo, metotrexato de biotina o glucocorticoides. Puede detectarse una sonda marcada con biotina mediante la utilización de estreptavidina marcada. Una sonda marcada con metotrexato puede detectarse mediante la utilización de dihidrofolato reductasa (DHFR) marcada. Una sonda marcada con un glucocorticoide puede detectarse mediante la utilización de un anticuerpo marcada o derivados de anticuerpo (por ejemplo fragmentos Fab, anticuerpos de cadena sencilla, diacuerpos, triacuerpos y tandabs) dirigidos contra dichos

60

glucocorticoides. Alternativamente, la sonda puede des-marcarse y detectarse con un anticuerpo marcado o con un derivado de anticuerpo marcado (por ejemplo un fragmento Fab, un anticuerpo de cadena sencilla, un diacuerpo, un triacuerpo o un tandab). Alternativamente, la sonda puede des-marcarse y detectarse con un anticuerpo no marcado o con un derivado de anticuerpo no marcado (por ejemplo un fragmento Fab, un anticuerpo de cadena sencilla, un diacuerpo, un triacuerpo o un tandab), que se detecte, por su parte, con un anticuerpo o derivado de anticuerpo marcada dirigida contra su parte Fc.

Además, la sonda puede conjugarse con un enzima que puede generar un color a partir de un precursor (por ejemplo una peroxidasa o fosfatasa alcalina). Dicha conjugación puede obtenerse mediante conjugación de la sonda covalentemente con el enzima o mediante conjugación de la sonda con una molécula de unión (por ejemplo digoxigenina o metotrexato) que pueda unirse a una proteína de fusión del enzima que pueda generar un color a partir de un precursor y una proteína de unión a la molécula de unión. Alternativamente, el enzima puede ser capaz de emitir luz mediante conversión química (quimioluminiscencia) (por ejemplo luciferasa).

Alternativamente, las sondas pueden marcarse radioactivamente. Por lo tanto, la sonda puede marcarse con, por ejemplo, ^3H , ^{32}P , ^{35}S , ^{15}C , lantánidos u otros marcajes radioactivos.

Alternativamente, las sondas pueden marcarse con átomos pesados detectables mediante resonancia magnética nuclear (RMN) o espectrometría de masas (por ejemplo EM-ESI o EM-MALDI). Por lo tanto, la sonda puede marcarse con, por ejemplo, ^2H o ^{13}C .

Además, el punto de fusión de la sonda o sondas puede encontrarse determinado. El experto en la materia podrá entender que un punto de fusión más alto se refiere a una unión más fuerte. De esta manera, una sonda de unión más fuerte a su secuencia diana que una sonda correspondiente que porta una o más posiciones nucleotídicas presenta menos desapareamientos con la secuencia diana que la sonda correspondiente. De esta manera, puede caracterizarse el genotipo de una o más muestras.

Preferentemente, una de las sondas de uno de los dos o más juegos de sondas muestra una correspondencia perfecta con la secuencia diana; las otras sondas del mismo juego de sondas muestran uno o más desapareamientos con la secuencia diana. Las sondas del juego o juegos correspondientes de sondas pueden mostrar un desapareamiento adicional en la posición discriminante.

Las sondas pueden ser sondas TaqMan[®]. Las sondas TaqMan[®] pueden utilizarse para llevar a cabo un ensayo TaqMan[®]. El ensayo TaqMan[®] es bien conocido por el experto en la materia. Tal como se utiliza en la presente memoria, las expresiones "sonda TaqMan[®]" y "sonda de hidrólisis" pueden entenderse intercambiamente. La sonda TaqMan[®] comprende un fluoróforo y un inhibidor. Preferentemente, el fluoróforo y el inhibidor se encuentran situados en una posición próxima a los extremos de la sonda, más preferentemente, el fluoróforo se encuentra situado en una posición próxima al extremo 5' y el inhibidor se encuentra situado en una posición próxima al extremo 3'. La expresión "3'-terminal" puede ser entendida en el sentido más amplio por el experto en la materia. Además, las expresiones "extremo 3'-terminal" y "extremo 3'" pueden entenderse intercambiamente tal como es conocido de la técnica. El experto en la materia observará que las expresiones "extremo 3'-terminal" y "extremo 3'" tal como se utiliza en la presente memoria puede referirse al extremo 5' de la cadena de nucleótidos, pero puede no excluir que en el extremo 3' se añada otra fracción molecular (tal como, por ejemplo, un fluoróforo, un inhibidor, una fracción de unión o similar) al extremo 3' de la sonda.

Tal como se utiliza en la presente memoria, la expresión "situado en una posición próxima" preferentemente se refiere a que la fracción fluorescente o inhibidora se encuentra situada a no más de 4, más preferentemente a no más de 3, todavía más preferentemente a no más de 2, y todavía más preferentemente en el primer nucleótido respecto al extremo respectivo.

La sonda TaqMan[®] puede hibridarse con su secuencia diana. Además, la composición utilizada puede comprender además una pareja de cebadores, de esta manera un cebador directo y un cebador inverso. Estos cebadores preferentemente no se encuentran marcados. Preferentemente, el cebador directo se une a una cadena arriba, mientras que el cebador inverso se une a una cadena abajo, de manera que la sonda TaqMan[®] se une a una secuencia que es una parte de la cadena que se amplifica. Se lleva a cabo una reacción de PCR tal como es bien conocido de la técnica. De esta manera, se funde el ADN diana y después se seleccionan condiciones que permitan la hibridación de los cebadores y la sonda con el ADN diana. A continuación, se seleccionan condiciones que permitan que la ADN polimerasa amplifique la cadena de ADN entre los cebadores. En el contexto del ensayo TaqMan[®], la ADN polimerasa preferentemente presenta una actividad de exonucleasa 5' a 3'. Más preferentemente, la ADN polimerasa es una polimerasa Taq o una variante funcional de la misma. Al llegar la ADN polimerasa a la sonda TaqMan[®], se escinde el extremo 5'. De esta manera, el fluoróforo o el inhibidor unido al nucleótido o nucleótidos 5'-terminales también son escindidos. Preferentemente se escinde el fluoróforo. En consecuencia, el fluoróforo y el inhibidor pueden difundirse en direcciones diferentes. La distancia espacial entre ambos puede incrementarse

significativamente y la fluorescencia producida por el fluoróforo se incrementa significativamente debido a que ya no resulta inhibida por el inhibidor oscuro. Preferentemente, el ensayo TaqMan[®] se analiza en tiempo real. Más preferentemente, el ensayo TaqMan[®] se lleva a cabo en un método de PCR en tiempo real. También puede llevarse a cabo cuantitativamente en una reacción de PCRc.

5 El ensayo TaqMan[®] utilizando las sondas de la presente invención puede utilizarse preferentemente para la discriminación de alelos, genotipado, ensayos de identificación bacteriana, cuantificación del ADN y la determinación de la carga vírica en especímenes clínicos, ensayos de expresión génica y verificación de resultados de micromatrices. Más preferentemente puede utilizarse para la discriminación de alelos, genotipado y ensayos de
10 identificación bacteriana. El genotipado preferentemente puede ser de polimorfismos de nucleótidos simples (PNS), por lo tanto la determinación de un genotipo en un locus de interés definido en una muestra, en el que el locus es un único nucleótido. Alternativamente, el genotipado puede ser genotipado de una variante de número de copia (VNC). Una variante de número de copia (VNC) es un segmento de ADN en el que se han encontrado diferencias de número de copia (número de copias de una secuencia de ADN o partes de la misma) mediante la comparación entre
15 dos o más genomas. Las secuencias y loci de los PNS y VNC conocidos de la técnica en la actualidad pueden obtenerse de bases de datos conocidas por el experto en la materia. Estas bases de datos son, por ejemplo, la base de datos de variantes genómicas (DGV), la base de datos NCBI dbSNP, el sitio de bioinformática genómica de UCSC, la base de datos Databases of Chromosomal Imbalance and Phenotype in Humans using Ensembl Resources (DECIPHER) [base de datos de desequilibrios cromosómicos y fenotipo en seres humanos utilizando recursos compartidos], el proyecto HapMap, el proyecto de variación de número de copia del instituto Sanger y el proyecto Human Structural Variation Project [proyecto de variación estructural humana].

También puede diseñarse un ensayo similar utilizando dos moléculas que muestran un efecto FRET en lugar de la inhibición de la fluorescencia. En éste, se escinde uno de los fluoróforos con una actividad de nucleasa 5' a 3' de la
25 ADN polimerasa. Preferentemente, la ADN polimerasa porta una actividad de exonucleasa 5' a 3', tal como la polimerasa Taq. Se incrementa la distancia espacial entre los fluoróforos. La eficiencia de FRET se reduce tras la escisión del fluoróforo situado en el extremo 5'. El fluoróforo situado en el extremo 5' puede ser el fluoróforo donante o el fluoróforo aceptor. Preferentemente, el fluoróforo situado en el extremo 5' puede ser el fluoróforo donante. Preferentemente, este ensayo se lleva a cabo en un método de PCR en tiempo real. También puede llevarse a cabo
30 cuantitativamente en una PCRc.

La determinación del genotipo puede llevarse a cabo mediante un método basado en la hibridación y/o la PCR, tal como se describe en la presente memoria. Preferentemente, la determinación del genotipo puede llevarse a cabo en un método basado en la PCR. Más preferentemente, la determinación del genotipo puede llevarse a cabo en un
35 método basado en la PCR y en TaqMan[®].

Puede llevarse a cabo una etapa de hibridación con dos o más sondas. Tal como se utiliza en la presente memoria, cada una de las sondas preferentemente se marcan con uno o más fluoróforos, dos fluoróforos diferentes o un fluoróforo y un inhibidor. Además, se detecta la señal producida por las sondas. Preferentemente, con la señal FRET se detecta el incremento de la fluorescencia. La sonda puede ser una sonda TaqMan[®] que comprende un fluoróforo y un inhibidor. La detección puede basarse en la presencia o ausencia de fluorescencia o en un incremento o
40 reducción de la fluorescencia.

Las sondas de la presente invención pueden utilizarse además para la hibridación *in situ*, en particular en la hibridación *in situ* comparativa y/o en la hibridación *in situ* fluorescente (FISH). En la presente memoria, la sonda puede marcarse fluorescente o radioactivamente o puede detectarse con un anticuerpo o derivado de anticuerpo que se encuentre marcado o que puede detectarse con un segundo anticuerpo dirigido contra su parte Fc. Además, la sonda puede conjugarse con un enzima que puede generar un color a partir de un precursor (por ejemplo una peroxidasa o fosfatasa alcalina).
45

Además, las sondas de la presente invención pueden utilizarse en otros métodos basados en la hibridación de sondas. Las sondas de la presente invención pueden utilizarse en métodos de micromatrices. Preferentemente, las sondas pueden utilizarse en la validación de micromatrices. Además, las sondas pueden utilizarse en ensayos de cuantificación mediante inactivación génica.
50

Además, el método de la presente invención puede utilizarse para PCR específica de alelo, en la que una de las sondas puede servir como cebador que se combina con un cebador adecuado o en la que dos sondas sirven como cebadores directo e inverso. Preferentemente una de las sondas puede servir como cebador que se une al locus en el que se encuentran situados los diferentes alelos o se supone que se encuentran situados.
55

Además, la sonda de la presente invención puede utilizarse para el perfilado de expresión de micromatrices, la investigación de ARN pequeños, la reparación génica/salto de exón, la detección de variantes de procesamiento,
60

ADNzimas y/o la hibridación genómica comparativa (HGC). Estos métodos son bien conocidos por el experto en la materia.

5 También puede utilizarse una sonda de la presente invención en un método para regular negativamente la expresión de un gen. Por lo tanto, la sonda puede servir como oligonucleótido antisentido, tal como ARNi (por ejemplo ARNip, ARNm_i, ARNnp) tal como es conocido de la técnica. El oligonucleótido antisentido puede interferir con el ARNm del organismo diana. El organismo diana preferentemente puede ser un animal, incluyendo un ser humano o una planta. Los oligonucleótidos antisentido son bien conocidos por el experto en la materia.

10 La composición de la presente invención puede combinarse con software informático para el análisis de los datos.

15 En un aspecto adicional, la presente invención se refiere a una biblioteca de por lo menos dos juegos de sondas, en la que la biblioteca comprende una pluralidad de juegos de sondas, presentando cada una de las sondas ocho nucleótidos con la estructura general 5'-D -L L -L -L -L -X -X-3' or 5'-D -L -L -L -L -X -X -X-3',
 20 en la que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es un nucleótido aleatorio de ANB, en el que dentro de un juego de sondas, todas las sondas presentan secuencias de nucleótidos idénticas con la excepción de los dos y/o tres nucleótidos aleatorios de ANB, en el que en cada posición de un nucleótido aleatorio de ANB, la base se selecciona independientemente de entre adenina, citosina, guanina y timina, y cualquier posible secuencia resultante de la variación o variaciones de secuencia en las dos posiciones se encuentra representada por una sonda en cada juego de sondas, y
 25 en el que un juego de sondas difiere de otro juego de sondas en la secuencia de por lo menos el nucleótido de ADN D o un nucleótido de ANB L.

30 Tal como se utiliza en la presente memoria, el término "biblioteca" se refiere a una colección de varios juegos de sondas, en la que los nucleótidos de ADN y los nucleótidos de ANB pueden encontrarse situados en la misma posición en cada una de las sondas. Además, en una biblioteca de juegos de sondas, las posiciones aleatorias y las posiciones determinadas, y la posición discriminante, pueden encontrarse situadas en la misma posición en cada sonda de la biblioteca.

35 Dicha biblioteca puede diseñarse para cubrir un gran número de dianas en el genoma de una especie o en el genoma de diferentes especies. Idealmente, dicha biblioteca de sondas puede ser una biblioteca de sondas para cualquier problema técnico.

40 Para juegos de sondas en los que tres nucleótidos son nucleótidos aleatorios y cinco nucleótidos son determinados, resultan suficientes 1.024 juegos de sondas para proporcionar una biblioteca de sondas adecuada para cualquier problema técnico. Debido a que se requieren dos juegos de sondas de los que se conoce por lo menos la posición discriminante de cada nucleótido, resulta suficiente una biblioteca que comprende 512 juegos de sondas para la aplicación a cualquier problema técnico.

45 Para juegos de sondas en los que dos nucleótidos son nucleótidos aleatorios y seis nucleótidos son determinados, resultan suficientes 4.096 juegos de sondas para proporcionar una biblioteca de sondas adecuada para cualquier problema técnico. Debido a que se requieren dos juegos de sondas de los que se conoce por lo menos la posición discriminante de cada nucleótido, resulta suficiente una biblioteca que comprende 2048 juegos de sondas para la aplicación a cualquier problema técnico.

50 En una realización preferente, la biblioteca de juegos de sondas se caracteriza porque:
 (i) el número de juegos es de por lo menos 64, preferentemente de 128, especialmente de por lo menos 256, particularmente de por lo menos 512, más preferentemente de por lo menos 1.024, todavía más preferentemente de 2.048, y/o
 (ii) los juegos de sondas se encuentran separados espacialmente entre sí.

55 La expresión "separados espacialmente entre sí" se refiere a que las sondas pueden almacenarse separados uno de otro; de esta manera, por ejemplo, pueden almacenarse en diferentes viales o en diferentes pocillos. Los viales pueden ser viales de plástico o viales de vidrio, tales como, por ejemplo, un vaso de plástico, un vial con tapa enroscable o un vial sellado. Un pocillo puede referirse a, por ejemplo, un pocillo de una placa multipocillo (placa multicámara, placa de multititulación), tal como, por ejemplo, una placa de 8 pocillos, una placa de 12 pocillos, una placa de 24 pocillos, una placa de 96 pocillos, una placa de 384 pocillos o una placa de 1.536 pocillos, tal como es conocido de la técnica. Pueden almacenarse bajo condiciones adecuadas para sondas de ADN y/o ANB de cadena sencilla de la longitud dada. Pueden almacenarse a temperatura ambiente, a 4°C, a -20°C, a -80°C o en nitrógeno líquido. Pueden almacenarse en agua, en un tampón acuoso, o en un solvente orgánico, tal como DMSO. Además,
 60 pueden liofilizarse.

Una biblioteca de dichas sondas puede comprender por lo menos 2, por lo menos 64, preferentemente 128, por lo menos 256, por lo menos 512, por lo menos 1.024, por lo menos 2.048, por lo menos 4.096 ó más sondas. La biblioteca de sondas puede comprender grandes partes del genoma del organismo o incluso el genoma completo. La biblioteca de sondas puede comprender uno, dos, tres, cuatro, cinco o más alelos diferentes para un locus particular de interés. Puede encontrarse presente cero, uno, dos, tres o más loci de interés en un único gen.

La invención se explica en mayor detalle mediante los ejemplos y figuras siguientes, los cuales pretenden ilustrar, aunque no limitar, el alcance de la presente invención.

10 Figuras

La fig. 1 muestra un esquema de un ensayo de PCR de color dual con dos sondas de hidrólisis. La primera sonda (con informador 1 denominado R1) se hibrida y se corta en presencia de una correspondencia perfecta (con la secuencia de tipo salvaje) en la muestra (parte superior), permitiendo que R1 produzca una señal; la segunda sonda (con informador 2, denominado R2) se hibrida y se corta en presencia de una correspondencia perfecta (con la secuencia mutante) en la muestra (parte inferior), permitiendo que R2 produzca una señal (Q=inhibidor).

La fig. 2 muestra un ensayo de PCR mono-color con parámetro 18S. Las curvas (triángulos) (parte superior) muestran un resultado Cp positivo y una curva de amplificación sigmoideal con correspondencia completa de la sonda a diferentes concentraciones de muestra. Las curvas (cruces) (parte inferior) muestran un resultado Cp negativo y una curva de amplificación no sigmoideal con desapareamientos de la sonda a diferentes concentraciones de muestra.

La fig. 3 muestra un ensayo de PCR mono-color con parámetro MNAT1. Las curvas (triángulos) (parte superior) muestran un resultado Cp positivo y una curva de amplificación sigmoideal con correspondencia completa de la sonda a diferentes concentraciones de muestra. Las curvas (cruces) (parte inferior) muestran un resultado Cp negativo y una curva de amplificación no sigmoideal con desapareamientos de la sonda a diferentes concentraciones de muestra.

La fig. 4 muestra un ensayo de PCR de color dual con parámetro 18S. Las curvas (círculos) (parte superior) muestran un resultado Cp positivo y una curva de amplificación sigmoideal con correspondencia completa de la sonda a diferentes concentraciones de muestra. Las curvas (estrellas) (parte inferior) muestran un resultado Cp negativo y una curva de amplificación no sigmoideal con desapareamientos de la sonda a diferentes concentraciones de muestra.

35 Ejemplos

Ejemplo 1:

El experimento se llevó a cabo con el fin de demostrar el poder discriminante de la sonda de correspondencia completa en comparación con la sonda de desapareamiento, en la que las sondas de correspondencia completa y la de desapareamiento presentan la misma secuencia de ácidos nucleicos excepto por una base de ácidos nucleicos en la parte central de la secuencia de la sonda. Las parejas de cebadores son iguales para ambas sondas, proporcionando el mismo amplificación durante la amplificación por PCR. Debido a que ambas sondas presentan el mismo pigmento informador, el experimento de PCR se llevó a cabo en diferentes pocillos en modo mono-color únicamente. Para demostrar que el rendimiento de la PCR resulta suficiente incluso con diferentes concentraciones de muestra, se llevó a cabo la PCR con dos diluciones diferentes de ADNc como diana (cada concentración por duplicado). El parámetro diana en el Ejemplo 1 era 18S.

Tal como se esperaba, la PCR con una sonda que presentaba una secuencia de correspondencia completa con la secuencia diana mostraba una curva de amplificación sigmoideal y, por lo tanto, un resultado de Cp positivo, mientras que la sonda que presentaba un desapareamiento con la secuencia diana mostraba una curva de amplificación no sigmoideal y un resultado de Cp negativo.

El experimento se llevó a cabo tal como se ilustra en la fig. 1. Ensayo de PCR mono-color con parámetro 18S, estructura general de la sonda: 5'-D -L -L -L -L -L -X -X-3'

secuencia de la sonda FAM de correspondencia completa: 5'-X-t-T***A***T*T*G*-(AGCT)*(AGCT)*-Z-3' secuencia de sonda FAM de desapareamiento: 5'-X-t-T***C***T*T*G*-(AGCT)*(AGCT)*-Z-3', mientras que X = FAM, t = dT T*, A*, C* o G* = LNA T, A, C o G

(AGCT)* = bases tambaleantes de ANB T, A, C y G (posición aleatoria)

Z = BHQ2-inhibidor (inhibidor Black Hole 2)

Las fuentes en negrita y cursiva indican una posición discriminante.

Secuencia del cebador directo: gacggaccagagcgaaag (SEC ID n° 1)

Secuencia del cebador inverso: cgtcttgcgaacctccgact (SEC ID n° 2)

ciclador de PCR: aparato de LightCycler® 480, bloque de 384 pocillos (Roche Applied Science, n° de cat.

04 545 885 001)

reactivos de PCR: sondas maestras LightCycler® 480 (Roche Applied Science, nº de cat.04 707 494 001);
placa multipocillo LightCycler® 480, 384 (Roche Applied Science, nº de cat. 04 729 749 001)

- 5 material de muestra: ADNc, transcripción inversa a partir de ARN (Takara/Clontech, nº de cat.636538) con kit de síntesis de ADNc transcriptor de primera cadena (Roche Applied Science, nº de cat.04 897 030 001). La etapa de síntesis de ADNc se llevó a cabo tal como se describe en el impreso en el paquete del kit correspondiente. Concentración de las muestras: 50 ng y 5 ng en cada pocillo en 10 ml de volumen final de PCR Protocolo de PCR: ver la Tabla, posteriormente.

10 **FAMKanal: 465:510**

	°C	Tiempo	RR(°C/s)(96)	AqMode	Ciclos
Preincubación	96	10min	4,8(4,4)	-	-
Amplificación desnat.	95	10s	4,8(4,4)	ninguno	30
Hibridación alarg.	37	30s	2,5(2,2)	ninguno	-
	50	30s	4,8(4,4)	único	-
Codificación	40	30s	2,2(2,5)	ninguno	-

Ejemplo 2

- 15 El experimento se llevó a cabo con el fin de demostrar el poder discriminante de la sonda de correspondencia completa en comparación con la sonda de desapareamiento, en la que las sondas de correspondencia completa y la de desapareamiento presentan la misma secuencia de ácidos nucleicos excepto por una base de ácidos nucleicos en la parte central de la secuencia de la sonda. Las parejas de cebadores son iguales para ambas sondas, proporcionando el mismo amplicón durante la amplificación por PCR. Debido a que ambas sondas presentan el mismo pigmento informador, el experimento de PCR se llevó a cabo en diferentes pocillos en modo mono-color únicamente. Para demostrar que el rendimiento de la PCR resulta suficiente incluso con diferentes concentraciones de muestra, se llevó a cabo la PCR con dos diluciones diferentes de ADNc como diana (cada concentración por duplicado). La diferencia del Ejemplo 2 en comparación con el Ejemplo 1 es un parámetro diana separado (el parámetro en el Ejemplo 2 es MNAT1).

- 25 Tal como se esperaba, la PCR con una sonda que presentaba una secuencia de correspondencia completa con la secuencia diana mostraba una curva de amplificación sigmoidal y, por lo tanto, un resultado de Cp positivo, mientras que la sonda que presentaba un desapareamiento con la secuencia diana mostraba una curva de amplificación no sigmoidal y un resultado de Cp negativo.

Ensayo de PCR mono-color con parámetro MNAT1 Estructura general de la sonda: 5'-D -L -L -L -L -L -X -X-3'

30 secuencia de la sonda FAM de correspondencia completa: 5'-X-t-T***G***A*T*G*-(AGCT)*(AGCT)*-Z-3'

secuencia FAM de sonda de desapareamiento: 5'-X-t-T***A***A*T*G*-(AGCT)*(AGCT)*-Z-3'

mientras que X = FAM, t = dT

T*,A*,C* o G*=LNA T, A, C o G

(AGCT)* = bases tambaleantes de LNA T, A, C y G (posición aleatoria)

35 Z = inhibidor BHQ2 (inhibidor Black Hole 2)

Las fuentes en negrita y en cursiva indican posición discriminante.

Secuencia del cebador directo: cccaacacctgtaaaaccagtg (SEC ID nº 3)

Secuencia del cebador inverso: ttgtgaatagtgccagttaa (SEC ID nº 4)

- 40 ciclador de PCR: aparato de LightCycler® 480, bloque de 384 pocillos (Roche Applied Science, nº de cat. nº 04 545 885 001)

reactivos de PCR: sondas maestras LightCycler® 480 (Roche Applied Science, nº de cat. nº 04 707 494 001);

placa multipocillo LightCycler® 480, 384 (Roche Applied Science, nº de cat. nº 04 729 749 001)

material de muestra: ADNc, transcripción inversa a partir de ARN (Takara/Clontech, nº de cat. nº 636538) con kit de síntesis de ADNc transcriptor de primera cadena (Roche Applied Science, nº de cat. nº 04 897 030 001). La etapa de

- 45 síntesis de ADNc se llevó a cabo tal como se describe en el impreso en el paquete del kit correspondiente. Concentración de las muestras: 50 ng y 5 ng en cada pocillo en 10 ml de volumen final de PCR Protocolo de PCR: ver la Tabla, posteriormente.

50 **FAMKanal:465:510**

	°C	Tiempo	RR(°C/s)(96)	AqMode	Ciclos
Preincubación	95	10min	4,8(4,4)	-	-
Amplificación desnat.	95	10s	4,8(4,4)	ninguno	30
Hibridación alarg.	37	30s	2,5(2,2)	ninguno	-
	50	30s	4,8(4,4)	Single	-
Codificación	40	30s	2,2(2,5)	ninguno	-

Ejemplo 3:

5 El experimento se llevó a cabo con el fin de demostrar el poder discriminante de la sonda de correspondencia completa en comparación con la sonda de desapareamiento, en la que las sondas de correspondencia completa y la de desapareamiento presentan la misma secuencia de ácidos nucleicos excepto por una base de ácidos nucleicos en la parte central de la secuencia de la sonda. Las parejas de cebadores son iguales para ambas sondas, proporcionando el mismo amplicón durante la amplificación por PCR. Debido a que en el presente ejemplo ambas sondas presentan pigmentos informadores diferentes, el experimento de PCR se llevó a cabo en diferentes pocillos en modo de color dual. Para demostrar que el rendimiento de la PCR resulta suficiente incluso con diferentes concentraciones de muestra, se llevó a cabo la PCR con dos diluciones diferentes de ADNc como diana (cada concentración por duplicado). La diferencia del Ejemplo 3 con el Ejemplo 1 es el modo de PCR mono-color a color dual (el parámetro en el Ejemplo 3 es 18S).

15 Tal como se esperaba, la PCR con una sonda que presentaba una secuencia de correspondencia completa con la secuencia diana mostraba una curva de amplificación sigmoideal en el canal de fluorescencia correspondiente y, por lo tanto, un resultado de Cp positivo, mientras que la sonda que presentaba un desapareamiento con la secuencia diana mostraba una curva de amplificación no sigmoideal en el canal de fluorescencia correspondiente y un resultado de Cp negativo.

20 Ensayo de PCR de color dual con parámetro 18S
 Estructuras generales de sonda: 5'-D -L -L -L -L -L -X -X-3'
 secuencia de la sonda LC amarillo 555 de correspondencia completa: 5'-Y-t-T*C*T*T*G*-(AGCT)*(AGCT)*-Z-3'
 secuencia de la sonda FAM de desapareamiento: 5'-X-t-T*A*T*T*G*-(AGCT)*(AGCT)*-Z-3'
 25 mientras que X = FAM, Y = LC amarillo 555 (LightCycler® amarillo 555), t = dT
 T*, A*, C* o G* = LNA T, A, C o G
 (AGCT)* = bases tambaleantes de LNA T, A, C y G (posición aleatoria)
 Z = inhibidor BHQ2 (inhibidor Black Hole 2)
 Las fuentes en negrita y en cursiva indican posición discriminante.
 30 Secuencia del cebador directo: gacggaccagagcgaag (SEC ID nº 1)
 Secuencia del cebador inverso: cgtcttgaacctccgact (SEC ID nº 2)
 Ciclador de PCR: aparato de LightCycler® 480, bloque de 384 pocillos (Roche Applied Science, nº de cat. nº 04 545 885 001)
 Reactivos de PCR: sondas maestras LightCycler® 480 (Roche Applied Science, nº de cat. nº 04 707 494 001);
 35 placa multipocillo LightCycler® 480, 384 (Roche Applied Science, nº de cat. nº 04 729 749 001)
 Material de las muestras: ADNc, transcripción inversa a partir de ARN (Takara/Clontech, nº de cat. nº 636538) con kit de síntesis de ADNc transcriptor de primera cadena (Roche Applied Science, nº de cat. nº 04 897 030 001). La etapa de síntesis de ADNc se llevó a cabo tal como se describe en el impreso en el paquete del kit correspondiente.
 Concentración de las muestras: 50 ng y 5 ng en cada pocillo en 10 ml de volumen final de PCR
 40 Protocolo de PCR: ver la Tabla, posteriormente.

FAMKanal: 465:510

	°C	Tiempo	RR(°C/s)(96)	Modo Aq	Ciclos
Preincubación	95	10min	4,8(4,4)	-	-
Amplificación desnat.	95	10s	4,8(4,4)	ninguno	30
Hibridación alarg.	37	30s	2,5(2,2)	ninguno	-
	50	30s	4,8(4,4)	único	-
Codificación	40	30s	2,2(2,5)	ninguno	-

45 LISTADO DE SECUENCIAS

<110> Roche Diagnostics GmbH
 F. Hoffmann-La Roche AG

50 <120> Nuevo tipo de sondas universales para la detección de variantes genómicas

<130> R67345EP

<160> 4

55 <170> PatentIn versión 3.5

<210> 1
<211> 18
<212> ADN
<213> Secuencia artificial
5
<220>
<223> Secuencia de cebador directo

<400> 1
10 gacggaccag agcgaaag 18

<210> 2
<211> 19
<212> ADN
15 <213> Secuencia artificial

<220>
<223> Secuencia de cebador inverso

20 <400> 2
cgcttcgaa cctccgact 19

<210> 3
<211> 21
25 <212> ADN
<213> Secuencia artificial

<220>
<223> Secuencia de cebador directo
30
<400> 3
cccaaacctg taaaaccagt g 21

<210> 4
<211> 21
35 <212> ADN
<213> Secuencia artificial

<220>
40 <223> Secuencia de cebador inverso

<400> 4
tttgaatag gtgccagtga a 21
45

REIVINDICACIONES

1. Composición que comprende un primer juego de sondas y un segundo juego de sondas, presentando cada una de las sondas ocho nucleótidos, en el que los ocho nucleótidos están compuestos de uno a tres nucleótidos de ADN y cinco a siete nucleótidos de ANB (ácido nucleico bloqueado), en el que todas las sondas del primer y segundo juegos de sondas presentan secuencias de nucleótidos idénticas con la excepción de:
- (i) la base o bases en una, dos o tres posiciones aleatorias de ANB, y
- (ii) la base en una posición discriminante,
- en el que una, dos o tres posiciones aleatorias de ADN y la posición discriminante se encuentran situadas en posiciones idénticas en todas las sondas del primer y segundo juegos, en el que en cada posición aleatoria de ANB la base se selecciona independientemente de entre adenina, citosina, guanina y timina y cualquier posible secuencia resultante de la variación o variaciones de secuencia en una, dos o tres posiciones aleatorias de ANB se encuentra representada por como mínimo una sonda en cada juego de sondas, y en el que la base en la posición discriminante es idéntica dentro de cada juego de sondas pero difiere entre el primer y el segundo juegos de sondas.
2. Composición según la reivindicación 1, en la que la posición discriminante se encuentra en la posición 2, 3, 4, 5, 6 ó 7 desde el extremo 5' en cada sonda, preferentemente en la posición 3, 4 ó 5, más preferentemente en la posición 4.
3. Composición según la reivindicación 1 ó 2, en la que el nucleótido en la posición 1 desde el extremo 5' es un nucleótido de ADN.
4. Composición según cualquiera de las reivindicaciones 1 a 3, en la que cada sonda consiste de:
- un nucleótido de ADN y siete nucleótidos de ANB,
 - dos nucleótidos de ADN y seis nucleótidos de ANB, o
 - tres nucleótidos de ADN y cinco nucleótidos de ANB.
5. Composición según cualquiera de las reivindicaciones 1 a 4,
- a) en la que el juego de sondas presenta una posición aleatoria de ANB en la posición 5, 6, 7 ó 8 desde el extremo 5', o
- b) en la que el juego de sondas presenta dos posiciones aleatorias de ANB en las posiciones 5 y 6, 5 y 7, 5 y 8, 6 y 7, 6 y 8, ó 7 y 8 desde el extremo 5', preferentemente en las posiciones 7 y 8, o
- c) en el que el juego de sondas presenta tres posiciones aleatorias de ANB en las posiciones 5, 6 y 7, 5, 6 y 8, ó 6, 7 y 8 desde el extremo 5', preferentemente en las posiciones 6, 7 y 8.
6. Composición según cualquiera de las reivindicaciones 1 a 5, en la que las sondas presentan la estructura general 5'-D -L -L -L -L -X -X -X-3', en la que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es una posición aleatoria de ANB.
7. Composición según cualquiera de las reivindicaciones 1 a 5, en la que las sondas presentan la estructura general 5'-D -L -L -L -L -L -X -X -3', en la que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es una posición aleatoria de ANB.
8. Composición según cualquiera de las reivindicaciones 1 a 7, en la que las sondas del primer juego de sondas se marcan con un primer marcador y las sondas del segundo juego de sondas se marcan con un segundo marcador, en los que el primer marcador es diferente del segundo marcador.
9. Composición según la reivindicación 8, en la que el primer y el segundo marcadores son marcadores fluorescentes.
10. Composición según la reivindicación 9, en la que las sondas son sondas de hidrólisis marcadas adicionalmente con un inhibidor.
11. Método para determinar el genotipo en un locus de interés en una muestra obtenida de un sujeto, comprendiendo el método:
- a) poner en contacto la muestra que comprende material genético con la composición según cualquiera de las reivindicaciones 1 a 10, y
- b) detectar la unión de una sonda del primer o del segundo juego de sondas al material genético, determinando de esta manera el genotipo en el locus.
12. Método según la reivindicación 11, en el que el locus es un único nucleótido.

13. Método según la reivindicación 11 ó 12, en el que el método comprende:
- a) llevar a cabo una etapa de amplificación que comprende la puesta en contacto de la muestra con un juego de cebadores con el fin de producir un producto de amplificación que incluye el locus de interés,
 - b) llevar a cabo una etapa de hibridación que comprende poner en contacto el producto de amplificación de la etapa a) con la composición según cualquiera de las reivindicaciones 1 a 10, y
 - c) detectar la hibridación de una sonda del primer o segundo juego de sondas con el material genético, determinando de esta manera el genotipo en el locus.
14. Método según las reivindicaciones 11 a 13, en el que:
- (i) la detección se realiza mediante la medición de la presencia o ausencia de fluorescencia,
 - (ii) la detección se realiza en tiempo real,
 - (iii) el marcador se selecciona de entre el grupo que consiste de fluoresceína, LC-amarillo 555, FAM, VIC, HEX, rodamina-B, rodamina-6G, LC-rojo 610, LC-rojo 640, LC-rojo 670, LC-rojo 705, Cy3, Cy3.5, Cy5, Cy5.5 y un inhibidor,
 - (iv) la etapa de amplificación utiliza un enzima polimerasa que presenta actividad de exonucleasa 5' a 3', y/o
 - (v) la muestra es una muestra biológica, preferentemente una muestra seleccionada de entre el grupo que consiste de un líquido corporal, una muestra de sangre, una muestra de orina, suero, mucosa, esputo, heces, muestra epidérmica, muestra de piel, frotis bucal, esperma, líquido amniótico, células en cultivo y muestra de médula ósea.
15. Biblioteca de por lo menos dos juegos de sondas, en la que la biblioteca comprende una pluralidad de juegos de sondas, presentando cada una de las sondas ocho nucleótidos con la estructura general 5'5'-D-L-L-L-L-L-X-X-3' o 5'-D-L-L-L-L-X-X-3', en las que D es un nucleótido de ADN, cada L es un nucleótido de ANB y cada X es un nucleótido aleatorio de ANB, en el que, dentro de un juego de sondas, todas las sondas presentan secuencias de nucleótidos idénticas con la excepción de los dos y/o tres nucleótidos aleatorios de ANB, en el que en cada nucleótido aleatorio de ANB, la base se selecciona independientemente de entre adenina, citosina, guanina y timina y cualquier secuencia posible resultante de la variación o variaciones de bases en las dos posiciones aleatorias se encuentra representada por una sonda en cada juego de sondas, y en el que un juego de sondas difiere de otro juego de sondas en la secuencia de por lo menos el nucleótido de ADN D o un nucleótido de ANB L.
16. Biblioteca de juegos de sondas según la reivindicación 15,
- (i) en la que el número de juegos es de por lo menos 64, preferentemente de 128, especialmente de por lo menos 256, particularmente de por lo menos 512, más preferentemente de por lo menos 1.024, todavía más preferentemente de 2.048, y/o
 - (ii) en el que los juegos de sondas se encuentran separados espacialmente entre sí.

Fig. 1

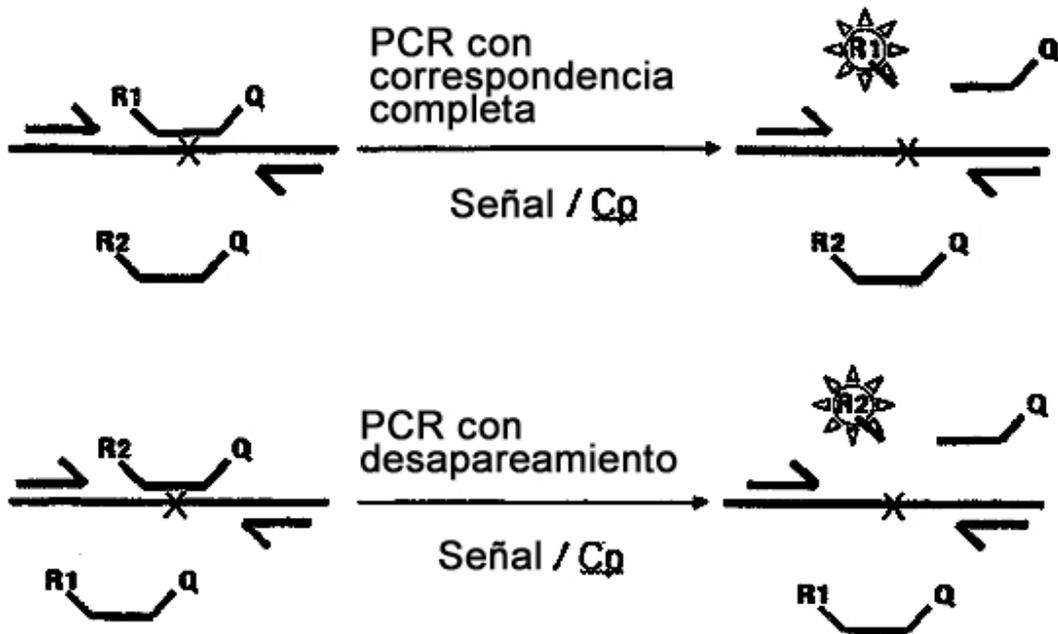


Fig. 2

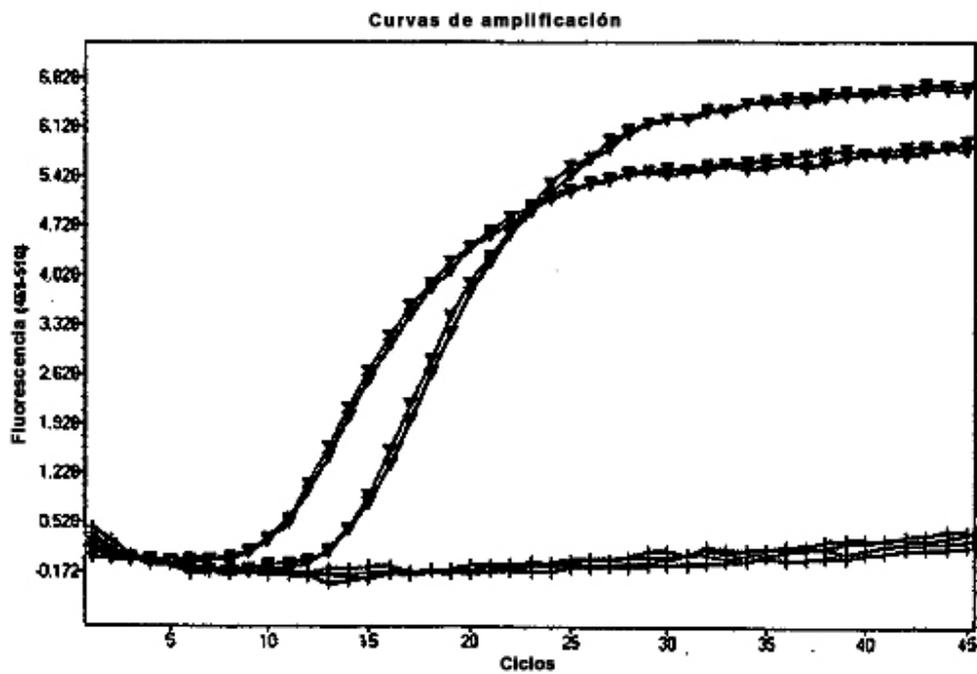


Fig. 3

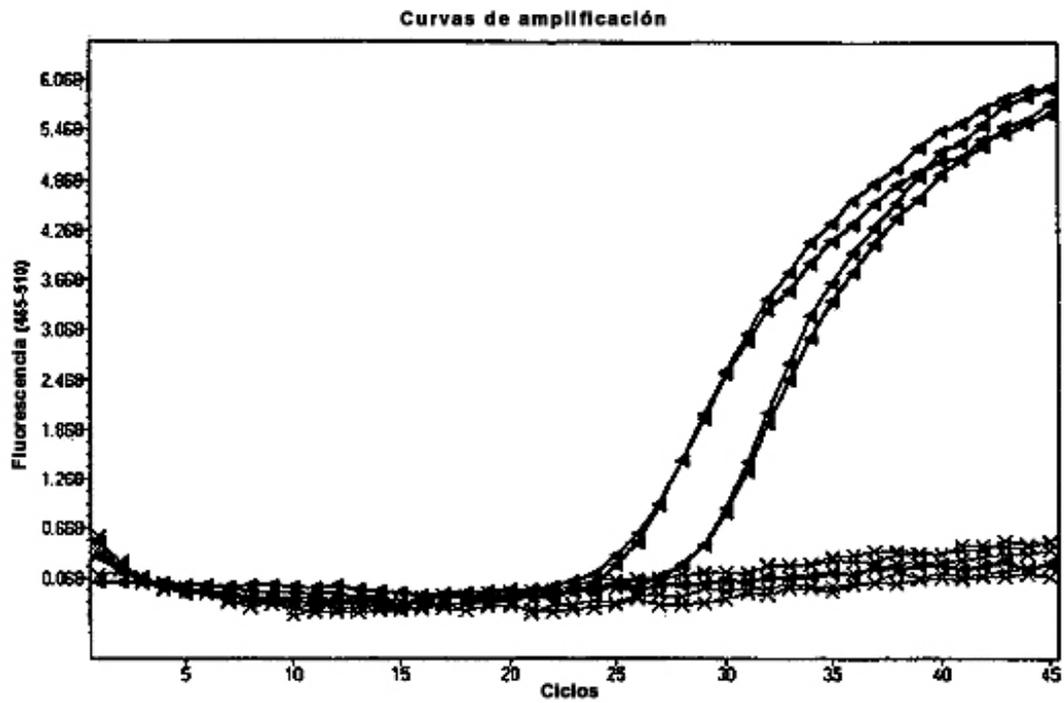


Fig. 4

