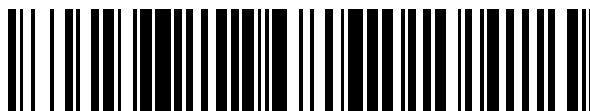


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 392 049**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04719497 .2**

96 Fecha de presentación: **11.03.2004**

97 Número de publicación de la solicitud: **1601798**

97 Fecha de publicación de la solicitud: **07.12.2005**

54 Título: **Biochips con sondas y sus métodos de uso**

30 Prioridad:

13.03.2003 FR 0303112

45 Fecha de publicación de la mención BOPI:

04.12.2012

45 Fecha de la publicación del folleto de la patente:

04.12.2012

73 Titular/es:

**GENEWAVE SAS (100.0%)
172 rue de Charonne
75011 Paris, FR**

72 Inventor/es:

**VALAT, CHRISTOPHE y
LE DOUGUET, CÉCILE**

74 Agente/Representante:

DE ELZABURU MÁRQUEZ, Alberto

ES 2 392 049 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

Biochips con sondas y sus métodos de uso

La invención se refiere al campo de la utilización de sondas de polinucleótidos sin marcar, capaces de formar horquillas, a biochips que contienen dichas sondas y a métodos para su uso. La presente descripción también se refiere a métodos para la creación de tales sondas y biochips. Más particularmente, la descripción describe el uso de dichas sondas sin marcar y biochips para la manipulación y el análisis de secuencias de polinucleótidos y posiblemente de sus moléculas asociadas. La presente descripción también describe métodos para preparar y usar tales sondas y biochips para el análisis de mutaciones, la secuenciación, la detección de variantes por corte y empalme alternativas, el análisis de la expresión génica, el análisis de desequilibrios alélicos y la pérdida de heterocigosidad, y la detección de cualquier ácido nucleico presente en los organismos o los restos procedentes de estos organismos.

La determinación rápida y precisa de la identidad y la cantidad de moléculas específicas en una muestra que contiene muchas moléculas diferentes, es de interés principal en los campos de las aplicaciones médicas y biológicas. Muchos tipos de sondas y sistemas de análisis han sido desarrollados para detectar moléculas específicas, tales como las sondas y los chips para la detección de cambios en la secuencia de ácidos nucleicos.

Un tipo de sonda tal para detectar ácidos nucleicos específicos, puede ser un "interruptor molecular" similar al descrito en el documento de patente de EE.UU. n° 5.118.801, que funciona según el principio de la imposibilidad que tienen los extremos de la sonda para interaccionar entre sí, cuando la porción central de la sonda se hibrida con una secuencia diana. Cada interruptor molecular tiene 2 extremos complementarios que tienen una longitud de al menos 10 nucleótidos, y una porción central de aproximadamente 15 a 115 nucleótidos de longitud. Dependiendo de las condiciones de la prueba, la sonda descrita puede adoptar una conformación "cerrada" o "abierta". Cuando la sonda está en su conformación cerrada, los extremos de la sonda se hibridan entre sí para formar un "vástago", en donde la secuencia central forma un "bucle". En la conformación abierta, la secuencia central de la sonda se hibrida con una secuencia diana complementaria, predeterminada para la que se ha diseñado la sonda. Esta conformación abierta lleva a la disociación de los extremos de la sonda y a la imposibilidad de que estos extremos interaccionen entre sí. Uno o ambos extremos de la sonda descrita contienen una porción de ácido nucleico biológicamente funcional para generar selectivamente una señal detectable, indicativa de la hibridación de la sonda con la secuencia diana predeterminada. Una porción preferida es, por ejemplo, un ARN que permite una replicación exponencial gracias a una ARN polimerasa, dependiente de ARN, en presencia de nucleótidos radioactivos, incorporados en los productos de la replicación.

Otro tipo de sonda, comúnmente llamado "baliza molecular" es similar a la que se ha descrito anteriormente, y se describe en los documentos de patente de EE.UU. n° 5.925.517 y europea EP0745690 (véase también y col., 1996, *Nat. Biotechnol.* 14:303-308; Tyagi y col., 1998, *Nat. Biotechnol.* 16:49-53; Matsuo T. 1998, *Biochimica Biophysica Acta.* 1379:178-184; Sokol y col. 1998, *Proc. Natl. Acad. Sci. USA.* 95:11538-11543; Leone y col. 1998, *Nucleic Acids Res.* 26:2150-2155; Piatek y col. 1998, *Nat. Biotechnol.* 16:359-363; Kostrikis y col. 1998, *Science* 279:1228-1229; Giesendorf y col. 1998, *Clin. Chem.* 44:482-486; Marras y col. 1999 *Genet. Anal.* 14:151-156; Vet y col. 1999, *Proc. Natl. Acad. Sci. USA* 96:6394-6399). Las "balizas moleculares" son moléculas que poseen un fluoróforo interno extinguido en el que la fluorescencia se restaura cuando se unen a un ácido nucleico diana. Están diseñadas de manera similar a un interruptor molecular, de modo que la parte del bucle de la molécula constituye una secuencia de sonda complementaria de una molécula nucleico diana y la parte del vástago es el resultado del apareamiento de las secuencias de los "brazos" complementarios, situadas en los extremos de la secuencia de la sonda. Un grupo fluorescente se fija a un extremo de un brazo. Un grupo extintor se fija al extremo del otro brazo. En ausencia de una molécula diana, estos dos grupos están muy próximos entre sí, provocando la extinción del fluoróforo. Cuando la sonda se hibrida con un ácido nucleico diana que es perfectamente complementario, la "baliza molecular" sufre un cambio conformacional que obliga a los brazos del vástago a disociarse, haciendo así que el fluoróforo y el extintor se alejen uno de otro. Esto permite detectar la hibridación de la sonda con su molécula diana a través de la aparición de fluorescencia.

Actualmente, los sistemas de biochips son ampliamente utilizados para la detección y la medición de determinadas sustancias en muestras complejas. En tales sistemas, la identidad y la cantidad de una sustancia diana en una muestra se determinan midiendo el nivel de asociación de la secuencia diana con sondas diseñadas específicamente para esta secuencia diana. En las tecnologías de biochips de ácidos nucleicos, un conjunto de sondas de ácidos nucleicos, en donde cada una tiene una secuencia definida, se inmoviliza sobre un soporte sólido de tal forma que cada sonda ocupa una posición predeterminada. Una vez que el conjunto de sondas está inmovilizado, el chip, se pone en contacto con una muestra, de modo que las secuencias complementarias se puedan asociar a una sonda inmovilizada, por ejemplo, hibridándose, asociándose o uniéndose a la sonda. Después de retirar el material no asociado, las secuencias asociadas se detectan y se miden.

Las tecnologías de biochips de ADN han hecho posible el seguimiento de los niveles de expresión de un gran número de transcritos genéticos al mismo tiempo (véase, por ejemplo, Schena y col., 1995, *Science* 270:467-470; Lockhart y col., 1996, *Nat. Biotechnology* 14:1675-1680; Blanchard y col., 1996, *Nat. Biotechnology* 14:1649; Ashby

y col., documento de patente de EE.UU. nº 5.569.588 expedida el 29 de octubre, 1996). La tecnología de biochips también se ha utilizado para secuenciar, determinar las huellas genéticas y cartografiar macromoléculas biológicas (documentos de patente de EE.UU. nº 6.270.961 expedida el 7 de agosto, 2001; 6.025.136 expedida el 15 de febrero, 2000; 6.018.041 expedida el 25 de enero, 2000; 5.871.928 expedida el 16 de febrero, 1999; 5.695.940 expedida el 9 de diciembre, 1997). Hay dos formatos principales de biochips de ADN. Para uno de esos formatos, los biochips de ADN se preparan depositando fragmentos de ADN que varían en tamaño desde unas pocas decenas de bases a unas pocas kilobases, sobre una superficie adecuada (véase, por ejemplo: DeRisi y col., 1996, *Nature Genetics* 14:457-460; Shalon y col., 1996, *Genome Res.* 6:689-645; Schena y col., 1995, *Proc Natl Acad Sci USA* 93:10539-11286, y Duggan y col., 1996, *Nature Genetics* Suplemento 21:10-14). Por ejemplo, en el análisis por "transferencia", como la transferencia de manchas "dot blot" o la hibridación de ADN/ADN sobre una membrana (Southern Blotting), las moléculas de ácido nucleico se pueden separar previamente mediante el uso de, por ejemplo, una separación en función del tamaño en una electroforesis en gel, seguida por su transferencia y su inmovilización sobre una membrana filtrante, tal como una membrana de nitrocelulosa o nailon, seguido de su hibridación con una secuencia única marcada (véase, por ejemplo., Nicoloso, M. y col., 1989, *Biochemical and Biophysical Research Communications* 159:1233-1241; Vernier, P. y col., 1996, *Analytical Biochemistry* 235:11-19). Los biochips de ADN complementario (ADNc) se preparan depositando sobre una superficie adecuada los productos de la reacción en cadena de la polimerasa (del inglés, "polymerase chain reaction", PCR) que se presentan en forma de fragmentos de ADN complementarios, cuyo tamaño varía entre 0,6 y 2,4 kb, procedentes de ADNc, ESTs (marcador de secuencia expresada, del inglés, "Expressed Sequence Tag") etc. (véase, por ejemplo: DeRisi y col., 1996, *Nature Genetics* 14:457-460; Shalon y col., 1996, *Genome Res.* 6:689-645; Schena y col., 1995, *Proc. Natl. Acad. Sci. USA* 93:10539-11286; y Duggan y col., 1996, *Nature Genetics* Suplemento 21:10-14). El otro formato de biochips de ADN, denominado matrices de alta densidad de oligonucleótidos contienen miles de oligonucleótidos complementarios de secuencias definidas, que ocupan posiciones conocidas sobre la superficie utilizada. Estos oligonucleótidos se sintetizan *in situ* sobre la superficie utilizando técnicas tales como la imagen litográfica (véase, por ejemplo, Fodor y col., 1991, *Science* 251:767-773; Pease y col., 1994, *Proc. Natl. Acad. Sci. USA* 91:5022-5026; Lockhart y col., 1996, *Nature Biotechnology* 14:1675-1680; los documentos de patentes de EE.UU. nº 5.578.832, 5.556.752, 5.510.270, 5.445.934, 5.744.305 y 6.040.138). Los métodos de fabricación de biochips que emplean tecnologías de chorro de tinta para la síntesis *in situ* de oligonucleótidos, también son bien conocidos por los expertos en la técnica (véase, por ejemplo, Blanchard, documento de solicitud de patente internacional WO 98/41531, publicado el 24 de septiembre 1998; Blanchard y col., 1996, *Biosensors and Bioelectronics*, 11:687-690; Blanchard, 1998, en *Synthetic DNA Arrays in Genetic Engineering*, vol. 20, J.K. Setlow, compilador, Plenum Press, Nueva York, páginas 111-123).

A pesar de estos avances tecnológicos, sigue habiendo una necesidad de biochips mejorados y de nuevos métodos útiles para la caracterización de alto rendimiento de los genes y de los productos de estos genes (por ejemplo, la detección de mutaciones puntuales como el "polimorfismo de un solo nucleótido" (SNP)), inserciones y/o deleciones de un solo nucleótido o múltiples y sucesivas, secuenciación de genes, análisis de la expresión génica, detección de corte y empalme alternativo, análisis de la pérdida de heterocigosidad, así como la detección de cualquier organismo o resto de organismo que contiene los nucleótidos. Entre los problemas encontrados con las tecnologías actuales de biochips, hay que mencionar el tamaño y el posicionamiento de las sondas, la falta de condiciones únicas para realizar el análisis con todas las sondas, la necesidad de sondas redundantes debido al contenido variable en bases G/C que llevan a la pérdida de espacio en el biochip, la necesidad de múltiples oligonucleótidos para la PCR de las secuencias detectadas, así como la necesidad de marcar cada secuencia diana. La presente invención cubre cada una de estas necesidades.

Sumario de la invención

La presente invención consiste en el kit de acuerdo con las reivindicaciones 1 a 12 y los métodos de acuerdo con las reivindicaciones 13 a 22.

La presente invención presenta sondas nucleicas capaces de formar horquillas si marcar, biochips que contienen una pluralidad de estas sondas y métodos que permiten su uso. Cada sonda polinucleotídica capaz de formar una horquilla está constituida por una secuencia de ácido nucleico monocatenario sin marcar, que comprende una secuencia específica de una diana, flanqueada a cada lado por unos extremos constituidos por dos secuencias complementarias entre sí. Cada sonda puede adoptar dos conformaciones. Una de estas conformaciones es la llamada "cerrada" o de "horquilla" y se caracteriza por la hibridación de los dos extremos complementarios de la sonda, que forman el "vástago", mientras que la secuencia específica de la diana forma un "bucle" (la secuencia específica de la diana no se puede hibridar con ninguna otra parte de la sonda). Esta conformación cerrada excluye la hibridación de otras secuencias en los brazos (extremos de la sonda que forman el vástago). La segunda conformación es la llamada "abierta" y se caracteriza por la hibridación de la secuencia diana de la sonda que forma el bucle, con la molécula diana, en donde esta hibridación fuerza a que los dos extremos complementarios se disocien entre sí. Esta conformación abierta evita que los dos extremos complementarios se unan entre sí y permite que otras secuencias, que son complementarias de los extremos, se hibriden con dichos extremos. La hibridación de la molécula diana se detecta mediante el uso de una molécula "informadora" que está prevista específicamente para hibridarse con al menos uno de los extremos libres de la sonda descrita en esta memoria.

Cada sonda descrita en esa memoria se puede proporcionar para discriminar una molécula diana específica que tiene una variación en un nucleótido (por ejemplo, una sustitución, una delección o una inserción de un nucleótido), o variaciones de varios nucleótidos separados por secuencias cortas, o también delecciones o inserciones extensas. Cada sonda en forma de horquilla descrita en esta memoria está diseñada para abrirse cuando se hibrida con la diana para la que se ha preparado específicamente, lo que permite distinguir y detectar las dianas. La capacidad de discriminación permite que múltiples secuencias se dirijan a las dianas en una única solución. En una realización preferida, un biochip comprende dos o más sondas tales como las que se describen en esta memoria, todos los híbridos perfectos formados entre las secuencias específicas de la diana de las sondas y las moléculas diana tienen una temperatura de fusión igual, dentro de un intervalo de 4°C. Preferiblemente, las temperaturas de fusión de los híbridos perfectos formados entre las secuencias específicas de la diana de sondas tales como las descritas en esta memoria y las moléculas diana son iguales, dentro de un intervalo de 3°C. Incluso más preferiblemente, las temperaturas de los híbridos perfectos formados entre las secuencias diana específicas de la diana de sondas tales como las descritas en esta memoria, y las moléculas diana, son iguales, dentro de un intervalo de 1°C.

La longitud de la secuencia específica de la diana de cada sonda descrita en esta memoria, comprende de 6 a 30 nucleótidos, preferiblemente entre 10 y 25, más preferiblemente entre 15 y 20 nucleótidos. Preferentemente, cada uno de los extremos que forman el vástago de cada sonda en forma de horquilla descrita en esta memoria, está constituido por una secuencia de menos de 10 nucleótidos de longitud, preferiblemente entre 5 y 9 nucleótidos de longitud.

La descripción describe también una sonda de ácido nucleico capaz de formar una horquilla, sin marcar, que comprende:

(a) una secuencia específica de una diana, de 6 a 30 nucleótidos de longitud;

(b) un primer brazo de menos de 10 nucleótidos de longitud, situado en 5' de dicha secuencia específica de una diana; y

(c) un segundo brazo de menos de 10 nucleótidos de longitud, situado en 3' de dicha secuencia específica de una diana,

dicha secuencia específica de una diana no es complementaria de ninguna otra parte de dicha sonda; y además, el primer brazo y el segundo brazo son completamente complementarios entre sí.

En una realización preferida, la secuencia específica de la diana tiene una longitud de 10-25 nucleótidos, más particularmente de 15 a 20 nucleótidos. Cuando una molécula diana se hibrida con la secuencia específica de la diana, la sonda adopta una conformación "abierta" y una molécula "informadora" se puede hibridar con el primer o el segundo brazo. La molécula "informadora" comprende menos de 10 nucleótidos perfectamente complementarios de la secuencia de ácido nucleico del primer brazo o del segundo brazo. La molécula "informadora" comprende un marcador detectable. El marcador detectable puede ser un análogo de nucleótido, un marcador fluorescente, biotina, iminobiotina, un antígeno, un cofactor, dinitrofenol, ácido lipoico, un compuesto olefínico, un polipéptido, una molécula rica en electrones, una enzima o un radioisótopo. Preferiblemente, la Dtm (diferencia entre la temperatura de fusión (Tm) del híbrido perfecto formado entre la secuencia específica de la diana de la sonda y la molécula diana, y la temperatura de fusión (Tm) del híbrido formado entre el primer brazo y el segundo brazo de la sonda) es mayor que 10, más preferiblemente es igual a 15. Alternativamente, la Dtm es menor que 10.

La descripción también describe una composición que comprende al menos una sonda tal como la descrita en esta memoria y al menos una molécula "informadora".

La descripción también describe un biochip de sondas capaces de formar horquillas que comprende:

(a) un soporte; y

(b) al menos dos sondas tales como las descritas en esta memoria.

En una realización preferida, la sonda se fija a un soporte. Opcionalmente, dicha sonda comprende además un espaciador y está fijada a dicho soporte a través de dicho espaciador. El soporte puede consistir en una superficie de vidrio funcionalizado, una superficie de plástico funcionalizado, una superficie de metal funcionalizado, una superficie conductora de metal, una superficie conductora de plástico, un soporte poroso, un metal poroso, una fibra óptica, un soporte obtenido a partir de fibra de vidrio, dióxido de silicio, una membrana lipídica funcional, un liposoma o una membrana de filtración. Más particularmente, la superficie de plástico funcionalizado puede ser de poliestireno; la superficie de metal funcionalizado puede ser de platino, oro o níquel; la superficie de plástico conductor puede ser un soporte a base de carbono, pudiendo ser este soporte un polímero; y el soporte poroso puede ser de vidrio. En una realización particular, los primeros brazos de cada sonda tienen una secuencia idéntica. Así, una misma molécula "informadora" se puede hibridar con cada sonda. Preferiblemente, todos los híbridos perfectos formados entre las secuencias específicas de la diana de las sondas y las moléculas diana, tiene una temperatura de fusión igual, dentro de un intervalo de 4°C, preferiblemente dentro de un intervalo de 1°C.

Preferiblemente, la diferencia entre la temperatura de fusión del híbrido formado entre la secuencia específica de la diana de cada sonda y la molécula diana, y la temperatura de fusión de la asociación entre la secuencia específica de la diana y una molécula para la que no se ha diseñado la secuencia específica de la diana de la sonda, es mayor o igual a 5°C, preferiblemente 8°C. En una realización particular, las Dtm de al menos dos sondas son iguales, dentro de un intervalo de 1°C.

La presente descripción describe igualmente un sistema de direccionamiento universal en el que un biochip normalizado, constituido por sondas en forma de horquilla inmovilizadas tal como la descrita en esta memoria, se usa con sondas en forma de horquilla o lineales, no inmovilizadas, específicas de las moléculas diana y que permiten la detección de esas moléculas diana. Preferiblemente, las sondas no inmovilizadas tienen forma de horquilla.

La descripción describe también un sistema de direccionamiento universal sin marcar, que comprende:

(a) al menos dos primeras sondas sin marcar, que comprenden:

(i) una secuencia específica de una diana, de 6 a 30 nucleótidos de longitud;

(ii) una secuencia denominada marcadora de 10 a 50 nucleótidos de longitud, conectada al extremo 5' de la secuencia específica de la diana,

en donde dicha secuencia específica de una diana no es complementaria de ninguna otra parte de dicha sonda sin marcar;

(b) un biochip tal como el descrito en esta memoria.

Preferiblemente, dichas primeras sondas sin marcar comprenden:

(i) una secuencia específica de una diana, de 6 a 30 nucleótidos de longitud;

(ii) un primer brazo de menos de 10 nucleótidos de longitud, situado en 5' de dicha secuencia específica de una diana;

(iii) un segundo brazo de menos de 10 nucleótidos de longitud, situado en 3' de dicha secuencia específica de una diana; y

(iv) una secuencia denominada marcadora de 10 a 50 nucleótidos de longitud conectada con el primer brazo o el segundo brazo,

en donde dicha secuencia específica de una diana no es complementaria de ninguna otra parte de dicha sonda sin marcar; y, además, en donde el primer brazo y el segundo brazo son perfectamente complementarios. Preferiblemente, todos los primeros brazos de las primeras sondas tienen una secuencia idéntica. Así, una misma molécula "informadora" se puede hibridar con una u otra de las sondas primera y segunda.

Preferiblemente, el biochip comprende al menos dos segundas sondas capaces de formar una horquilla que comprende:

(i) una secuencia específica de una segunda diana, de 6 a 30 nucleótidos de longitud;

(ii) un segundo primer brazo de dicha segunda sonda de menos de 10 nucleótidos de longitud y situado en 5' de dicha segunda secuencia específica de dicha diana;

(iii) un 2º segundo brazo de dicha segunda sonda de menos de 10 nucleótidos de longitud y situado en 3' de dicha segunda secuencia específica de dicha diana; y

(iv) un espaciador que conecta el primer o el segundo brazo con dicho soporte,

en donde dicha secuencia específica de dicha diana de la segunda sonda no es complementaria de ninguna otra parte de dicha segunda sonda; en donde dicho primer brazo y dicho segundo brazo de dicha segunda sonda son completamente complementarios entre sí; y además en donde dicha secuencia marcadora de una de las primeras sondas completada de la secuencia del brazo conectado a este marcador es complementaria de dicha segunda secuencia diana de una de las segundas sondas.

La descripción también describe métodos para la preparación de sondas y biochips tales como los descritos en esta memoria.

La descripción describe adicionalmente métodos de uso de sondas y biochips tal y como se describen en esta memoria. Los aspectos particulares se refieren al análisis de mutaciones, la secuenciación, el análisis de la expresión génica, el análisis de la pérdida de heterocigosidad y desequilibrio alélico, y la detección de cualquier

organismo o resto de un organismo que contiene ácidos nucleicos.

La descripción describe también un método para detectar ácido nucleico que comprende:

- (a) poner en contacto *ex vivo* una muestra de ácido nucleico con un biochip que comprende al menos dos sondas tales como las descritas en esta memoria; y
- 5 (b) detectar una señal que proviene de al menos una de dichas sondas del biochip que ha adoptado una conformación abierta, después del contacto establecido durante la etapa (a).

También describe un método de detección *ex vivo* de una variante genética en una muestra de ácido nucleico que comprende:

- (a) poner en contacto una muestra de ácido nucleico con un biochip que comprende al menos dos sondas tales como las descritas en esta memoria, en donde al menos una sonda del biochip es una sonda específica de una variante genética, que tiene una región de bucle perfectamente complementaria con dicha variante genética, y
- 10 (b) detectar una señal que proviene de la sonda específica de la variante genética, en donde la señal detectada en la etapa (b) indica la presencia de una variante genética en la muestra de ácido nucleico.

Preferiblemente, la variante genética detectada es un polimorfismo de un solo nucleótido.

- 15 También se describe un método de detección *ex vivo* de cualquier organismo que contiene ácidos nucleicos o cualquier resto de es dentro de un intervalo de os organismos que comprende:

- (a) poner en contacto una muestra de ácido nucleico con un biochip que comprende al menos dos sondas tales como las descritas en esta memoria, en donde al menos una sonda es una sonda específica del ácido nucleico de un organismo o de un resto de ese organismo; y
- 20 (b) detectar una señal que proviene de dicha sonda específica de un organismo que contiene ácido nucleico, en donde la señal detectada en la etapa (b) indica la presencia del organismo que contiene el ácido nucleico o restos de ese organismo.

Preferiblemente, el organismo que contiene el ácido nucleico es un virus o una bacteria.

- 25 La descripción también describe un método de detección *ex vivo* de un producto del corte y empalme alternativo de un gen en una muestra de ácido nucleico que comprende:

- (a) poner en contacto una muestra de ácido nucleico con un biochip con sondas que comprenden al menos dos sondas tales como las descritas en esta memoria, en donde al menos una sonda del biochip es una sonda específica de un exón del gen o de la unión de dos exones; y
- (b) detectar una señal que proviene de dicha sonda específica del exón o de la unión de dos exones,
- 30 en donde la señal detectada en la etapa (b) indica la presencia del producto por corte y empalme alternativo de un gen en la muestra de ácido nucleico.

Preferiblemente, la muestra de ácido nucleico comprende ARNm o ADNc.

Además se describe un método de secuenciación *ex vivo* de un oligonucleótido que comprende:

- 35 (a) poner en contacto una muestra que contiene el oligonucleótido con un biochip que comprende al menos dos sondas tales como las descritas en esta memoria; y
- (b) detectar una señal procedente de al menos una sonda de dicho biochip;

en donde la señal detectada en la etapa (b) se utiliza para determinar la secuencia del oligonucleótido.

- 40 La detección de la etapa (b) de los métodos mencionados anteriormente, se obtiene mediante el uso de una molécula "informadora" marcada con un marcador detectable. La muestra de la etapa (a) puede marcarse antes con un marcador detectable.

La descripción describe también un método de detección *ex vivo* de desequilibrios alélicos y pérdidas de heterocigosidad en una muestra de ácido nucleico, que comprende:

- (a) amplificar al menos una región de ADN cromosómico de tipo microsatélite mediante una pareja de cebadores a partir de al menos dos muestras de ácido nucleico procedentes de tejidos o fluidos biológicos, en donde al menos una de las muestras procede de células o de tejidos que no presentan desequilibrio alélico o pérdida de
- 45

heterocigosidad, y en donde cada tejido o fluido está marcado de manera diferente durante la amplificación;

(b) eliminar estos cebadores después de la amplificación;

(c) poner en contacto dichos productos de la amplificación con un biochip con sondas que comprende al menos dos sondas tales como las descritas en esta memoria, en donde al menos una sonda del biochip es complementaria de un cebador utilizado para amplificar dicha región de ADN cromosómico; y

(d) detectar las señales procedentes de al menos una sonda de dicho biochip;

en donde las señales detectadas en la etapa (d) se utilizan para determinar la presencia de un desequilibrio alélico o una pérdida de heterocigosidad en una de las muestras de ácido nucleico.

La descripción describe kits que comprenden biochips de sondas tales como los descritos en esta memoria, para determinar la presencia o ausencia de diversas moléculas en muestras biológicas o médicas. Cada kit contiene al menos un biochip de sondas tal como el descrito en esta memoria y uno o más reactivos adicionales necesarios para la detección de moléculas diana específicas. El kit puede comprender además un conjunto de sondas tales como las descritas en esta memoria, sin inmovilizar. Además, el kit comprende una molécula "informadora".

Descripción de las Figuras

La Fig. 1 representa las características de las sondas en forma de horquilla.

La Fig. 2 representa las diferentes conformaciones de una sonda en forma de horquilla inmovilizada.

La Fig. 3 representa cómo una señal puede ser generada por la hibridación una sonda en forma de horquilla con la secuencia diana.

La Fig. 4 representa la detección de una diana con una sonda en forma de horquilla y un sistema de marcador que emplea una polimerasa.

La Fig. 5 representa un sistema de direccionamiento universal.

La Fig. 6 representa la secuencia de un ARN mensajero, de una variante de ese ARN y las regiones de ese ARNm que son dianas de las sondas en forma de horquilla.

La Fig. 7 representa las secuencias de ácido nucleico utilizadas para mostrar cómo las sondas en forma de horquilla interaccionan con las dianas y las moléculas "informadoras".

La Fig. 8 representa un gráfico de barras de las intensidades de la fluorescencia después de la hibridación de las secuencias diana marcadas, con las sondas en forma de horquilla.

La Fig. 9 representa un gráfico de barras de las intensidades de la fluorescencia tras la hibridación sucesiva de una secuencia diana marcada, con las sondas en forma de horquilla, y después con una molécula "informadora".

La Fig. 10 representa un gráfico de barras de las intensidades de la fluorescencia tras la hibridación sucesiva de una secuencia diana marcada, con las sondas en forma de horquilla, a continuación, con una molécula "informadora", seguido de la eliminación de las interacciones no específicas.

La Fig. 11 representa un gráfico de barras de las intensidades de la fluorescencia después de la hibridación de varias dianas marcadas, con las sondas en forma de horquilla.

La Fig. 12 representa un gráfico de barras de las intensidades de la fluorescencia después de la hibridación de varias dianas marcadas, con las sondas en forma de horquilla, en condiciones de hibridación más rigurosas que las de la Fig. 11.

La Fig. 13 representa un método de análisis de la pérdida de heterocigosidad, después de las etapas de amplificación con PCR de las muestras, hasta la hibridación de los productos de la PCR en el biochip.

La Fig. 14 representa la evolución de la relación de discriminación de las sondas en forma de horquilla (Rd), como una función de la diferencia de Tm entre el vástago y el bucle (Dtm).

La Fig. 15 representa la hibridación de los productos de la PCR resultante de la amplificación del exón 3 del gen PMP22 en un biochip en forma de horquilla.

Descripción detallada de la invención

La descripción describe sondas capaces de formar horquillas sin marcar, biochips que contienen una pluralidad de sondas tales como las descritas y se refiere a métodos para preparar y usar sondas y biochips tales como los

descritos. Esto se describe en detalle y se ilustra con los ejemplos en la sección siguiente.

Sondas susceptibles de formar horquillas

La descripción describe un biochip de sondas capaces de formar horquillas que comprende:

- (a) un soporte; y
- 5 (b) al menos dos sondas que pueden formar horquillas sin marcar, en donde cada sonda sin marcar comprende:
 - (i) una secuencia específica de una sonda, de 6 a 30 nucleótidos de longitud;
 - (ii) un primer brazo de menos de 10 nucleótidos de longitud situado en 5' de la secuencia específica de la diana;
 - 10 (iii) un segundo brazo de menos de 10 nucleótidos de longitud situado en 3' de la secuencia específica de la diana;
 - (iv) un espaciador que conecta el primer o el segundo brazo a dicho soporte;

15 en donde la secuencia específica de una diana no es complementaria de ninguna otra parte de la sonda sin marcar; y en donde dicho primer brazo y dicho segundo brazo son perfectamente complementarios entre sí, en cada una de las sondas.

Por "sin marcar" se entiende que la sonda no contiene un marcador detectable. Por ejemplo, la sonda no contiene un análogo de nucleótido, un marcador fluorescente, biotina, iminobiotina, un antígeno, un cofactor, dinitrofenol, ácido lipoico, un compuesto olefínico, un polipéptido, una molécula rica en electrones, una enzima o un isótopo radioactivo, que permita la detección directa de la sonda.

20 Dos ejemplos de sondas sin marcar en forma de horquilla tales como las descritas en esta memoria, se ilustran en las Figs. 1A y 1B. En el contexto de esta descripción, una sonda de horquilla es una secuencia de ácido nucleico monocatenario que comprende un "bucle" 101, en donde la longitud de la secuencia específica de una diana puede variar de 6 a 30 nucleótidos, preferiblemente una longitud de entre 10 y 25 nucleótidos, y más preferiblemente de 15 a 20 nucleótidos, y dos extremos, o "brazos" que flanquean la secuencia específica de la diana. Un brazo 102 está unido al extremo 3' de la secuencia específica de la diana, y el otro brazo 103 está unido al extremo 5' de la secuencia específica de la diana. Cada brazo comprende preferiblemente menos de 10 nucleótidos perfectamente complementarios de los del otro brazo. Incluso más preferiblemente, cada brazo comprende de 4 a 9 nucleótidos, perfectamente complementarios de los del otro brazo. Convenientemente, cada brazo comprende 6-8 nucleótidos completamente complementarios de los del otro brazo. Si dos o más sondas en forma de horquilla se tienen que utilizar juntas (por ejemplo, en un biochip), todas las sondas en forma de horquilla se pueden diseñar para que tengan brazos cuyas secuencias tengan la misma composición.

Las sondas en forma de horquilla tales como las descritas en esta memoria también pueden comprender un "espaciador" 104. En el contexto de su uso, un espaciador se corresponde a una molécula que está fijada a uno de los brazos de la sonda de horquilla, y sirve como un medio de inmovilización permanente de la sonda de horquilla sobre un soporte 105 como, por ejemplo, pero no limitado a un portaobjetos de vidrio, para formar un biochip con sondas de horquilla. Como se muestra en la Fig. 1B, las sondas de horquilla tales como las descritas en esta memoria, pueden comprender además una secuencia "marcadora" 106 que se corresponde preferiblemente a 5-50 nucleótidos adicionales, incluso más preferiblemente, 5 a 20 nucleótidos, y que se puede utilizar en un sistema de direccionamiento universal. Tal secuencia marcadora se puede utilizar para "dirigir" una sonda de horquilla a una ubicación específica en un biochip, a través de su hibridación con una secuencia inmovilizada sobre el biochip que es perfectamente complementaria de la del marcador o de una parte del marcador.

Diseño de sondas en forma de horquilla

Las sondas en forma de horquilla se pueden diseñar para utilizar con una concentración determinada de sales, una temperatura determinada y una concentración determinada de uno o varios agentes químicos adicionales, tales como formamida, dimetilsulfóxido, cloruro de tetrametilamonio y detergentes. La temperatura fija a la que se pueden utilizar las sondas de horquilla, se puede obtener calculando en primer lugar la temperatura de fusión ("T_m") de la secuencia específica de la diana y de la molécula diana, utilizando el método de cálculo de la T_m más cercana, denominado "vecino más próximo" con el programa informático Meltcalc (véase, Schütz, E. y col., 1999, *Biotechniques* 27:1218-1224; Allawi H.T. y col. 1998, *Biochemistry* 37(8):2170-9; Allawi H.T. y col. 1997, *Biochemistry* 36:10581-94; Peyret N. y col. 1999, *Biochemistry* 38:3468-77) u Oligo 6 (Molecular Biology Insights, Inc., Cascade, CO); Meltcalc., (www.meltcalc.de). Meltcalc es un programa informático que se presenta en forma de una hoja de cálculo que permite calcular el punto de fusión termodinámica después de la hibridación de oligonucleótidos que tienen o no desemparejamientos. Meltcalc da la temperatura de fusión de un oligonucleótido variando las concentraciones de sales y de oligonucleótidos. La temperatura de fusión preferida del híbrido formado

entre la diana y la secuencia específica de la diana ("T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana") es 60°C +/-10°C para concentraciones de secuencias específicas de la diana de 100 nM y concentraciones de sales de 100 mM, en donde el sodio (NaCl) y el magnesio (MgCl) son las sales preferidas.

Cuando dos o más sondas de horquilla están diseñadas para utilizarlas conjuntamente, se prefiere que cada T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana de la sonda de horquilla, tenga un valor más o menos igual a las de los híbridos perfectos formados entre las secuencias específicas de las dianas y las moléculas diana de todas las otras sondas de horquilla. Preferiblemente, la diferencia del valor de las T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, es igual o inferior a 4°C. Más preferiblemente, la diferencia del valor de las T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, es igual o inferior a 3°C. Incluso más preferiblemente, la diferencia del valor de las T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, es igual o menor que 2°C. Lo más preferible, la diferencia del valor de las T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, es igual o inferior a 1°C. Las variaciones de cada T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, se pueden obtener mediante diversas manipulaciones de la secuencia específica de la diana. Por ejemplo, y sin ninguna limitación, la T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, se puede aumentar o reducir variando el número de nucleótidos en la secuencia específica de la diana. Un aumento del número de nucleótidos presentes en la secuencia específica de la diana, aumentando el número de nucleótidos que se asocian a la molécula diana, aumenta la T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana. Al contrario, una disminución del número de nucleótidos de la secuencia específica de la diana, reduce la T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana. Otros medios para modificar las T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, se describen posteriormente.

Una vez que se ha diseñado una secuencia del bucle, con una T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana, en donde el valor se sitúa en un intervalo aceptable, los brazos se pueden diseñar para que sean complementarios entre sí, y después se añaden a la secuencia del bucle. Preferiblemente, el contenido en G/C del vástago es del 80% con un valor preferido de T_m del vástago de 40°C-80°C. Una vez que se ha diseñado cada sonda, se utiliza el programa mfold (Zuker, M. mfold-2.3. <ftp://snark.wustl.edu>) para verificar que no esté presente ningún bucle interno adicional en la sonda de horquilla y que la sonda adopta bien una conformación de horquilla solo por la asociación de los nucleótidos del vástago. La falta de homología entre cada bucle y cualquier otra secuencia de nucleótidos (por ejemplo, un producto de la PCR diferente del analizado, un ADNc diferente del analizado), utilizada en el análisis, se comprueba entonces. La alineación de todas las sondas de horquilla con cada producto de la PCR que se puede utilizar en el análisis, se debería realizar también utilizando herramientas de alineación tal como LALIGN (<http://fasta.bioch.virginia.edu/fasta/lalign.html>) o LFASTA (<http://www.2.igh.cnrs.fr/fasta/lfasta-query.html>).

Para cada sonda complementaria de una secuencia que se va a analizar, la diferencia entre la T_m del bucle y la del vástago de la sonda, se calcula mediante la expresión: D_{tm} = T_m s- T_m l con T_m s = T_m del vástago, y T_m l = T_m del bucle, en donde las T_m se determinan tal y como se ha descrito anteriormente. En particular, la T_m del vástago se corresponde a la temperatura de fusión (T_m) del híbrido formado entre el primer brazo y el segundo brazo de la sonda. La T_m del bucle se corresponde a la diferencia entre la temperatura de fusión (T_m) del híbrido perfecto formado entre la secuencia específica de la diana de la sonda y la molécula diana.

En una realización preferida, esta D_{tm} debe ser para cada sonda, mayor que 10, con una preferencia por un valor cercano a 15 (por ejemplo, entre 12 y 17, preferiblemente entre 13 y 16 o entre 14 y 16), por ejemplo, igual a 15. Sin embargo, para aplicaciones en las que la sensibilidad de la señal debe estar favorecida sobre la capacidad de discriminación de las sondas, este valor de D_{tm} puede ser menor que 10.

Para un conjunto de sondas que se tienen que utilizar conjuntamente en un solo experimento (en particular, la aplicación en un mismo biochip), la distribución de los valores de D_{tm} para el conjunto de las sondas debe estar centrada preferiblemente en un valor mayor que 10, preferiblemente un valor medio de 15 con una desviación estándar de 5. Sin embargo, para aplicaciones en las que se debe favorecer la sensibilidad de la señal sobre la capacidad de discriminación de las sondas, este valor medio de D_{tm} puede ser inferior a 10, con una desviación estándar de 5.

Para dos sondas que se tienen que utilizar conjuntamente para analizar dos alelos de una secuencia (polimorfismo o mutación), las D_{tm} de estas sondas deben ser lo más similar posible. Las D_{tm} de dos sondas utilizadas para analizar dos alelos diferentes de una secuencia son idénticas, dentro de un intervalo de 2, preferiblemente dentro de un intervalo de 1. Estos valores de la D_{tm} son los valores preferidos que pueden ser aceptables para unos valores de desviación <1 o >2 para una pareja de sondas, en algunos casos en donde la composición de bases de las sondas no permite que se obtengan valores de D_{tm} idénticos.

Inmovilización de las sondas de horquilla y los biochips

Las sondas de horquilla se pueden inmovilizar o incluir sobre un soporte para preparar productos tales como, en

particular, biochips con sondas de horquilla, específica de una o varias moléculas diana. Las sondas de horquilla también se pueden inmovilizar sobre una variedad de soportes, tales como especialmente perlas o partículas de tamaño inferior a la micra. Cada sonda de horquilla específica de una diana se puede identificar después de la inmovilización por una propiedad física, química u óptica que distingue cada sonda de horquilla de las otras. Las sondas también se pueden sintetizar directamente sobre una superficie sólida utilizando un sistema bidimensional o el sistema de direccionamiento individual. Para lograr la inmovilización de las sondas de horquilla, estas sondas contienen al menos un grupo fijado, un "espaciador", tal y como se ha descrito anteriormente. Las sondas de horquilla que tienen un espaciador también se pueden inmovilizar gracias a un grupo inmovilizado sobre el soporte y que presenta afinidad hacia el espaciador.

La detección de las moléculas diana se mide preferiblemente mediante el uso de biochips con sondas tales como las descritas en esta memoria. Un biochip es un conjunto organizado de sitios de unión direccionables en términos de posición sobre un soporte; por ejemplo, un soporte sólido en el que cada sonda de horquilla está inmovilizada sobre la superficie del soporte. Preferiblemente, los soportes utilizados son preferentemente portaobjetos de vidrio modificados o funcionalizados por silanización, con el fin de crear grupos químicos adecuados para la inmovilización covalente de las sondas de horquilla. Estos grupos pueden ser preferentemente aminas primarias, tioles, carboxilos o aldehídos. Otras superficies útiles para la inmovilización pueden ser preferentemente dióxido de silicio, polímeros plásticos como el poliestireno, o soportes conductores de corriente como las superficies metálicas o de carbono vítreo. El soporte puede ser poroso o no poroso. Por ejemplo, las sondas de horquilla se pueden inmovilizar sobre filtros o membranas de nitrocelulosa o nailon. La descripción describe por tanto sondas de horquilla sin marcar, caracterizada porque el soporte consiste en una superficie de vidrio funcionalizada obtenida por silanización con un silano que es portador de un grupo que permite un enlace covalente con las sondas aminadas o tioladas, una superficie de plástico funcionalizada obtenida por la creación de grupos funcionales que permiten un enlace covalente con las sondas aminadas o tioladas, como una oxidación química o electroquímica, una superficie metálica funcionalizada por la creación de grupos funcionales que permiten un enlace covalente con las sondas aminadas o tioladas, como una oxidación química o electroquímica, una superficie conductora de metal, una superficie conductora de plástico, un soporte poroso (preferentemente de vidrio), un metal poroso, una fibra óptica, un soporte obtenido a partir de una fibra de vidrio, dióxido de silicio, una membrana lipídica funcional, un liposoma o una membrana de filtración. La superficie de plástico funcionalizada se realiza preferentemente en poliestireno. La superficie de metal funcionalizada se realiza preferentemente en platino, en oro o en níquel. La superficie de plástico conductora es preferiblemente un soporte a base de carbono. El soporte a base de carbono es preferiblemente un polímero.

Tales métodos de fijación de las sondas son bien conocidos en la técnica (véase, por ejemplo, Cass y col. compiladores, 1998, *Immobilized Biomolecules in Analysis, A practical approach*, Ed. Oxford University Press, Great Clarendon Street, Oxford)

Los biochips con sondas de horquilla se pueden preparar utilizando cualquier método conocido en la técnica anterior. Sin embargo, los biochips producidos presentan generalmente ciertas características. Los biochips de horquilla descritos en esta memoria son reproducibles en términos de fabricación, lo que permite la producción de múltiples copias de un biochip dado, y su comparación. La inmovilización de las sondas de horquilla sobre un biochip se puede conseguir a través de una transferencia de fluido con un "spotter" convencional, o por transferencia electrocinética o por electropolimerización de oligonucleótidos electroactivos, o a través de cualquier medio de unión química directa de las sondas sobre una superficie, o la inclusión de la sonda en un soporte sólido, o por síntesis *in situ* por procedimientos de fotolitografía utilizando un conjunto de máscaras prefabricadas (utilizando bases modificadas adaptadas a este procedimiento.) Los biochips descritos en esta memoria están hechos de materiales estables bajo las condiciones utilizadas para la hibridación de ácidos nucleicos. Los biochips son generalmente pequeños, por ejemplo de 1 cm² a 25 cm², preferiblemente de 1 a 3 cm². Sin embargo, los biochips de tamaño mayor o menor también se consideran en esta memoria. Los biochips de mayor tamaño se pueden preferir para la determinación simultánea de un gran número de dianas diferentes. En la presente descripción, la densidad de las sondas de horquilla presente en el biochip puede variar. La densidad puede oscilar entre varias sondas (por ejemplo, 3, 10, 30) y hasta 100 sondas diferentes (es decir, no idénticas) por cm² o más.

En esta descripción, las sondas de horquilla se pueden inmovilizar sobre una superficie sólida a fin de constituir un biochip de sondas que tienen especificidades diferentes para las secuencias diana. La inmovilización de las sondas se realiza de tal manera que las sondas no inhiben la capacidad de los brazos de cada sonda para autoasociarse, o asociarse con las moléculas diana o los reactivos utilizados en el análisis. Como se ha mencionado anteriormente, las sondas descritas en esta memoria se han diseñado para adoptar dos tipos de estructuras conformacionales en función de las condiciones del análisis, y la presencia o ausencia de secuencias de ácido nucleico complementarias. Una de las estructuras es la conformación "cerrada" (de horquilla) en la cual las secuencias de los dos brazos se hibridan juntas. Esta estructura se obtiene en condiciones de análisis favorables y en ausencia de una molécula diana que se asocie con la secuencia específica de la diana de la sonda de hibridación. La segunda estructura es la conformación "abierta" en la que las secuencias de los brazos no se hibridan entre sí. Esta conformación se puede detectar cuando la secuencia específica de la diana se asocia con su diana específica. Un método preferido para realizar un biochip tal como el descrito en esta memoria, se obtiene mediante la fijación de las sondas descritas en

esta memoria, por depósito directo o "impresión" sobre los portaobjetos de vidrio, tal y como describen en general Schena y col. 1995, *Science* 270:467-470. Este método es particularmente útil para preparar biochips con ADN complementario (véase, por ejemplo, DeRisi y col., 1996, *Nature Genetics* 14:457-460; Shalon y col., 1996, *Genome Res.* 6:639-645; Schena y col., 1995, *Proc. Natl. Acad. Sci. USA* 93:10539-11286). Otros métodos para la fabricación de biochips, por ejemplo, el uso de máscaras, también se pueden utilizar (Maskos y Southern, 1992, *Nucl. Acids. Res.* 20:1679-1684). En principio, es tal y como se ha mencionado anteriormente, se puede utilizar cualquier tipo de biochip, como las "transferencias de puntos" realizadas por hibridación sobre una membrana de nailon (véase Sambrook y col., compiladores, 1989, *Molecular Cloning: A Laboratory Manual*, vols. 1-3, 2ª Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York).

Los biochips descritos en esta memoria se pueden fabricar con un dispositivo de impresión con chorro de tinta que permite la síntesis de oligonucleótidos, por ejemplo mediante el uso de los métodos y sistemas descritos por Blanchard en el documento de patente internacional WO 98/41531, publicada el 24 septiembre de 1998; Blanchard y col., 1996, *Biosensors and Bioelectronics*, 11:687-690; Blanchard, 1998, *Synthetic DNA Arrays in Genetic Engineering*, vol. 20, J.K. Setlow, compilador, Plenum Press, Nueva York, págs. 111-123; y el documento de patente de EE.UU. nº 6.028.189. Más específicamente, las sondas de horquilla de tales biochips se sintetizan preferiblemente, por ejemplo, sobre un portaobjetos de vidrio por deposición en serie de bases de nucleótidos individuales, en la forma de "microgotas" de disolvente con tensión superficial elevada, tal como el carbonato de polipropileno. Las microgotas tienen volúmenes pequeños (por ejemplo 100 pL o menor, preferiblemente 50 pL o menor) y están separadas entre sí sobre el biochip (por ejemplo, utilizando dominios hidrófobos) para formar pozos de tensión superficial circular, que definen las posiciones de los elementos del biochip (es decir, las diferentes sondas). Las sondas de polinucleótidos de horquilla se pueden fijar covalentemente a la superficie por los extremos 3' o 5' del polinucleótido de horquilla.

Moléculas diana

Las moléculas diana que se pueden analizar por los métodos y composiciones descritos en esta memoria, incluyen moléculas de ácido nucleico, en particular moléculas de ARN, tales como, por ejemplo, moléculas de ARN mensajero (ARNm), moléculas de ARN ribosómico (ARNr), moléculas de ARNc (es decir, moléculas de ARN preparadas mediante transcripción *in vitro* a partir de moléculas de ADN complementario) y fragmentos de estas moléculas. Las moléculas diana que se pueden analizar también por los métodos y composiciones descritos en esta memoria, incluyen principalmente moléculas de ADN, tales como moléculas de ADN genómico, moléculas de ADNc y fragmentos de estas moléculas que incluyen oligonucleótidos, ESTs, STSs, secuencias microsatélites, etc. Otras moléculas diana que también se pueden analizar por los métodos y composiciones descritos en esta memoria, incluyen principalmente proteínas, complejos que comprenden proteínas, otras moléculas y complejos de estas moléculas que tienen afinidad hacia el ADN, tales como, por ejemplo, los factores de transcripción, las proteínas del sistema de reparación del ADN y los anticuerpos anti-ADN (Fang, W. y col. 2000. *Anal. Chem.* 72:3280-5; Bar-Ziv, R. y col. 2001. *P.N.A.S. U.S.A.* 98:9068-73; Hamaguchi, N. y col. 2001. *Anal. Biochem.* 294:126-31). Para los factores de transcripción, las sondas de horquilla se podrían utilizar para determinar qué secuencia que tiene supuestamente afinidad hacia un factor de transcripción, se puede diseñar en el bucle de una sonda de horquilla e utilizar para determinar qué secuencia tiene la mayor afinidad hacia el factor de transcripción.

Las moléculas diana pueden tener cualquier origen. Por ejemplo, las moléculas diana pueden ser moléculas naturales de ácido nucleico, tales como moléculas de ADN genómico o extragenómico, aisladas a partir de un organismo, o moléculas de ARN, tales como moléculas de ARN mensajero aisladas a partir de un organismo. Alternativamente, las moléculas diana se pueden sintetizar. Estas moléculas diana pueden ser, por ejemplo, moléculas de ácido nucleico sintetizadas con ayuda de enzimas *in vivo* o *in vitro*, tales como moléculas de ADNc, moléculas de polinucleótidos sintetizadas por PCR, moléculas de ARN sintetizadas por transcripción *in vitro*, etc. Los métodos de la PCR son bien conocidos por un experto en la técnica y se describen, por ejemplo, en Innis y col., compiladores, 1990, *PCR Protocols: A Guide to Methods and Applications*, Academic Press Inc., San Diego, CA. La muestra de moléculas diana puede contener, por ejemplo, moléculas de ADN, ARN o copolímeros de ADN y ARN. Las moléculas diana descritas en esta memoria, se pueden corresponder a genes específicos, a transcritos de genes específicos o a fragmentos específicos de un transcrito de gen (por ejemplo, las secuencias de ciertos ARNm expresados en células o secuencias de ciertos ADNc derivadas de estas secuencias de ARNm).

Las moléculas diana que se van a analizar también se pueden preparar *in vitro* a partir de ácidos nucleicos extraídos de células. Por ejemplo, el ARN se puede extraer de células (particularmente el ARN celular total, la fracción de poli(A)⁺ del ARNm o fracciones de estos últimos) y el ARN mensajero purificado a partir del extracto de ARN total. Los métodos para preparar el ARN total y el ARN poli(A)⁺ son bien conocidos en la técnica anterior y se describen de forma general en Sambrook y col., compiladores, 1989, *Molecular Cloning: A Laboratory Manual*, vols. 1-3, 2ª ed. Cold Spring Harbor Laboratory, Cold Spring Harbor, Nueva York. El ARN se puede extraer de células que tienen diferentes tipos de interés, mediante la lisis de las células con tiocianato de guanidinio, seguido de centrifugación en presencia de CsCl, y la purificación por afinidad con oligo dT (Chirgwin y col., 1979, *Biochemistry* 18:5294-5299). El ARN también se puede extraer de las células mediante una lisis con tiocianato de guanidinio, seguido de purificación en columnas Rneasy (Qiagen). Los ADNc se pueden sintetizar entonces a partir del ARNm purificado, utilizando oligo dT, por ejemplo, o cebadores aleatorios. Las moléculas diana pueden ser ARNc preparados a partir de ARN

mensajeros purificados o ARN total extraído de células. El ARNc descrito en esta memoria se refiere a los ARN complementarios del ARN de la fuente. Los ARN extraídos se pueden amplificar a continuación usando un proceso en el que cadenas dobles de ADNc se sintetizan a partir de los ARN usando un cebador unido a un promotor de la ARN polimerasa en una dirección, para dirigir la transcripción del ARN no codificante. Los ARN no codificantes o ARNsc se pueden transcribir entonces a partir de la segunda cadena de los ADNsc, utilizando una polimerasa de ARN (véase, por ejemplo, los documentos de patente de EE.UU. n° 5.891.636; 5.716.785; 5.545.522 y 6.132.997). Los cebadores de oligo-dT que contienen un promotor de la ARN polimerasa, se pueden utilizar. El ARN total se puede utilizar como material de partida para la síntesis de ARNc. Un cebador de oligo-dT que contiene una secuencia del promotor de la polimerasa T7, se puede utilizar para iniciar la síntesis de la primera hebra del ADNc, y los cebadores hexámeros aleatorios se pueden utilizar para iniciar la síntesis de la segunda hebra de ADNc con la transcriptasa inversa MMLV. Esta reacción produce un ADNc de doble cadena que contiene el promotor de la polimerasa de T7 en el extremo 3'. Los ADNc de doble cadena se pueden traducir a continuación en ARNc con la polimerasa de ARN de T7. La concentración de la diana sintetizada puede variar de 1 µM a 50 nM. La concentración de las dianas sintetizadas con la PCR puede variar de 5 ng/µL a 50 ng/µL.

Las moléculas diana pueden ser cebadores de PCR o estar incluidas en los cebadores de PCR. Por lo tanto, las sondas descritas en esta memoria pueden permitir la detección de un producto de la amplificación con PCR. Más particularmente, esta detección es un excelente testigo negativo para la PCR. Estos testigos negativos se realizan mediante la amplificación de una secuencia en ausencia de ADN genómico, pero en presencia de los reactivos necesarios para la PCR y una pareja de cebadores de los cuales al menos uno es portador de una secuencia perfectamente complementaria de la secuencia específica de la diana de la sonda de horquilla.

Las moléculas diana que se van a analizar por los métodos y composiciones descritos en esta memoria, se pueden marcar opcionalmente de modo que se puedan detectar. Por ejemplo, los ADNc se pueden marcar directamente, es decir, mediante el uso de análogos de nucleótidos, o indirectamente, por ejemplo, mediante la preparación de una segunda hebra de ADNc marcada utilizando la primera hebra como molde. Alternativamente, el ADNc de doble cadena se puede transcribir en ARNc y marcarlo después de su transcripción. El marcador de detección puede ser un marcador fluorescente que actúa, por ejemplo, incorporando análogos de nucleótidos. Otros marcadores que se pueden utilizar en esta memoria pueden incluir en particular biotina, iminobiotina, antígenos, cofactores, dinitrofenol, ácido lipoico, compuestos olefinicos, polipéptidos detectables, moléculas ricas en electrones, enzimas capaces de generar una señal detectable actuando en presencia de un soporte e isótopos radiactivos. Los isótopos radiactivos favoritos incluyen en particular ³²P, ³⁵S, ¹⁴C, ¹⁵N e ¹²⁵I. Las moléculas fluorescentes que se pueden utilizar en esta memoria incluyen preferentemente la fluoresceína y sus derivados, la rodamina y sus derivados, el rojo Texas, la 5' carboxi-fluoresceína ("FAM"), la 2',7'-dimetoxi-4',5'-dicloro-6-carboxi-fluoresceína ("JOE"), la N,N,N',N'-tetrametil-6-carboxi-rodamina ("TAMRA"), 6'-carboxi-X-rodamina ("ROX"), HEX, TET, IRD40 y IRD41. Otras moléculas adicionales fluorescentes, adaptadas y utilizables en esta memoria, incluyen además: cianinas (incluyendo especialmente, Cy3, Cy3.5 y Cy5), moléculas BODIPY (incluyendo especialmente BODIPY-FL, BODIPY-TM, BODIPY-630/650 y BODIPY-650/670) y las moléculas fluorescentes de la serie ALEXA (incluyendo especialmente, ALEXA-488, ALEXA-532, ALEXA-546, ALEXA-568 y ALEXA-594), así como cualquier otra molécula fluorescente conocida por una persona experta en la técnica. Las moléculas indicadoras ricas en electrones que se pueden utilizar en esta memoria, incluyen especialmente ferritina, hemocianina y/o oro coloidal. Alternativamente, las moléculas diana se pueden marcar formando complejos de forma específica con un primer grupo sobre las moléculas diana. Un segundo grupo, unido covalentemente a una molécula indicadora, que tiene afinidad hacia el primer grupo, se puede utilizar para detectar las moléculas diana de forma indirecta. En tal configuración, los compuestos utilizables en el primer grupo incluyen especialmente biotina e iminobiotina. Los compuestos utilizables como segundo grupo incluyen especialmente avidina y estreptavidina. Las moléculas fluorescentes que se pueden intercalar en la molécula de ADN, como el bromuro de etidio o SYBR Green I, también se pueden utilizar para controlar el aumento de la señal fluorescente después de la hibridación de una diana polinucleotídica con una sonda de horquilla, o la hibridación de la molécula "informadora" en la sonda de horquilla (ver más abajo). Este aspecto puede ser particularmente útil para el análisis de la hibridación en tiempo real.

Sin embargo, de acuerdo con una realización preferida del análisis de acuerdo con la invención, las moléculas diana no están marcadas, y los marcadores detectables descritos anteriormente están unidos a una molécula "informadora" (véase más abajo).

Métodos de hibridación utilizables con las sondas de horquilla

Las sondas de horquilla descritas en esta memoria están diseñadas para adoptar durante la detección de las moléculas diana, dos conformaciones alternativas. Estas dos conformaciones, la conformación "cerrada" y la conformación "abierta" se describen en la Fig. 2. Tanto la Fig. 2A como Fig. 2B muestran una sonda de horquilla fijada sobre un soporte. La Fig. 2A muestra una sonda de horquilla en su conformación cerrada. En ausencia de la diana específica de una sonda de horquilla, los brazos de la sonda se hibridan entre sí para formar un "vástago" 201, mientras que la secuencia específica de la diana forma un "bucle" 202. La sonda no puede adoptar otra conformación, porque los brazos de la sonda de horquilla están diseñados para hibridarse únicamente entre sí. La temperatura fija preferida para detectar una molécula diana con una sonda de horquilla es 5 a 10°C menor que la del híbrido formado entre la secuencia específica de la diana y la molécula diana. Esta temperatura puede variar

dependiendo de la longitud, la posición y la presencia o ausencia de mutaciones adicionales en la estructura del bucle. Una secuencia específica de la diana más larga o más corta, así como un vástago más largo o más corto, también se pueden utilizar para modificar la temperatura óptima de la hibridación. Durante la asociación con la molécula diana, la sonda de horquilla sufre un cambio de la conformación estructural. La asociación de una porción de la secuencia del bucle con su molécula diana (es decir, la hibridación de la secuencia específica de la diana con su secuencia diana complementaria de ácido nucleico) obliga a la sonda de horquilla a abrirse.

La competencia entre la hibridación de los brazos complementarios entre sí y la secuencia específica de la diana con su molécula diana, es un criterio clave para distinguir entre una asociación perfecta entre la secuencia específica de la diana y su molécula diana, y una asociación imperfecta con otras moléculas diana. Cuando una sonda de horquilla se pone en presencia de la molécula diana para la que se ha diseñado (es decir, que coincide perfectamente), la interacción entre la secuencia específica de la diana y la molécula diana conduce a la disociación del vástago, y a la aparición de la conformación "abierta", de la sonda. La Fig. 2B muestra una sonda de horquilla en una conformación "abierta" resultante de la asociación con su molécula diana 203. La secuencia específica de la diana 204 se hibrida con una parte 205 de la molécula diana 203. Además está previsto que una interacción entre una sonda de horquilla y una secuencia diana se pueda analizar con un bucle completamente hibridado con la molécula diana, o una porción de la misma. Esta hibridación entre la secuencia específica de la diana 204 y la molécula diana 203 no tiene que incluir necesariamente toda la secuencia de la molécula diana 203. Cuando la secuencia específica de la diana 204 se hibrida con una porción 205 de la molécula diana 203, los brazos de la sonda, el brazo 206 del lado 5' de la secuencia específica de la diana, y el brazo 207 del lado 3' de la secuencia específica de la diana, se disocian y están disponibles para interactuar con las moléculas "informadoras".

Durante la hibridación, la afinidad de la asociación de los ácidos nucleicos diana con las secuencias de las sondas, depende tanto de la similitud de las diferentes secuencias diana en la muestra, como de las condiciones de rigurosidad de la hibridación, en particular, la temperatura de la hibridación, la concentración de sales y la presencia de agentes químicos que reducen la afinidad. La cinética de la asociación también depende de las concentraciones relativas de los diferentes ácidos nucleicos presentes en la muestra. Para las sondas de polinucleótidos que se dirigen a (es decir, complementarias de) especies de poca abundancia o que se dirigen a ácidos nucleicos con un alto grado de similitud de secuencias (es decir, homólogos), las reacciones de hibridación cruzada pueden tener un impacto significativo en términos de contaminación, y distorsionar los resultados de la medición de la hibridación. Detección de SNP es un ejemplo en el que las hibridaciones cruzadas son un problema, ya que la secuencia que se va a detectar (más específicamente, el SNP que se va a analizar) se tiene que diferenciar de otras secuencias que solo tienen un solo de nucleótido de diferencia.

Varios enfoques de la técnica anterior se han diseñado para reducir las hibridaciones cruzadas y se pueden utilizar en el marco de los métodos descritos en esta memoria. Las hibridaciones cruzadas se pueden reducir regulando a la vez las condiciones de rigurosidad durante la hibridación y/o durante los lavados efectuados después de la hibridación. Por ejemplo, unas condiciones de lavado muy rigurosas se pueden utilizar para desestabilizar todos los híbridos formados, excepto los dúplex más estables, de modo que las señales detectadas se corresponden sólo a los híbridos perfectos entre la secuencia específica de una sonda de horquilla y su secuencia específica. Un ejemplo de condiciones muy rigurosas es una hibridación del ADN en un tampón SSC 5X que contiene 1% de dodecil sulfato sódico (SDS), EDTA 1 mM a 65°C, y un lavado en un tampón SSC 0,1X, que contiene 0,1% de SDS (Ausubel y col., compiladores, 1989, *Current Protocols in Molecular Biology*, vol. I, Green Publishing Associates, Inc., y John Wiley & Sons, Inc., Nueva York, NY, páginas 2.10.3). Estos lavados se pueden realizar 5-10°C por debajo de la T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana. Las condiciones muy rigurosas permiten detectar variantes alélicas de una secuencia de nucleótidos, por ejemplo, un desemparejamiento cada 10-30 nucleótidos. Alternativamente, unas condiciones de lavado "moderadas" o "poco rigurosas" también se pueden utilizar. Estas condiciones moderadas o poco rigurosas son también bien conocidas en la técnica anterior (véase, por ejemplo, Sambrook J. y col., compiladores, 1989, *Molecular Cloning: A Laboratory Manual*, 2ª ed., Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY, páginas 9.47-9.51 y 11.55-11.61; Ausubel y col., compiladores, 1989, *Current Protocols in Molecular Biology*, vol. I, Green Publishing Associates, Inc., y John Wiley & Sons, Inc., Nueva York, páginas 2.10.1-2.10.16). Ejemplos de condiciones de lavado de rigurosidad moderada incluyen, por ejemplo, un lavado en tampón SSC 0,2X que contiene 0,1% de SDS a 42°C (Ausubel y col., 1989, más arriba). Ejemplos de condiciones de lavado poco rigurosas, por ejemplo, incluyen un lavado en tampón SSC 5X o 0,2X, que contiene 0,1% de SDS, o SSC 5X, a temperatura ambiente (Ausubel y col., 1989, más arriba).

Las contribuciones de las hibridaciones cruzadas a las señales de las sondas de horquilla, se pueden rastrear y eliminar sustrayendo las señales de las sondas de referencia apropiadas. La contribución de las hibridaciones cruzadas se puede medir, por ejemplo, usando sondas de horquilla que no se hibridan con ninguna de las moléculas diana presentes en el análisis. La señal no específica así obtenida se puede sustraer de la señal de hibridación específica.

Según la descripción, las hibridaciones cruzadas también se pueden reducir mediante el uso de sondas de horquilla que tienen pequeñas diferencias entre ellas de la T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana, es decir, unas diferencias inferiores o iguales a 1°C. Por el contrario, cuando las diferencias en la T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana aumentan, es decir, son

superiores a 1°C, las hibridaciones cruzadas se vuelven un factor más significativo. En ese caso, se pueden utilizar enfoques similares a los mencionados anteriormente, el número de sondas de horquilla utilizado para detectar cada molécula diana se puede aumentar o se puede utilizar una combinación de estos enfoques.

- Se pueden utilizar una o varias secuencias únicas testigo y la temperatura se puede determinar mediante el análisis de qué sonda de horquilla se hibrida con la secuencia testigo. Una secuencia testigo puede ser complementaria de 5-10 sondas de horquilla diferentes, que difieren entre sí por la longitud de la secuencia específica de la diana y, por lo tanto, por su T_m. Por ejemplo, una sonda de horquilla puede ser complementaria de ocho nucleótidos situados en medio de la secuencia testigo, y tener una T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana, de 40°C. Una segunda sonda de horquilla puede ser complementaria de doce nucleótidos de la secuencia testigo, y tener una T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana, de 45°C. Otras sondas de horquilla pueden ser complementarias de otras regiones de la secuencia testigo y tener unas T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, diferentes. Para un usuario es posible determinar la temperatura del ensayo basándose en la identificación de la sonda de horquilla que se hibrida con la secuencia testigo.
- La descripción describe también un biochip de sondas tales como las descritas en esta memoria, caracterizado porque todos los híbridos perfectos formados entre las secuencias específicas de la diana de las sondas y las moléculas diana, tienen una temperatura de fusión igual, dentro de un intervalo de 4°C.
- También se describe un biochip de sondas tales como las descritas en esta memoria, caracterizado porque cada híbrido formado entre la secuencia específica de la diana de cada sonda y la molécula diana, tiene una temperatura de fusión igual, dentro de un intervalo de 1°C, de las temperaturas de fusión de los híbridos formados entre la secuencia específica de la diana y la molécula diana de todas las otras sondas.
- Adicionalmente se describe un biochip de sondas tales como las descritas en esta memoria, caracterizado porque la diferencia entre cada temperatura de fusión del híbrido, formado entre la secuencia específica de la diana de cada sonda y la molécula diana, y una segunda temperatura de fusión de la asociación entre la secuencia específica de la diana y una molécula para la que la no se ha diseñado la secuencia específica de la diana, es mayor o igual a 5°C. La diferencia entre cada temperatura de fusión del híbrido formado entre la secuencia específica de la diana de cada sonda y la secuencia diana, y la segunda temperatura de fusión, es claramente mayor o igual a 8°C.

Moléculas "informadoras"

- La presente invención también comprende el uso de una o varias moléculas "informadoras", sin inmovilizar, marcadas o no, tales como secuencias de ácido nucleico que tienen una longitud igual a la del brazo de una sonda. La invención describe por tanto un biochip con sondas de horquilla, caracterizado porque una molécula "informadora" se puede hibridar con cada sonda de horquilla. Una sonda de horquilla descrita en esta memoria, se puede usar junto con moléculas "informadoras" que solo se pueden hibridar con la forma abierta de la sonda, lo que permite una diferenciación entre las conformaciones abiertas y cerradas, y que genera una señal sólo cuando una molécula diana se asocia con la secuencia específica de la diana de una sonda de horquilla.
- Las moléculas "informadoras" se pueden diseñar para que tengan una estructura de vástago-bucle con un vástago de 4-10 nucleótidos, preferiblemente de 6 nucleótidos o menos, y un bucle que comprende entre 12-20 nucleótidos, en el que la secuencia específica de la diana puede ser perfectamente complementaria de un brazo de la sonda de horquilla.
- Las moléculas "informadoras" se pueden diseñar para que tengan una estructura lineal que comprende una secuencia perfectamente complementaria de un brazo de la sonda de horquilla.
- La invención describe por tanto una sonda de horquilla sin marcar, caracterizada porque una molécula diana se hibrida con la secuencia específica de la diana y porque una molécula "informadora" se hibrida con el primer o el segundo brazo. La molécula "informadora" comprende preferiblemente menos de 10 nucleótidos perfectamente complementarios con la secuencia de ácido nucleico del primer brazo o del segundo brazo.
- Las moléculas "informadoras" se pueden diseñar con una molécula que se fija en el surco menor del ADN ("minor groove binder") fijada a un extremo de la molécula "informadora". Las moléculas "informadoras" se marcan de la misma manera que las secuencias diana (véase más arriba), con un marcador fluorescente en el extremo 3' y/o 5'.
- Una vez que la muestra se ha puesto en contacto con una sonda de horquilla, o un conjunto de sondas de horquilla, el brazo de una sonda puede ser accesible para una molécula "informadora" complementaria, debido a la asociación del bucle con su secuencia diana, y entonces se puede emitir una señal. La molécula "informadora" puede ser en particular una secuencia de ácido nucleico, una secuencia de péptido-ácido nucleico, un ácido nucleico bloqueado ("locked nucleic acid, LNA"), una secuencia oligopeptídica, una proteína o una enzima acoplada a una secuencia de ácido nucleico. La molécula "informadora" se caracteriza preferentemente por su capacidad para unirse a un brazo de la sonda descrita en esta memoria, únicamente con la conformación "abierta". La molécula "informadora" contiene preferiblemente al menos una secuencia perfectamente complementaria de una de las secuencias de los

brazos de una sonda de horquilla. Durante la adición de una molécula "informadora", la asociación entre la molécula "informadora" y una sonda de horquilla se produce sólo con un brazo de la sonda que está en su conformación abierta. Esta estructura abierta es una consecuencia de la asociación del bucle de la sonda de horquilla con una molécula diana. Para las sondas de horquilla inmovilizadas sobre el mismo soporte, las sondas pueden estar diseñadas para tener la misma secuencia del brazo, de modo que solo sea necesaria una molécula "informadora" para analizar la asociación de todas las sondas de horquilla con sus moléculas diana específicas.

Las Fig. 3 muestran cómo se realiza la asociación de una molécula "informadora" con una sonda descrita en esta memoria, para generar una señal durante la hibridación de la sonda de horquilla con su molécula diana. La Fig. 3A muestra las etapas de la asociación de una molécula "informadora" con una sonda de horquilla. En primer lugar, la sonda 301 está en su conformación cerrada, en la que el brazo 302 situado en el extremo 5' de la secuencia específica de la diana y el brazo 303, situado en el extremo 3' de la secuencia específica de la diana, se hibridan, formando el vástago 304. En presencia 306 de la molécula diana 305 (es decir, un ácido nucleico) para la que ha sido diseñada la sonda de horquilla, la secuencia específica de la diana 307 se hibrida con la molécula diana 305. Esto hace que la sonda de horquilla adopte una conformación abierta en la que los brazos 302 y 303 de la sonda se disocian entre sí y se vuelven accesibles para la molécula "informadora". Al añadir 308 de la molécula "informadora" 309, que comprende un ácido nucleico, a la sonda descrita en esta memoria en la conformación abierta, la molécula "informadora" 309 se puede hibridar con el brazo 302 situado en el extremo 5' de la secuencia específica de la diana, ya que una porción de la molécula "informadora" 309 ha sido diseñada para ser el complemento del brazo 302 situado en el extremo 5' de la secuencia específica de la diana.

Puesto que cada sonda de horquilla tiene dos brazos, dos moléculas "Informadoras" se pueden diseñar y utilizar para generar una señal cuando una sonda de horquilla se hibrida con una secuencia diana. La Fig. 3B describe una sonda tal como la descrita en esta memoria, con la conformación abierta en la que los dos brazos 310 y 311 de la sonda de horquilla se hibridan con dos moléculas "informadoras" 312 y 313. Cuando se utilizan dos moléculas "informadoras" y una molécula "informadora" está diseñada para contener una secuencia de ácido nucleico complementaria de un brazo de la sonda de horquilla, y la otra molécula "informadora" está diseñada para contener una secuencia de ácido nucleico complementaria del otro brazo de la sonda de horquilla, puede ser necesario añadir las moléculas "informadoras" una después de la otra, intercalando una etapa de lavado entre las etapas de adición de las moléculas "informadoras". Esto es debido a que cada molécula "informadora" contiene una secuencia de ácido nucleico perfectamente complementaria de la de la otra molécula "informadora". Si las dos moléculas "informadoras" se añadieran al mismo tiempo, estas moléculas "informadoras" podrían hibridarse entre sí, así como con la sonda de horquilla.

La molécula "informadora" comprende un marcador detectable. Preferiblemente, el marcador detectable se selecciona entre un análogo de nucleótido, un marcador fluorescente, biotina, iminobiotina, un antígeno, un cofactor, dinitrofenol, ácido lipoico, un compuesto olefínico, un polipéptido, una molécula rica en electrones, una enzima o un isótopo radioactivo. Las moléculas "informadoras" pueden ser secuencias marcadas, por ejemplo, con un fluorocromo (referencia F en las Figs. 3A, 3B, 5B y 5C) o secuencias sin marcar con modificaciones adicionales que permiten aumentar la afinidad para su asociación con los brazos a los que son complementarios. Por ejemplo, esta modificación puede ser una molécula que se fija en el surco menor del ADN ("DNA minor groove binder") fijada al extremo 3' o 5' del "informador", o una modificación química del "informador" que permite el establecimiento de un enlace covalente entre las hebras de ADN hibridadas después de la aplicación de un tratamiento específico. Estas modificaciones también pueden incluir la incorporación de un grupo fosforotioato en la cadena fosfato de la molécula "informadora" (Cogoi, S. y col. 2001 *Biochemistry* 40:1135-43; Xodo, L. y col. 1994 *Nucleic Acids Res.* 22:3322-30).

Las moléculas "informadoras" también se pueden detectar por su masa, utilizando la espectrometría de masas y el análisis de tiempo de vuelo. Por ejemplo, las sondas descritas en esta memoria, inmovilizadas sobre un soporte e hibridadas con sus secuencias diana y las moléculas "informadoras" se pueden analizar directamente mediante espectrometría de masas, orientando la fuente de láser del espectrómetro de masas sobre una región definida de la superficie, correspondiente a una sonda específica, y recogiendo la masa y la cantidad o la abundancia relativa de cada molécula "informadora". La molécula "informadora" puede variar en longitud y en masa, o puede estar unida a varias especies moleculares de masa conocida, detectable por espectrometría de masas (Laken, S.J. y col. 1998. *Nat. Biotechnol.* 16:352-6; Little, D.P. y col., 1997. *Nat. Med.* 3:1413-6, Little, D.P. 1997. *Eur. J. Clin. Chem. Clin. Biochem.* 35:545-8; Braun, A. y col. 1997. *Clin. Chem.* 43:1151-8).

La molécula "informadora" puede estar acoplada a moléculas electroquímicamente activas, tales como los derivados de ferroceno o de cobaltoceno que generan una corriente eléctrica de óxido-reducción cuando se les aplicaba una corriente (Umek, R.M. y col. 2001. *J. Mol. Diagn.* 3:74-84; Padeste, C. y col. 2000 *Biosens. Bioelectron.* 15:431-8; Tsai, W.C. y col. 1995 *Analyst* 120:2249-54). Las moléculas "informadoras" también pueden ser una enzima capaz de generar especies electroactivas, por actividad electrocatalítica, o por escisión de un sustrato electroquímicamente inactivo en un producto electroquímicamente activo. (Valat, C. y col. 2000 *Analytica Chimica Acta* 404:187-94; Oliver, B. y col. 1997. *Anal. Chem.* 69:4688-94; Limoges, B. 1996 *Anal. Chem.* 68:4141-48; Bourdillon, C. y col. 1996. *J. Am. Chem. Soc.* 115:12264-69). Para tales moléculas "informadoras", las sondas descritas en esta memoria se pueden inmovilizar sobre superficies conductoras, tales como dióxido de silicio, grafito, carbono vítreo, superficies de óxido de selenio-indio, superficies metálicas o superficies conductoras a base de polímeros plásticos, o superficies

no conductoras que tienen zonas individualizadas de materiales conductores polimerizados o depositados previamente. Este método de detección puede usar la electropolimerización del detector o de una parte del mismo en la superficie, como medio para detectar una señal en forma de una corriente o de una variación de potencial.

Otro sistema "informador" para detectar la asociación de una sonda de horquilla con una molécula diana, utiliza un sistema de marcado con una polimerasa. La Fig. 4 describe el principio del funcionamiento de este sistema. La figura 4A muestra una sonda de horquilla en su conformación cerrada, prolongada en su extremo 5' por una secuencia de 10 a 20 nucleótidos, el "cebador" 401. Los cebadores de todas las sondas de horquilla que se deben utilizar en una condición de ensayo específico, se diseñan preferiblemente para que tengan la misma secuencia de nucleótidos. La descripción también describe el uso de diferentes cebadores en una condición de ensayo específico para añadir un nivel adicional de diferenciación entre las señales generadas por la hibridación de una sonda de horquilla con su molécula diana específica. Para cada cebador utilizado, la secuencia nucleica del cebador 401 está controlada para garantizar la ausencia de homología de secuencia con todas las otras sondas de horquilla, y con los otros reactivos de la polimerasa. Preferiblemente, el cebador tiene una longitud de 18 a 20 nucleótidos y una temperatura de fusión de 59°C +/- 2°C. El cebador 401 se utiliza para iniciar la incorporación de al menos un nucleótido marcado, con una polimerasa. Durante la asociación de la secuencia específica de la diana con la molécula diana, el brazo 5' de la secuencia específica de la diana y el cebador son accesibles a los reactivos de la reacción de polimerización. La reacción de polimerización se lleva a cabo sin la desnaturalización del ADN y a 37-60°C, dependiendo de la polimerasa utilizada. Ejemplos de polimerasas utilizables para este sistema de marcado es la polimerasa I de *E. coli*, el fragmento Klenow, las polimerasas termoestables Taq, la polimerasa de ADN de T4, y todas las polimerasas naturales o modificadas genéticamente, con la excepción de las polimerasas "hot-start" (Sambrook J. y col., 1989, *Molecular Cloning: A Laboratory Manual*, 2ª Ed., Cold Spring Harbor Laboratory, Cold Spring Harbor, NY, páginas 5.44-5.47; Ausubel F.M. y col. 1997, *Current Protocols in Molecular Biology*, vol. I, Green Publishing Associates, Inc., y John Wiley & Sons, Inc., Nueva York, páginas 3.5.11 y 3.5.12).

La Fig. 4B muestra una sonda tal de horquilla en su conformación abierta, en la que el brazo 402 y el cebador 401 situados en el extremo 5' de la secuencia específica de la diana, son accesibles para participar en la reacción con la polimerasa. Una molécula "informadora" 403, complementaria de una porción del brazo 402 y del cebador 401, se puede hibridar entonces con las secuencias. La Fig. 4C muestra una hibridación tal entre un brazo 402, un cebador 401 y una molécula "informadora" 403. La molécula "informadora" 403 puede ser complementaria de una porción o de la totalidad del brazo 402. Sin embargo, la molécula "informadora" 403 no puede ser complementaria de la totalidad del cebador 401. La porción de la molécula "informadora" 403 que completa la porción del brazo 402 y del cebador 401, no se debe asociar o tener nucleótidos que sean complementarios de una porción del extremo 3' del cebador 401. Preferiblemente, el tamaño de la secuencia de nucleótidos de la molécula "informadora" 403 es 1 a 5 nucleótidos más corta que la del cebador 401 con la que se hibrida. Esta característica está representada por la secuencia no hibridada 404 del cebador 401, y permite que una polimerasa prolongue la molécula "informadora" 403, mediante la adición de nucleótidos complementarios del cebador 401. Cuando una polimerasa 405 prolonga la molécula "informadora" 403 (véase la Fig. 4D), los nucleótidos marcados 406 se pueden incorporar para generar una señal que permita detectar la hibridación de la secuencia específica de la diana 407 con la molécula diana 408.

Sistema de direccionamiento universal

La presente descripción también describe un sistema en el que las sondas de horquilla no inmovilizadas se pueden usar junto con sondas de horquilla inmovilizadas. Designado por la expresión "sistema de direccionamiento universal, o SAU", este sistema de detección ex vivo tiene la ventaja de aplicar una batería de sondas de horquilla o lineales inmovilizadas, prefabricadas y utilizables con las sondas de horquilla no inmovilizadas, suministradas por el usuario final, lo que le permite evitar la fabricación del biochip. El SAU puede, por ejemplo, proporcionar un biochip universal en donde las secuencias específicas de las dianas de horquilla inmovilizadas, comprenden un conjunto de sondas de horquilla predeterminadas, diseñadas para asociarse con las sondas de horquilla o lineales no inmovilizadas que están marcadas de forma específica. Esto permite la construcción de un biochip único para todas las condiciones del análisis. Los ácidos nucleicos diana que se van a detectar, en primer lugar son capturados con las secuencias específicas de las dianas de las sondas de horquilla o lineales, no inmovilizadas, que están diseñadas, a su vez, para asociarse, a través de su secuencia marcadora específica, con las sondas de horquilla inmovilizadas.

En una configuración preferida, el SAU comprende dos conjuntos de sondas de horquilla. Una sonda de horquilla, designada por la expresión "primera sonda de horquilla" no está inmovilizada sobre un soporte y tiene una secuencia marcadora en el extremo de uno de sus brazos 3' o 5'. El marcador se puede hibridar con una secuencia de ácido nucleico monocatenaria de 6 a 30 nucleótidos, o con una sonda de horquilla denominada "segunda sonda de horquilla", que está inmovilizada sobre un soporte y puede ser parcial o totalmente complementaria del marcador. Cuando la especie inmovilizada es una segunda sonda de horquilla, la secuencia específica de la diana de la segunda sonda puede ser complementaria de la secuencia marcadora, de una porción de la secuencia marcadora y/o de una porción del brazo adyacente al marcador. El uso de segundas sondas que tienen secuencias específicas de la diana, dirigidas contra una porción del marcador y una porción del brazo adyacente al marcador, permite que solo las primeras sondas que tienen su secuencia específica de la diana asociada con una molécula diana, sean capaces de hibridarse con la segunda sonda de horquilla. Si se utilizan múltiples secuencias marcadoras con el

SAU, y se tienen que diferenciar, una diferencia de al menos 2 nucleótidos situada en el centro del bucle de la segunda sonda puede ser necesaria, para que haya suficientes diferencias con las otras secuencias marcadoras para hibridar específicamente la secuencia específica de la diana complementaria en las sondas de horquilla inmovilizadas.

5 La descripción describe así un sistema de direccionamiento universal sin marcar, que comprende:

(a) al menos dos primeras sondas sin marcar lineales que comprenden:

(i) una secuencia específica de una diana, de 6 a 30 nucleótidos de longitud;

(ii) una secuencia denominada marcadora de 10 a 50 nucleótidos de longitud, conectada al extremo 5' de la secuencia específica de la diana,

10 en donde dicha secuencia específica de una diana no es complementaria de ninguna otra parte de dicha sonda sin marcar; y

(b) al menos dos segundas sondas capaces de formar una horquilla que comprende:

(i) una secuencia específica de una segunda diana, de 6 a 30 nucleótidos de longitud;

15 (ii) un segundo primer brazo de dicha segunda sonda de menos de 10 nucleótidos de longitud y situado en 5' de dicha segunda secuencia específica de dicha diana;

(iii) un segundo segundo brazo de dicha segunda sonda de menos de 10 nucleótidos de longitud y situado en 3' de dicha segunda secuencia específica de dicha diana; y

(iv) un espaciador que conecta el primer o el segundo brazo con dicho soporte,

20 en donde dicha secuencia específica de dicha diana de la segunda sonda no es complementaria de ninguna otra parte de dicha segunda sonda; en donde dicho primer brazo y dicho segundo brazo de dicha segunda sonda son perfectamente complementarios entre sí; y además en donde dicha secuencia marcadora de una de las primeras sondas, es complementaria de dicha segunda secuencia diana de una de las segundas sondas.

Preferiblemente, la descripción describe así un sistema de direccionamiento universal sin marcar que comprende:

(a) al menos dos primeras sondas sin marcar capaces de formar una horquilla que comprende:

25 (i) una secuencia específica de una diana, de 6 a 30 nucleótidos de longitud;

(ii) un primer brazo de menos de 10 nucleótidos de longitud, situado en 5' de dicha secuencia específica de una diana;

(iii) un segundo brazo de menos de 10 nucleótidos de longitud, situado en 3' de dicha secuencia específica de una diana; y

30 (iv) una secuencia denominada marcadora de 10 a 50 nucleótidos de longitud, conectada con el primer brazo o el segundo brazo,

en donde dicha secuencia específica de una diana no es complementaria de ninguna otra parte de dicha sonda sin marcar; y además, en donde el primer brazo y el segundo brazo son perfectamente complementarios entre sí, y

(b) al menos dos segundas sondas capaces de formar una horquilla que comprende:

35 (i) una secuencia específica de una segunda diana, de 6 a 30 nucleótidos de longitud;

(ii) un segundo primer brazo de dicha segunda sonda de menos de 10 nucleótidos de longitud y situado en 5' de dicha segunda secuencia específica de dicha diana;

(iii) un segundo segundo brazo de dicha segunda sonda de menos de 10 nucleótidos de longitud y situado en 3' de dicha segunda secuencia específica de dicha diana; y

40 (iv) un espaciador que conecta el primer o el segundo brazo a dicho soporte,

en donde dicha secuencia específica de dicha diana de la segunda sonda no es complementaria de ninguna otra parte de dicha segunda sonda; en donde dicho primer brazo y dicho segundo brazo de dicha segunda sonda son perfectamente complementarios entre sí; y además en donde dicha secuencia marcadora de una de las primeras sondas completada de la secuencia del brazo conectado a este marcador, es complementaria de dicha segunda secuencia diana de una de las segundas sondas.

45

En una realización preferida, la secuencia específica de la diana de las primeras sondas de horquilla, tiene una longitud de 10 a 25 nucleótidos, más preferiblemente de 15 a 20 nucleótidos.

La secuencia específica de la diana de las segundas sondas de horquilla tiene preferiblemente una longitud de 10 a 25 nucleótidos, más preferiblemente de 15 a 20 nucleótidos.

- 5 La descripción también describe un sistema de direccionamiento universal sin marcar, que se caracteriza porque todos los primeros brazos tienen una secuencia idéntica.

También describe un sistema de direccionamiento universal sin marcar, caracterizado porque todos los primeros brazos de la segunda sonda de horquilla, tienen una secuencia idéntica.

- 10 La descripción también describe un sistema de direccionamiento universal sin marcar, que se caracteriza porque una misma molécula "informadora" se puede hibridar con una u otra de las sondas primera y segunda descritas en esta memoria.

El primer brazo de la segunda sonda descrita en esta memoria y el segundo brazo de la segunda sonda descrita en esta memoria, se pueden disociar entre sí, cuando una porción del marcador y una porción del primer o del segundo brazo conectado al marcador, se asocian con la secuencia específica de la diana de la segunda sonda.

- 15 En particular la descripción describe un sistema de direccionamiento universal sin marcar, caracterizado porque el marcador se asocia con la secuencia específica de la diana de la segunda sonda.

- La Fig. 5 describe el funcionamiento de una de las configuraciones del SAU. En esta configuración, la secuencia específica de la diana de la segunda sonda de horquilla, está diseñada para hibridarse con el marcador y una porción del brazo adyacente del marcador de la primera sonda de horquilla. La Fig. 5A muestra el SAU 501, en ausencia de molécula diana. La primera sonda de horquilla 502, que comprende un marcador 503, y la segunda sonda de horquilla 504, fijada al soporte 505 por un espaciador 506, están en su conformación cerrada. La Fig. 5B muestra cómo la primera sonda 502, adopta una conformación abierta durante la asociación con su diana 507, y una molécula "informadora" 508 marcada. La Fig. 5C muestra la primera sonda de horquilla 502, con su marcador 503, y una porción del brazo 509 adyacente al marcador hibridado con la secuencia específica de la diana 510 de la segunda sonda de horquilla 504. La posición de la molécula diana se puede determinar mediante la detección de la posición de la molécula "informadora" sobre el soporte.

Detección de ácidos nucleicos

La descripción describe especialmente un método para la detección de ácido nucleico que comprende:

- 30 (a) poner en contacto *ex vivo* una muestra de ácido nucleico con un biochip que comprende al menos dos sondas descritas en esta memoria; y

(b) detectar una señal que proviene de al menos una de dichas sondas del biochip que ha adoptado una conformación abierta, después del contacto establecido durante la etapa (a).

- La detección de la etapa (b) se obtiene preferiblemente mediante el uso de una molécula "informadora" marcada con un marcador detectable. El marcador detectable se selecciona preferiblemente entre un análogo de nucleótido, un marcador fluorescente, biotina, iminobiotina, un antígeno, un cofactor, dinitrofenol, ácido lipoico, un compuesto olefínico, un polipéptido, una molécula rica en electrones, una enzima o un radioisótopo.

Según una realización preferida, la muestra de ácido nucleico empleada en la etapa (a), está marcada previamente con un marcador detectable.

- 40 Preferiblemente, todos los híbridos perfectos formados entre las secuencias específicas de la diana de las sondas y las moléculas diana, tienen una temperatura de fusión igual, dentro de un intervalo de 4°C, preferiblemente dentro de un intervalo de 1°C. La diferencia entre una temperatura de fusión del híbrido formado entre la secuencia específica de la diana y la molécula diana, y una segunda temperatura de fusión para la asociación entre la secuencia específica de la diana y una molécula para la que no se ha diseñado la secuencia específica para cada sonda, es preferiblemente superior o igual a 5°C, más preferiblemente superior o igual a 8°C.

Análisis de la mutación

- La descripción también proporciona un método para diseñar y usar biochips de las sondas descritas en esta memoria, para el análisis de secuencias múltiples de ADN que pueden presentar variaciones alélicas. Más particularmente, la presente invención proporciona una herramienta para la determinación de haplotipos en segmentos cortos de ADN que presentan múltiples variaciones puntuales de un solo nucleótido en una secuencia de ADN. Las mutaciones analizables por este método incluyen transiciones (o sustituciones), deleciones o inserciones de un nucleótido, o las variaciones múltiples separadas por una secuencia corta, así como las deleciones e inserciones grandes. El análisis de tales mutaciones puede requerir más de una sonda específica de los alelos.

Hasta cuatro sondas específicas de los alelos se pueden diseñar e inmovilizar para el análisis de una sola transición, en función del grado de alelismo observado para cada variación genética, y el grado de confianza deseado para el análisis. La Tabla I muestra, por ejemplo, las secuencias de horquilla que se pueden utilizar para analizar una transición presente en el codón 142 del gen TCF1, que codifica el factor nuclear hepático 1 (Linder, T. y col. 1999. *Hum. Molec. Genet.* 8:2001-8), 142CC Cy5 que es la molécula diana marcada con Cy5 de la sonda de horquilla StrNCb, y 142CT Cy5, que es la molécula diana marcada con Cy5 de la sonda de horquilla StrNTb. Las letras en negrita representan la secuencia de los bucles, mientras que las letras subrayadas indican la posición del sitio de la mutación. La mutación que tiene origen natural es una transición C/T o G/A en la hebra complementaria. Las sondas de horquilla necesarias para el análisis de la presencia de esta mutación, son sondas portadoras de una C, StrNCb, y una T, StrNCT. Las otras dos sondas de horquilla, StrNGb y StrNAb, se pueden utilizar pero no son indispensables.

TABLA I

Nombre	Secuencia	SEQ ID NO
StrNCb	5'- GCG AGC CAA CCA GTC CCA CCT GTC GCT CGC -3'	1
StrNTb	5'- GCG AGC CAA CCA GTT CCA CCT GTC GCT CGC -3'	2
StrNGb	5'- GCG AGC CAA CCA GTG CCA CCT GTC GCT CGC -3'	3
StrNAb	5'- GCG AGC CAA CCA GTA CCA CCT GTC GCT CGC -3'	4
142CC CY5	Cy5-5'- AGG TGT TGG GAC AGG TGG <u>GAC</u> TGG TTG AGG CCA GTG GTA TCG -3'	5
142CT CY5	Cy5-5'- AGG TGT TGG GAC AGG TGG <u>AAC</u> TGG TTG AGG CCA GTG GTA TCG -3'	6

Se hace referencia a todos los polimorfismos asociados con la región estudiada mediante un mapa suficientemente preciso de la secuencia genómica. Para las variaciones que se deben analizar para la variación alélica, se diseñó un conjunto de cebadores de PCR, utilizando las directrices generales conocidas por el experto en la técnica. Para evitar el uso de cebadores de PCR que contienen variantes alélicas y maximizar la eficacia y la especificidad de las reacciones de PCR, se presta especial atención al hecho de no diseñar parejas de cebadores en una región genómica que contiene polimorfismos conocidos. Para cada mutación que requiere un análisis de la variación alélica, se diseñan y se utilizan al menos dos sondas de horquilla. La posición de la variación alélica está situada de forma la más cercana posible al centro de la secuencia específica del bucle (véase 101 en la Fig. 1), pero puede variar dependiendo de las características de la secuencia. Cuando la localización de la variante está situada lejos del centro del bucle, la T_m del híbrido formado entre la secuencia específica de la diana y una molécula diana que no es perfectamente complementaria de la misma, disminuye. En el diseño de sondas de horquilla para ser utilizadas en el mismo ensayo, el cambio de la longitud de la secuencia específica de la diana (véase más arriba) se utiliza de modo que todas las T_m de los híbridos formados entre la secuencia específica de la diana y la molécula diana, sean iguales dentro de un intervalo de 1-4°C. Por ejemplo, la variación alélica que se va a detectar puede estar localizada a 1 hasta 10 bases de distancia del centro del bucle, preferiblemente de 1 a 5 bases de distancia del centro del bucle. Una vez que las secuencias específicas de la diana diseñadas para todas las sondas de horquilla, de modo que tengan una T_m situada en un intervalo aceptable (véase más arriba), se calcula la T_m de cada secuencia específica de la diana durante la hibridación con otros alelos (heterodúplex) utilizando Meltcalc (véase más arriba). La variación de la T_m entre los dúplex de ADN perfectos e imperfectos se mantiene lo más elevada posible, y se considera suficiente para una discriminación correcta, cuando la T_m del híbrido formado entre una secuencia específica de la diana y una molécula diana que no se asocian perfectamente, es decir, un híbrido que tiene al menos un desemparejamiento de una base, es al menos 5°C menor que la T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana. De forma incluso más preferible, la T_m del híbrido formado entre una secuencia específica de la diana y una molécula diana que no se asocian perfectamente, es al menos 8°C menor que la T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana.

Las secuencias diana pueden ser de ADN genómico, o secuencias de ADN marcadas o sin marcar, amplificadas por PCR. El tamaño de la secuencia analizada en términos de número de nucleótidos puede ser el mismo que el de la parte del bucle de las sondas de hibridación, pero se prefiere un mayor número de nucleótidos, sin limitación en cuanto al tamaño máximo de la secuencia analizada. Cada hebra de estas secuencias diana bicatenarias se puede separar térmicamente e hibridar como una especie individual o como una mezcla, sobre el biochip con sonda de horquilla, de modo que sólo la secuencia perfectamente complementaria se hibrida con su sonda de horquilla específica inmovilizada.

La abundancia relativa de cada alelo detectado en una muestra se puede determinar mediante la medición de la señal relativa para cada sonda específica de cada alelo. Para una transición simple correspondiente a un bialelismo único, dos sondas de horquilla entre las cuatro posibles (teniendo cada una una base de diferencia que se corresponde a un alelo) deben dar una señal para una muestra heterocigota. Este alelo se puede determinar por la posición de las sondas de horquilla en la superficie sólida. Para una muestra homocigota, sólo una de las sondas de horquilla de las cuatro, debe emitir una señal.

De este modo, la descripción describe un método de detección *ex vivo* de una variante genética en una muestra de ácido nucleico que comprende:

- 5 (a) poner en contacto la muestra con un biochip de sondas descritas en esta memoria que comprende al menos dos sondas, en donde al menos una sonda del biochip es una sonda descrita en esta memoria, específica de una variante genética y que tiene una región en bucle perfectamente complementaria de la variante genética, y
 - (b) detectar una señal que proviene de dicha sonda de horquilla de la variante genética,
- en donde la señal detectada en la etapa (b) indica la presencia de la variante genética en la muestra de ácido nucleico.

10 Preferiblemente, la variante genética detectada es un polimorfismo de un solo nucleótido, una delección o una inserción de una o más bases, o la duplicación de una o más bases.

Según una realización preferida, una molécula "informadora" se puede hibridar con cada sonda de horquilla del biochip. La detección de la etapa (b) se obtiene, por lo tanto, utilizando una molécula "informadora" marcada con un marcador detectable. La muestra de la etapa (a) puede estar marcada previamente con un marcador detectable.

Secuenciación

- 15 La descripción describe también un método de secuenciación *ex vivo* de muestras de oligonucleótidos que permite evitar las etapas de preparaciones manuales, implicadas en los métodos de secuenciación que emplean una degradación química y una terminación de cadena con nucleótidos didesoxi. La descripción describe el diseño y el uso de sondas de horquilla, y biochips con sondas de horquilla en donde la secuencia específica de la diana de cada sonda contiene una secuencia idéntica a una porción al menos de otra secuencia específica de la diana de otra
- 20 sonda inmovilizada sobre el mismo biochip. Dependiendo del tamaño de la secuencia que se va a secuenciar, las secuencias específicas de la diana de las sondas de horquilla inmovilizadas, se diseñan además para hibridarse con cualquier secuencia de oligonucleótido posible cuyo tamaño sea por lo menos igual a la longitud de las secuencias específicas de las dianas de las sondas de horquilla. Con este método, se puede diseñar un biochip para secuenciar cualquier oligonucleótido de un tamaño predeterminado. La detección de las hibridaciones que se producen entre las
- 25 porciones de la muestra y las sondas de horquilla, se puede controlar, y la secuencia se puede determinar de acuerdo con las descripciones realizadas en esta memoria y las de los documentos de patente de EE.UU. nº 6.270.961; 6.025.136; 6.018.041; 5.871.928; y 5.695.940.

La descripción describe así un método de secuenciación *ex vivo* de un oligonucleótido que comprende:

- 30 (a) poner en contacto una muestra que contiene el oligonucleótido con un biochip que comprende al menos dos sondas descritas en esta memoria; y
 - (b) detectar una señal que proviene de al menos una de dichas sondas del biochip;
- en donde la señal detectada en la etapa (b) se utiliza para determinar la secuencia del oligonucleótido.

35 De acuerdo con una realización preferida de la invención, una molécula "informadora" se puede hibridar con cada sonda de horquilla del biochip. Detección en la etapa (b) se obtiene utilizando una molécula "informadora" marcada con un marcador detectable. La muestra de la etapa (a) se puede marcar previamente con un marcador detectable.

Análisis de la expresión de los genes

La descripción describe también un método para el diseño y el uso de sondas y biochips de sondas tales como las descritas en esta memoria, para el análisis de la expresión génica.

- 40 Esta aplicación puede ser útil para el análisis de la expresión diferencial de ácidos nucleicos sin marcar, para los que se tiene que analizar un gran número de secuencias diferentes, en términos de cantidades relativas entre sí. Para cada gen que se va a analizar, se puede diseñar al menos una sonda de horquilla para cada parte correspondiente a los exones del gen.

45 La descripción también describe un biochip de sondas tales como las descritas en esta memoria, que comprenden al menos dos sondas inmovilizadas sobre un soporte. Preferiblemente, una estructura de vástago es idéntica para cada una al menos de las dos sondas de horquilla. Incluso más preferiblemente, una molécula "informadora" se puede hibridar con cada sonda descrita en esta memoria, del biochip.

50 Con el fin de vigilar los niveles de expresión de los extractos celulares o de las muestras de tejido, el ARN total, los ARNm y los ADNc y ARNc marcados o sin marcar, se pueden utilizar como secuencia diana para las sondas de horquilla, o los biochips de las sondas de horquilla. Si los ADNc o ARNc se utilizan para el análisis de la expresión, estas secuencias diana se pueden marcar mediante la incorporación de marcadores fluorescentes durante la transcripción inversa, o de etapas adicionales de PCR.

Para el análisis de la expresión diferencial, se pueden comparar varias condiciones. Estas condiciones pueden ser de estrés fisiológico, de grados de diferenciación diferentes, la activación de una vía metabólica o la actividad transcripcional de un fármaco. Cada una de las dos mezclas de ácidos nucleicos, tales como ARNm o ARN total sin marcar, extraídos de las células o de los tejidos, o ADNc se pueden tratar, por ejemplo, por separado e hibridar con un biochip con sondas de horquilla, diseñado específicamente para cada condición o con un conjunto de primeras sondas diseñadas específicamente para un análisis con el SAU.

Detección de las variantes por corte y empalme alternativo y las condiciones para la medición de las cantidades relativas

La descripción describe también un método para analizar los productos del corte y empalme alternativo y la determinación de la abundancia relativa de cada producto del corte y empalme alternativo en diferentes estados fisiológicos del mismo tejido, o la abundancia relativa de cada producto por corte y empalme alternativo para un mismo gen en varios tejidos diferentes. Esta aplicación puede ser útil para el análisis de variantes múltiples de varios genes en donde los productos pueden interaccionar con un esquema definido.

Así, la descripción describe un método de detección *ex vivo* de un producto de corte y empalme alternativo de un gen en una muestra de ácido nucleico que comprende:

(a) poner en contacto la muestra con un biochip de sondas tal como la descrita en esta memoria, que comprende al menos dos sondas tales como las descritas en esta memoria, en donde al menos una de dichas sondas del biochip es una sonda específica de un exón del gen o de la unión de dos exones; y

(b) detectar una señal que proviene de la sonda específica del exón o de la unión de dos exones, en donde la señal detectada indica la presencia del producto por corte y empalme alternativo del gen en la muestra de ácido nucleico.

Preferiblemente, la muestra de ácido nucleico comprende ARNm.

Las secuencias específicas de las dianas de sondas de horquilla están diseñadas para ser complementarias de secuencias de ARNm o ADNc. Más específicamente, están diseñadas para ser complementarias de diferentes productos por corte y empalme ya conocidos de cada secuencia diana. Las variantes por corte y empalme que contienen grandes deleciones o deleciones simples de un nucleótido, se pueden analizar sobre el mismo biochip que es portador de las sondas de horquilla específicas de estas variaciones.

Para cada secuencia de exón supuesta o conocida de un gen estudiado, es posible diseñar al menos una sonda de horquilla específica para detectar la presencia o la ausencia del exón, de la porción del exón, o de la unión de dos exones en el producto génico. El diseño de las sondas de horquilla se realiza de la misma manera que para el análisis de la expresión génica, y se puede utilizar el SAU para dirigir estas secuencias en un soporte. Además, cada variación de la longitud o de la composición del ARNm o del ADNc analizado, puede estar vigilada mediante varias sondas de horquilla, en donde las secuencias específicas de la diana se diseñan para que se hibriden con el centro de la región observada. El análisis de la pérdida o del cambio de un nucleótido entre dos ARNm, se puede analizar, por ejemplo, mediante el diseño de dos sondas de horquilla que tengan un nucleótido de diferencia situado en el centro de las secuencias específicas de la diana.

El diseño de estas sondas de horquilla es similar al descrito para el análisis de mutaciones. Para la detección de grandes deleciones en una secuencia de ARNm, las secuencias específicas de las dianas de dos sondas de horquilla 601 y 602, se pueden diseñar para hibridarse con una secuencia de ARNm (Fig. 6). La secuencia específica de la diana de la primera sonda de horquilla 601, se puede diseñar para que se hibride con una región 604 de un ARNm 603 que puede estar delecionada, formando de este modo una variante por corte y empalme. La secuencia específica de la diana de la segunda sonda de horquilla 602, se puede diseñar para detectar la forma alternativa por corte y empalme 605, hibridándose con las regiones 5' y 3' justo por debajo y por encima de la región delecionada 604. El diseño de estas sondas es similar al de las sondas diseñadas para el análisis de la expresión. Un biochip se puede utilizar entonces para buscar la presencia o ausencia de los exones transcritos de un mismo gen.

El procedimiento para el análisis diferencial de las variantes por corte y empalme es el mismo que para el análisis de la expresión diferencial. Los ARNm procedentes de dos estados fisiológicos de un tejido o de un estado fisiológico de dos tejidos, o de dos tejidos diferentes, se transcriben de forma inversa en ADNc por separado y se marcan (un marcador diferente para cada estado o tejido). Los ADNc marcados de dos fuentes, se mezclan y se hibridan sobre los biochips de las sondas de horquilla o se hibridan con las sondas de horquilla de SAU, y luego se hibridan sobre los biochips de las segundas sondas.

Según una realización preferida de la invención, una molécula "informadora" se puede hibridar con cada sonda de horquilla del biochip. La detección en la etapa (b) se obtiene utilizando una molécula "informadora" marcada con un marcador detectable. La muestra de la etapa (a) se puede marcar previamente con un marcador detectable.

Sondas de horquilla para la detección de moléculas de origen foráneo

La descripción describe también el uso para la detección de moléculas procedentes de cualquier organismo que contiene nucleótidos o restos procedentes de estos organismos, tales como microorganismos, agentes patógenos, virus, parásitos o cualquier modificación que se encuentra en los organismos modificados genéticamente ("OMG") a través de la detección del ADN, ARN o cualquier otra molécula que se asocia con una sonda de horquilla. En esta memoria, la expresión "cualquier resto de organismo que contiene los nucleótidos" se refiere a cualquier molécula o a cualquier molécula resultante de esta molécula, procedente de un organismo que contiene nucleótidos que ha penetrado en otra célula, por ejemplo, un ADN vírico en una célula infectada, o un ARNm producido a partir de este ADN vírico. Esto puede ser particularmente útil en situaciones en las que se tienen que detectar diferentes moléculas de origen foráneo, en la misma muestra, o para analizar variantes genéticas en una muestra biológica.

La descripción también describe un método de detección *ex vivo* de cualquier organismo que contiene ácido nucleico o de cualquier resto de esos organismos que comprende:

(a) poner en contacto una muestra de ácido nucleico con un biochip que comprende al menos dos sondas tales como las descritas en esta memoria, en donde al menos una de dichas sondas es una sonda específica de ácido nucleico de un organismo o de los restos de ese organismo; y

(b) detectar una señal que proviene de dicha sonda específica de ácido nucleico de un organismo,

en donde la señal detectada en la etapa (b) indica la presencia del organismo que contiene el ácido nucleico o los restos de ese organismo.

El organismo que contiene ácido nucleico es preferiblemente un virus o una bacteria.

Según una realización preferida de la invención, una molécula "informadora" se puede hibridar con cada sonda de horquilla del biochip. La detección en la etapa (b) se obtiene mediante el uso de una molécula "informadora" marcada con un marcador detectable. La muestra de la etapa (a) se puede marcar previamente con un marcador detectable.

Preferiblemente, el ARN total, los ARNm o los ADNc marcados o sin marcar, así como los ADNc marcados o no amplificados por PCR, se pueden utilizar sobre biochips de sondas de horquilla para detectar la presencia o ausencia del agente patógeno considerado, los microorganismos, los virus, o los OMGs. Para cada molécula de origen foráneo que se va a analizar, se diseña al menos una sonda de horquilla para al menos una parte exónica de los genes conocidos porque se expresan a niveles constitutivos en la muestra analizada.

Análisis del desequilibrio alélico y de la pérdida de heterocigosidad

La descripción describe también un método para diseñar y usar sondas y biochips de sondas tales como las descritas en esta memoria, para el análisis del desequilibrio alélico debido a deleciones o inserciones de fragmentos cromosómicos y al análisis de la pérdida de heterocigosidad.

Esta aplicación puede ser útil para el análisis de deleciones cromosómicas, para las que los puntos de corte del cromosoma son conocidos o desconocidos. Los puntos de corte del fragmento de ADN en donde se produce la deleción, son frecuentemente desconocidos, y a veces específicos de un individuo. (Mateo M. y col. (1999) *Am. J. Path.*, 154 (5), 1583-1589). Este tipo de alteración genética se encuentra de nuevo en patologías cancerosas, tales como el cáncer de próstata, de mama, ciertas formas de cáncer de colon y ciertos linfomas (Larsson C.M. y col. (2001) *Molecular Diagnosis*, 6 (3), 181-188).

La descripción describe la comparación de secuencias genómicas de células sanas (linfocitos de la sangre, por ejemplo), con las de células sospechosas de ser tumorales (por ejemplo, células procedentes de biopsias o de líquidos como la orina o el líquido cefalorraquídeo, secreciones) utilizando como medio de análisis un biochip de sondas, tal como el descrito en esta memoria, y un método de análisis diferencial.

Este método de análisis consiste en amplificar con PCR, usando métodos conocidos para el experto en la técnica, las secuencias microsatélites, e hibridar los productos de la amplificación sobre un biochip tal como se describe en esta memoria. Estas secuencias microsatélites son secuencias repetidas, altamente polimórficas, presentes a intervalos regulares en el genoma humano, y flanqueadas por secuencias que permiten situar los cebadores de la PCR (Goldstein, D. y C. Schlotterer, compiladores, 1999. *Microsatellites: Evolution and Applications*. Oxford University Press). Las secuencias microsatélites que se van a amplificar se seleccionan del área putativa de deleción, para que se amplifiquen cuando la deleción está ausente, y no generen productos de la PCR cuando la deleción está presente (para la selección de las secuencias microsatélites y sus cebadores para la PCR: *Genome Database*, <http://www.gdb.org>). Las secuencias microsatélites de células sanas y de células presuntamente tumorales se amplifican individualmente y se marcan utilizando cebadores de la PCR marcados de forma diferente para cada tipo, sano o tumoral. Se prefiere una amplificación múltiple de los marcadores microsatélites (PCR multiplex). Se pueden amplificar así de una a varias secuencias microsatélites, dependiendo de la patología y del

análisis que se va a realizar. Una marcación preferida se obtiene con la amplificación de las células sanas con un cebador para la PCR marcado con Cy3 y el cebador complementario sin marcar, y la amplificación de las células tumorales con un cebador para la PCR marcado con Cy5 y el cebador complementario sin marcar. Los productos de la PCR de cada uno de los marcadores microsatélites, marcados con Cy3 para las células sanas y Cy5 para las células tumorales, se tratan a continuación para eliminar los cebadores de la PCR que no se han incorporado durante la amplificación. Se consideran diversos métodos de eliminación de cebadores, con una preferencia por un tratamiento con exonucleasa I y fosfatasa alcalina de camarón (exo/sap), de acuerdo con el protocolo proporcionado, por ejemplo, en el kit exo/sap ExoSAP-IT (USB Corporation, 26111 Miles Road, Cleveland Ohio 44128, EE.UU.).

Al final de este tratamiento, los productos de la amplificación de las secuencias microsatélites se reúnen en una sola solución y se hibridan sobre un biochip de sondas, tal como el descrito en esta memoria. Esta etapa de montaje de los productos de la amplificación de secuencias microsatélites, también se puede hacer antes del tratamiento de eliminación de los cebadores. El biochip descrito en esta memoria incluye al menos una sonda. Esta sonda es al menos parcialmente complementaria de uno de los cebadores de la PCR, usados en la amplificación de la secuencia microsatélite, y preferiblemente, perfectamente complementaria de una parte o a la totalidad del cebador. Se pueden utilizar varios tipos de sondas en el biochip descrito en esta memoria, o preferiblemente sondas de horquilla descritas en esta memoria, modificadas por la adición de un espaciador, o sondas del sistema SAU. La secuencia de la sonda del método preferido se obtiene de acuerdo con los criterios de la presente descripción, particularmente cuando se utilizan varias sondas con varias secuencias microsatélites para analizar, la T_m de todos los híbridos formados entre las secuencias diana de las sondas y las secuencias microsatélites que son complementarias de las mismas, es igual dentro de un intervalo de 4°C, preferentemente, dentro de un intervalo de 1°C. De manera similar, cuando se utilizan varias sondas en un biochip, uno de los brazos de todas las sondas en el método preferido pueden tener la misma secuencia. Cuando se tienen que analizar varias secuencias microsatélites, si se requiere la presencia en el biochip de varias sondas, estas pueden ser de varios tipos, es decir, sondas lineales y/o sondas de horquilla y/o sondas del sistema SAU.

En el método preferido, las sondas del sistema SAU se utilizan para incorporar testigos negativos de la amplificación durante la PCR de las secuencias microsatélites. Estos testigos negativos se realizan mediante la amplificación de una secuencia microsatélite en ausencia de ADN genómico de las células sanas o tumorales, pero en presencia de los reactivos necesarios para la PCR, y una pareja de cebadores de los cuales al menos uno es una primera sonda lineal de SAU que tiene una secuencia marcadora perfectamente complementaria de la secuencia específica de la diana de la segunda sonda inmovilizada.

Durante la hibridación de un producto de la PCR procedente de la amplificación de una secuencia microsatélite de las células sanas y presuntamente tumorales, en una sonda del biochip, la sonda del biochip se hibrida con las secuencias de los cebadores de la PCR, presentes en los extremos de los productos de la PCR de las células sanas y tumorales. Una marcación diferencial con Cy3 y Cy5 de los amplicones procedentes de estas células, resulta en la producción para cada sonda en el chip de una señal compuesta de una emisión para Cy3, debida a la hibridación de la sonda con un producto de la PCR obtenido de las células sanas, y una emisión para Cy5, debida a la hibridación de la sonda con un producto de la PCR obtenido de células presuntamente tumorales. Mediante la medición de la parte relativa de la emisión para cada uno de los marcadores, se deduce la cantidad relativa de productos de la PCR específicos de una secuencia microsatélite, obtenida a partir de células sanas y tumorales. Las secuencias microsatélites analizadas que no se amplifican cuando se produce una delección en la región cromosómica de donde proviene el microsatélite, se observa como resultado una disminución de la emisión de Cy3 (célula presuntamente tumoral) en comparación con la emisión de Cy5 (célula sana) cuando se lleva a cabo esta delección.

En este método, el análisis de la cantidad relativa de productos de la PCR para la amplificación de una secuencia microsatélite se obtiene comparando los valores de la emisión de señales en Cy5 y Cy3 después de la hibridación, con una curva de calibración interna. Esta curva de calibración se obtiene por amplificación e hibridación competitiva sobre las sondas del biochip descrito en esta memoria, del ADN de marcadores microsatélite no delecionados, localizados en el cromosoma Y (marcadores de heterocigosidad) y otros marcadores en otros cromosomas, y no delecionados (marcadores de homocigosidad). En el método preferido, una amplificación de estos marcadores se lleva a cabo en ambos tipos de muestras (células sanas y presuntamente tumorales) al mismo tiempo que la amplificación de los otros marcadores microsatélite, respetando la marcación diferencial con Cy3 y Cy5, tal y como se describió anteriormente. Los ADNs usados para amplificar los marcadores de heterocigosidad y homocigosidad pueden provenir de una extracción del ADN a partir de células sanas y tumorales del individuo o de preparaciones extemporáneas de estos ADNs a partir de individuos de sexo masculino, con genotipos conocidos (que no contienen delecciones para estos marcadores) para los marcadores microsatélite analizados en esta amplificación.

Alternativamente, la curva de calibración se puede obtener amplificando por PCR dos o más marcadores microsatélite sin delección en la patología considerada. Estos marcadores se amplifican en presencia de una dilución en serie de los ADNs extraídos de cada uno de los dos tipos celulares comparados (células sanas y presuntamente tumorales), y al mismo tiempo que los otros marcadores microsatélite. La curva de calibración se establece, después de completar el análisis mediante el trazado en un gráfico que tiene en la abscisa la intensidad de la fluorescencia de Cy3 y en la ordenada la intensidad de la fluorescencia de Cy5, en donde los valores respectivos de estas dos

fluorescencias para cada una de las sondas del biochip se corresponden a los marcadores de calibración.

El análisis de cada uno de los productos de la PCR en términos de presencia o ausencia de un desequilibrio alélico, se realiza estableciendo en este mismo gráfico los valores de la fluorescencia en Cy3 y Cy5, obtenidos después de la hibridación de los productos de la amplificación de las secuencias satélite en el biochip.

- 5 La descripción también describe kits de biochips tales como los descritos en esta memoria, de sondas y reactivos, incluyendo reactivos de RT-PCR para el análisis de desequilibrios alélicos en:

- cánceres de próstata, de riñón y de vejiga

- cánceres de mama

- cánceres de colon

- 10 - cánceres de pulmón

- análisis de la presencia de deleciones importantes en enfermedades hereditarias tales como la miopatía de Duchenne y enfermedades adquiridas.

Comentarios de las Figuras:

- 15 Fig. 13A: el cebador de PCR 1302, marcado en su extremo 5' con una molécula fluorescente (Cy5), está situado en 3' de la hebra codificadora de la secuencia microsatélite 1301, y se usa con el cebador 1303 sin marcar, situado en 3' de la hebra no codificadora de la secuencia microsatélite 1301, para amplificar esta secuencia microsatélite en las muestras de ADN genómico extraídas de células sanas 1304, y para producir el producto de la PCR 1305.

- 20 El cebador de PCR 1307, marcado en su extremo 5' con una molécula fluorescente (Cy3), está situado en 3' de la hebra codificadora de la secuencia microsatélite 1301, y se usa con el cebador 1308 sin marcar, situado en 3' de la cadena no codificadora de la secuencia microsatélite 1301, para amplificar esta secuencia microsatélite en las muestras de ADN genómico, extraídas de células presuntamente tumorales 1309, y para producir el producto de la PCR 1310.

- 25 Fig. 13B: El biochip de la sonda se pone en contacto 1315 con la mezcla que contiene productos de la PCR procedentes de las etapas de amplificación 1304 y 1309 de la Fig. 3A. El biochip de sonda se representa con dos sondas de horquilla 1311 y 1314, idénticas, ambas inmovilizadas sobre el soporte 1312 a través de un espaciador 1313. La secuencia específica de la diana 1319 de estas dos sondas es totalmente complementaria de la secuencia de los cebadores de la PCR, codificadora 1318, incluida en los productos de la PCR 1316 y 1317. El producto de la PCR 1316 se marca con Cy5 y el producto de la PCR 1317 se marca con Cy3. El producto de la PCR 1316 se obtiene mediante la desnaturalización del producto de la PCR 1305 (Figura 3A), y el producto de la PCR 1317, por desnaturalización del producto de la PCR 1310 (Figura 3A).

Kits de biochips de horquilla

- 35 La descripción describe además el suministro de kits para detectar la presencia de moléculas diana en una muestra. Dichos kits comprenden típicamente dos o más compuestos necesarios para llevar a cabo el análisis. Dichos compuestos pueden incluir un biochip de sondas de horquilla y uno o varios reactivos adicionales, necesarios para la detección de las moléculas diana seleccionadas. Alternativamente, el kit puede comprender un conjunto de sondas de horquilla inmovilizadas y/o no inmovilizadas y reactivos tales como los utilizados en el SAU. Un kit preferido comprende el biochip de sondas de horquilla, un conjunto de sondas de horquilla no inmovilizadas y/o una molécula "informadora" y uno o varios reactivos adicionales. Otro kit preferido comprende el biochip de sondas lineales, un conjunto de sondas de horquilla no inmovilizadas y/o una molécula "informadora", y uno o varios reactivos adicionales.

Ejemplos: Hibridaciones de ADN diana con las sondas de horquilla inmovilizadas

Los siguientes ejemplos se presentan a modo de ilustración de la presente invención, y no pretenden limitar la invención de ninguna manera. Los ejemplos presentados a continuación describen el uso de biochips de sondas de horquilla para la detección de secuencias diana.

Las moléculas de ácido nucleico

- 45 La mutación del codón 137 en el exón 2 del gen TCF2 es una deleción de 75 pares de bases que tiene como resultado la aparición de un fenotipo diabético de tipo II, asociado con una disfunción renal grave (Linder, T. y col. 1999. *Hum. Molec. Genet.* 8:2001-8). La secuencia diana usada en este ejemplo se corresponde a la región 266 a 290 de la secuencia delecionada del codón 137 en el exón 2 del gen TCF2. La Fig. 7A muestra la posición de la mutación del codón 137 del exón 2. La parte en negrita y entre corchetes, comprende las bases 276 a 350, que están delecionadas en los individuos mutantes. La Fig. 7B muestra la secuencia del exón 2 que es portadora de la

mutación.

La secuencia diana, Tg2X2C137M1, utilizada en este ejemplo, la Fig. 7C, es la secuencia complementaria, o la hebra negativa de la secuencia en negrita de la Fig. 7B. La secuencia específica de la sonda de horquilla 2X2C137M1, que se muestra en la Fig. 7D, se diseñó mediante la selección de 18 nucleótidos en el centro de la secuencia diana mutante, que es perfectamente complementaria de la secuencia diana. La molécula "informadora" diseñada para este análisis, se muestra en la Fig. 7E y tiene una temperatura de fusión de 23°C en presencia de una solución de sodio 1 M y una concentración de la solución de oligonucleótidos de 1 µM. El vástago de la sonda de horquilla 2X2137M1 tiene una T_m de 72°C. La secuencia específica de la diana de las sondas de horquilla se ha diseñado para tener una T_m del híbrido formado entre la secuencia específica de la diana y la molécula diana de 60°C ± 2°C, utilizando el programa Meltcalc, con los parámetros fijos en 100 mM en términos de sodio y para una concentración de oligonucleótidos de 100 mM (la T_m del dúplex Tg2X2C137M1/2X2C137M1 = 60,8°C).

El programa Meltcalc se utiliza para la predicción de las temperaturas de fusión durante las hibridaciones de oligonucleótidos. Este programa permite ajustar una T_m a través de variaciones en la concentración salina y en DMSO, así como de oligonucleótidos. Las sondas de horquilla de estos ejemplos se han diseñado utilizando concentraciones de 0,1 mM en sal de sodio y 0,1 mM en oligonucleótidos. La asociación sobre sí mismos de los brazos de la sonda de horquilla, se evalúa con mfold, utilizando los siguientes parámetros: temperatura de la asociación 37°C, sodio 1 M, sin magnesio, corrección de tipo oligómero, 5% de suboptimalidad, y un límite superior de asociación informatizada de 50. Todas las sondas de horquilla incluidas en mfold tienen un vástago continuo con una longitud máxima de 6 nucleótidos, con una T_m de 65-72°C, determinada utilizando los parámetros anteriores. La T_m de los híbridos perfectamente complementarios formados entre las moléculas diana y las secuencias específicas de las dianas de tipo silvestre y mutantes es 60°C, con un ΔT_m de al menos 5°C entre las hibridaciones perfectas e imperfectas, y una T_m de 65-72°C para todos los vástagos.

Inmovilización de las sondas de horquilla

Todas las sondas de horquilla, sintetizadas con un espaciador aminado en su extremo 5', se depositaron sobre portaobjetos de vidrio activados con N-hidroxi-succinimida ("NHS") (NoAb, Diagnostics, Mississauga, Ontario, Canadá) en un tampón de depósito, que fue proporcionado con los portaobjetos, durante 45 minutos a 30°C, en una cámara humidificada. Los restantes grupos NHS fueron desactivados mediante reacción con una solución de saturación proporcionada también con los portaobjetos, a temperatura ambiente durante 30 minutos. Los portaobjetos se lavaron a continuación en tampón citrato salino a pH 7, que comprendía NaCl 0,1 M, y se enjuagaron con agua milliQ.

Hibridación de las moléculas diana

Cada biochip de sondas inmovilizadas se trató con 20 µL de una solución 400 nM de la molécula diana, Tg2X2C137M1, marcada con Cy5, en un tampón SSC 6X que contenía 50% de formamida. Después de depositar un cubreobjetos sobre cada portaobjetos, la solución de la molécula diana se dejó hibridar durante una hora a temperatura ambiente. Cada biochip se lavó a continuación durante cinco minutos con SSC 4X y luego se lavó durante 2 minutos con tampón SSC 0,1X. Finalmente, cada portaobjetos se analizó en un escáner de biochip (Axon Instruments, Inc., Union City, California, EE.UU.), utilizando tanto longitudes de onda de excitación y de emisión de Cy5 y Cy3.

La Fig. 8 muestra las intensidades fluorescentes de Cy5 y Cy3 después de la hibridación de una molécula diana marcada en Cy5. Cada valor es una media de las mediciones de 5 réplicas por cada sonda de horquilla. Las diez sondas de horquilla utilizadas en esta prueba de detección están diseñadas para el análisis de cinco mutaciones diferentes presentes en diferentes exones del gen TCF2. Sólo la sonda 2X2C137M1 ("137M1" en la Fig. 8) es perfectamente complementaria de la diana Tg2X2C137M1. El testigo es una secuencia monocatenaria lineal marcada en 5' en Cy3 y modificada en su extremo 3' con un espaciador aminado. Este testigo se utiliza para situar los depósitos y para controlar la calidad de los depósitos. La señal en Cy5 asociada con la hibridación de la molécula diana marcada Cy5, se observa esencialmente en los depósitos correspondientes a las secuencias perfectamente complementarias de la diana, el bucle de la sonda de horquilla 2X2C137M1.

Hibridación de la molécula "informadora"

Tras la hibridación con la molécula diana, cada biochip de sondas de horquilla inmovilizadas se trató con 20 µL de una solución 1 M de molécula "informadora", marcada en Cy3 en un tampón SSC 6X que contenía 50% de formamida. Después de depositar un cubreobjetos sobre cada portaobjetos, la solución de la molécula "informadora" se dejó hibridar a temperatura ambiente, durante dos horas. Cada portaobjetos se lavó a continuación dos veces durante cinco minutos, en tampón SSC 4X, y luego se lavó durante dos minutos en tampón SSC 0,1X. Cada portaobjetos se analizó a continuación en un escáner de biochip utilizando longitudes de onda de excitación y de emisión de Cy5 y Cy3.

La Fig. 9 muestra las intensidades medias de la fluorescencia de Cy5 y Cy3, calculadas basándose en las mediciones efectuadas sobre 5 réplicas, después de la hibridación con la molécula "informadora" marcada en Cy3.

Las señales de Cy3 asociadas a la hibridación de la molécula "informadora" marcada en Cy3, se observan fundamentalmente en los puntos en donde la secuencia diana Tg2X2C137M1, marcada en Cy5, se hibrida con la secuencia que es perfectamente complementaria de ella, es decir, el bucle de la sonda de horquilla 2X2C137M1. Los resultados muestran que la molécula "informadora" se fija preferencialmente sobre las sondas abiertas y que estas sondas abiertas sólo se observan cuando las moléculas diana se hibridan con las sondas de horquilla que contienen la secuencia específica de la diana.

Especificidad de las sondas de horquilla

Las interacciones no específicas entre las sondas de horquilla y las secuencias diana se pueden reducir variando las condiciones de la hibridación. En la Fig. 8, se observa una interacción no específica evidente entre la diana marcada Cy5 y la sonda 151M1. En la Fig. 9, se observa una interacción no específica evidente entre la molécula "informadora" y todas las sondas de horquilla, con la excepción de la sonda 2X2C137M1. Una forma de eliminar las interacciones no específicas, es añadir etapas de lavado. La Fig. 9 muestra que con lavados adicionales, se reduce la interacción no específica entre la secuencia diana marcada en Cy5 y la sonda de horquilla 151M1. Otro método para disminuir las interacciones no específicas, es utilizar condiciones más rigurosas. La Fig. 10 muestra cómo un lavado más riguroso durante 15 a 30 minutos con tampón SSC 4X a 28°C, que es el valor de la T_m del "informador" +5°C, elimina la mayor parte de las interacciones no específicas entre la molécula "informadora" y todas las sondas de horquilla, excepto la sonda de horquilla 2X2C137M1.

Discriminación de un nucleótido con las sondas de horquilla

La Tabla II muestra las secuencias de dieciséis sondas de horquilla que se diseñaron para analizar mutaciones en los genes de la lipasa hepática ("LIPC"), la colesterol éster transferasa ("CEPT"), la lipoproteína lipasa ("LPL") y TCF1. En la Tabla II se resume el nombre de cada sonda de horquilla, las correspondencias con los nombres de los genes y los codones, la secuencia de cada sonda de horquilla, y el número y el tipo de desemparejamientos que presenta la sonda de horquilla, cuando se hibrida con las moléculas diana. Las secuencias específicas de la diana de las sondas de horquilla o bien son perfectamente complementarias de la secuencia diana especificada, o no son perfectamente complementarias y presentan uno o dos desemparejamientos; o no son en absoluto complementarias de la molécula diana, indicándose el caso con "-". Las sondas de horquilla se han diseñado mediante la selección de una secuencia de bucle de 18 a 24 nucleótidos, complementarios de la secuencia diana que tiene un polimorfismo de un nucleótido. Todas las variantes de SNP se sitúan en el centro del bucle (caracteres en negrita) con 9 a 11 nucleótidos que flanquean la posición polimórfica. La T_m de cada bucle de tipo silvestre o mutante hibridado con su secuencia diana, se ha fijado en 60°C ± 2°C, usando el programa informático Meltcalc con los parámetros de concentración establecidos en sodio 100 mM y oligonucleótidos 100 mM. La secuencia de las sondas de horquilla se ajustó mediante la adición y/o eliminación de uno o más nucleótidos en cada extremo del bucle, para llegar a la T_m de 60°C, con una variación de la temperatura de fusión entre los heterodúplex y homodúplex de al menos 5°C.

La inmovilización de las sondas de horquilla se realizó tal y como se ha descrito anteriormente. Cada biochip de sondas de horquilla inmovilizadas se trató con 20 µL de una solución tampón de sodio 1 M, que contenía diana M19 400 nM marcada en Cy3, y diana 142CC 400 nM marcada en Cy5, durante una hora en tampón SSC 6X a 20°C, cubierto con un cubreobjetos de cristal. El biochip se lavó a continuación durante cinco minutos con SSC 4X y luego se lavó con SSC 0,1X durante dos minutos. El biochip se analizó después tal y como se ha descrito anteriormente, usando las longitudes de onda de excitación y de emisión de Cy5 y Cy3, y utilizando un factor fotomultiplicador ("PMT") de 600x600 para cada fluorocromo.

La Fig. 11 muestra las intensidades medias de la fluorescencia de Cy5 y Cy3, calculadas basándose en las mediciones de 5 réplicas, después de la hibridación de las secuencias diana marcadas con ayuda de Cy5 y Cy3. La señal en Cy3 corresponde a la hibridación de la diana M19, y la señal en Cy5 corresponde a la hibridación de la diana 142 CC. Sólo las sondas de horquilla parcial o totalmente complementarias de las secuencias diana fueron capaces de generar una señal tras la hibridación de las secuencias diana, bajo condiciones poco rigurosas. Para discriminar los homodúplex de los heterodúplex, cada biochip se lavó durante 20 minutos en tampón SSC 6X a 50°C, seguido por un lavado con SSC 6X y de nuevo un escaneo con las longitudes de onda de fluorescencia de Cy5 y Cy3, con un PMT de 600x600, para los dos fluorocromos.

La Fig. 12 muestra las intensidades medias de la fluorescencia de Cy5 y Cy3, calculadas basándose en mediciones realizadas para 5 réplicas, después de someter cada biochip a una etapa de lavado en condiciones más rigurosas. Las sondas de horquilla perfectamente complementarias de las dianas marcadas, StrM19G y StrNCb, tienen el mayor nivel de hibridación. Las sondas de horquilla que tienen un único desemparejamiento, tienen una señal de hibridación más débil pero intermedia, mientras que las sondas que presentan dos desemparejamientos tienen la señal de hibridación más débil. Esta diferenciación en la hibridación se obtiene a través de una mayor estabilidad de los híbridos perfectos formados entre las sondas de horquilla y las secuencias diana.

TABLA II

Gen/Número de codón en el gen	Nombre de la sonda	Secuencia	Desemparejamientos con la diana	SEQ ID NO
LPL/477 S447X C/G	StrM11C	5'-GCG AGC GAA TAA GAA GTA GGC TGG TGA GC GCT CGC-3'	-	12
	StrM11G	5'-GCG AGC GAA TAA GAA GTA GGC TGG TGA GC GCT CGC-3'	-	13
LPL/188 Gly188Glu G/A	StrM16G	5'-GCG AGC CAC CAG AGG GTC CCC TGG GCT CGC-3'	-	14
	StrM16A	5'-GCG AGC CAC CAG AGA GTC CCC TGG GCT CGC- 3'	-	15
LIPC/480 C-480T C/T	StrM6C	5'-GCG AGC TTT TGA CAG GGG GTG AAG G GCT CGC- 3'	-	16
	StrM6T	5'-GCG AGC TTT TGA CAG GGG GTG AAG G GCT CGC-3'	-	17
CETP IV/WIAF- 10949 G/A	StrM19G	5'-GCG AGC CCG AGT CCG TCC AGA GCT GCT CGC-3'	diana M19: perfectamente complementaria	17
	StrM19A	5'-GCG AGC CCG AGT CCA TCC AGA GCT GCT CGC-3'	diana M19: 1 desemparejamiento C/A	19
TCF1/142	StrNCb	Véase Tabla 1	diana 142CC: perfectamente complementaria	1
	StrNTb	Véase Tabla 1	diana 142CC: 1 desemparejamiento T/G	2
	StrNGb	Véase Tabla 1	diana 142CC: 1 desemparejamiento G/G	3
	StrNAb	Véase Tabla 1	diana 142CC: 1 desemparejamiento A/G	4
	StrNC	5'-GCG AGC CAA CCA CTC CAC CTG TC GCT CGC-3'	diana 142CC: 1 desemparejamiento T/C	20
	StrNT	5' -GCG AGC CAA CCA TTC CAC CTG TC GCT CGC- 3'	diana 142CC: 2 desemparejamientos T/G y T/C	21
	StrNG	5'-GCG AGC CAA CCA GTC CAC CTG TC GCT CGC-3'	diana 142CC: 2 desemparejamientos G/G y T/C	22
	StrNA	5'-GCG AGC CAA CCA ATC CAC CTG TC GCT CGC- 3'	diana 142CC: 2 desemparejamientos G/G y T/C	23

Ejemplo de diseño de sondas de horquilla

5 Las sondas mostradas en la tabla III se han diseñado para analizar las mutaciones presentes en individuos afectados por la enfermedad de Charcot Marie Tooth.

La enfermedad de Charcot Marie Tooth es la enfermedad hereditaria más frecuente del sistema nervioso periférico. De naturaleza heterogénea (multigénica), se caracteriza por una multitud de trastornos neuropáticos motores y sensoriales (<http://molgen-www.uia.ac.be/CMTMutations/CMT.cfm>).

10 A nivel molecular, se conocen actualmente 14 genes que están implicados, a través de sus mutaciones, en la aparición de las neuropatías de Charcot Marie Tooth. Entre ellos, los genes de la proteína 22 de la mielina periférica (PMP22), de la proteína cero de la mielina (MPZ) y de la conexina 32 (GJB1), en los que las mutaciones son dominantes, se utilizan en este ejemplo para el diseño de sondas que permiten el análisis de 5 mutaciones, enumeradas a continuación:

Nombre de la mutación	Tipo	Gen
C42R T>C	Sustitución T>C	PMP22
W140R T>C	Sustitución T>C	PMP22
T124M C>T	Sustitución C>T	MPZ
V113F G>T	Sustitución G>T	MPZ
S26L C>T	Sustitución C>T	GJB1

El diseño de estas sondas se ha realizado de acuerdo con las etapas descritas en la sección de diseño de las sondas, en la descripción detallada de la invención.

Estas etapas son:

- 5 1- Determinar la secuencia de los bucles de las sondas mutantes y de tipo silvestre, y calcular las Tms de los híbridos perfectos e imperfectos con el programa Meltcalc.
- 2- Añadir los brazos de las sondas (vástagos), comprobar la estructura de las horquillas y calcular la Tm de las horquillas con Mfold.

3- Verificar la alineación de las sondas con sus dianas respectivas, y la ausencia de homología entre las horquillas diseñadas y las dianas no complementarias (Blastn).

- 10 El conjunto de las sondas diseñadas para el análisis de estas mutaciones se enumeran en la tabla III. Estas sondas se han utilizado para diseñar un biochip, mediante el siguiente procedimiento:

Cada sonda se ha sintetizado con un grupo amino en 5' (Sigma Genosys, GB). Estas sondas se depositaron sobre portaobjetos de vidrio activados (Genescore, Francia) mediante el depósito de las muestras por "spotting" (robot microgrid III Biorobotics). Después del depósito, se forma un enlace covalente se forma entre el soporte de vidrio y el extremo aminado de la sonda. Se realizaron cinco réplicas de cada depósito.

- 15

Los biochips así constituidos se utilizaron para hibridar un conjunto de dianas de oligonucleótidos (tabla IV) perfectamente complementarias de las sondas indicadas en la tabla III, y marcadas con una moléculas fluorescente (cianina 3) en su extremo 5' (Sigma Genosys).

TABLA III

Sondas de horquilla	5'-3'	SEQ ID NO
Neu1S26LW6	GCGAGCAGTATGGCTCTCGGTCATGCTCGC	24
Neu1S26LM6	GCGAGCAGTATGGCTCTTGGTCATGCTCGC	25
Neu4V113FW2	GCGAGCGCTCCATTGTCATACACAAGCTCGC	26
Neu4V113FM2	GCGAGCGCTCCATTTTCATACACAAGCTCGC	27
Neu4T124MW3	GCGAGCCAATGGCACGTTCACTTGCTCGC	28
Neu5C42RW1	GCGAGCCAGAACTGTAGCACCGCTCGC	29
Neu5C42RM1	GCGAGCCAGAACCGTAGCACCGCTCGC	30
Neu6W140RW4	GCGACGTCCTGGCCTGGGTGCGTCGC	31
Neu6W140RM4	GCGACGTCCTGGCCCGGGTGCGTCGC	32

20

Se realizaron dos series de experimentos en dos biochips diferentes. En el primer chip, se hibridó el conjunto de dianas que se correspondían a los alelos de tipo silvestre (WT, tabla IV (a)). En el segundo biochip se hibridaron todas las dianas correspondientes a los alelos mutantes (MT, tabla IV (b)).

TABLA IV (a)

Dianas WT	5'-3'	SEQ ID NO
CiS26LW	ATGAAGATGACCGAGAGCCATACTCGGCCA	33
CiV113FW	TCTAGGTTGTGTATGACAATGGAGCCATCC	34
CiT124MW	CGTCACAAGTGAACGTGCCATTGTCACTGT	35
CiC42RW	GAAGAGGTGCTACAGTTCTGCCAGAG	36
CiW140RW	GAAGGCCACCCAGGCCAGGATGTAGG	37

25

TABLA IV (b)

Dianas MT	5'-3'	SEQ ID NO
CiS26LM	ATGAAGATGACCAAGAGCCATACTCGGCCA	38
CiV113FM	TCTAGGTTGTGTATGAAAATGGAGCCATCC	39
CiC42RM	GAAGAGGTGCTACGGTTCTGCCAGAG	40
CiW140RM	GAAGGCCACCCGGGCCAGGATGTAGG	41

Para cada una de estas hibridaciones, las condiciones se establecieron en: dianas 100 nM en tampón SSC 6X, NaCl 1 M, hibridadas durante 1 hora a temperatura ambiente, 20 µL de diana por biochip. Valor medio de 5 replicados, escaneado en un escáner GST Lumonics, láser y PMT a 500x500.

30

Los valores medios de la fluorescencia después del escaneo se analizaron en cada serie de experimentos para las sondas de tipo silvestre y mutantes, y se utilizaron para calcular el índice de discriminación de cada sonda (Rd).

El índice de discriminación de las sondas (Rd) se define como:

$Rd = \text{intensidad de la fluorescencia de un híbrido perfecto} / \text{intensidad de la fluorescencia de un híbrido imperfecto}$.

- 5 Con una intensidad de la fluorescencia de híbrido perfecto = intensidad medida en la hibridación de la diana perfectamente complementaria de la horquilla considerada (por ejemplo, la hibridación de la diana CiS26LW con la horquilla Neu1S26LW6, o la hibridación de la diana CiS26LM con la horquilla Neu1S26LM6) y la intensidad de la fluorescencia de un híbrido imperfecto = la intensidad medida en la hibridación de la diana con una base diferente (por ejemplo, la hibridación de la diana CiS26LW con la horquilla Neu1S26LM6, o la hibridación de la diana CiS26LM con la horquilla Neu1S26LW6).

El índice de discriminación (Rd) tiene el siguiente significado:

por ejemplo, un Rd de 10 para una horquilla dada, significa que la hibridación de una diana perfectamente complementaria de esta horquilla, generará una señal 10 veces más elevada que la hibridación de otra diana que difiere en una base de la primera.

- 15 Estos valores de Rd se trasladan a la tabla V, así como el valor calculado de la Tm de los bucles, y los vástagos de cada horquilla.

El parámetro Dtm también se calcula para cada una de estas sondas, con la $Dtm = Tm \text{ del vástago} - Tm \text{ del bucle}$. Este parámetro describe la "fuerza relativa" de la discriminación de las sondas. Cuanto más elevada es la Dtm (valor de la Tm del vástago elevada, y valor de la Tm del bucle baja), mayor será la capacidad para discriminar dos secuencias con homología elevada.

- 20 Todos estos valores se representan en una gráfica (Figura 14) que muestra la relación entre la evolución del valor de Dtm y el valor del índice de discriminación Rd.

TABLA V

Nombre	Tm del bucle	Tm del vástago	Dtm	Rd
Neu4V113FW2	63,1	69,1	6	9,95
Neu6W140RW4	67,3	73,8	6,5	9,61
Neu6W140RM4	63,6	71,6	8	9,9
Neu1S26LW6	65	73,9	8,9	11,65
Neu4T124MW3	62,8	72,6	9,8	13,16
Neu4V113FM2	61,1	72	11	10,04
Neu1S26LM6	62,7	73,9	11	7,93
Neu5C42RM1	60	73,4	13	8,76
Neu5C42RW1	56,7	73,4	17	12,48

- 25 Esta relación entre Rd y Dtm se puede utilizar para seleccionar los bucles y los vástagos durante el diseño de las sondas, con un tamaño y una composición adaptadas para el análisis deseado. Por ejemplo, si se requiere un Rd de 10 para el diseño de todas las sondas, el valor de Dtm que servirá como criterio para el diseño de las sondas, será de 9.

Ejemplo de resultado del análisis de mutaciones en una muestra de ADN genómico humano

- 30 Este ejemplo se refiere específicamente al análisis del alelo de tipo silvestre de la mutación W140R del gen PMP22 (exón 3), implicado en la enfermedad de Charcot Marie Tooth.

Todas las sondas de horquilla cuya secuencia se muestra en la tabla III, se inmovilizan sobre un biochip.

El ADN genómico del exón 3 del gen PMP22 se amplificó con PCR, utilizando una pareja de cebadores, en donde uno de ellos estaba marcado en su extremo 5' con una molécula de cianina 3 fluorescente.

- 35 La PCR se realizó del modo siguiente: 100 ng de ADN extraído de una persona sana se utilizaron con un kit de PCR (High fidelity expand PCR, Roche), siguiendo el protocolo proporcionado por el fabricante, y los cebadores específicos de PCR durante 40 ciclos. Un testigo negativo (pocillo de PCR sin ADN genómico) se incluyó también durante la PCR.

- 40 Los productos de la PCR se controlaron por electroforesis en un gel de agarosa, para verificar el tamaño del producto de la PCR, así como la falta de contaminación en el testigo negativo. El ADN amplificado se cuantificó por

adsorción de UV a 260 nm, después de la purificación en microcolumna (Quiaquick, Qiagen).

Se hibridaron 1,4 µg de ADN marcado en la cianina 3, procedente de la PCR del exón 3 del gen PMP22, durante 1 hora a temperatura ambiente sobre el biochip portador de las sondas de horquilla.

Tras la hibridación, el biochip se escaneó y los resultados se presentan en la Fig. 15.

- 5 Estos resultados muestran que los productos de la PCR se hibridan principalmente con las sondas que son complementarias de los mismos, o que difieren en una base (sondas Neu6W140RM4 y Neu6W140RM4). El Rd es 6,61 (señal de la sonda WT/MT).

El genotipo dado para este análisis es homocigoto de tipo silvestre para la mutación W140R, que se corresponde al fenotipo "sano" del individuo sometido a ensayo.

10 Referencias citadas

La mención de una referencia no se interpretará como una admisión de que tal referencia pertenece a la técnica anterior de la presente invención.

LISTA DE SECUENCIAS

5 <110> VALAT, CHRISTOPHE
 <120> Biochips con sondas y sus métodos de uso
 <130> B0175wo

10 <150> FR 03 03112
 <151> 13--03-2003
 <160> 41

15 <170> PatentIn versión 3.1
 <210> 1
 <211> 30
 <212> DNA
 20 <213> Secuencia artificial
 <220>
 <223> sonda de horquilla

25 <220>
 <221> horquilla (stem_loop)
 <222> (1)..C6)
 <223>

30 <220>
 <221> unión a sonda (probe_bind)
 <222> (7)..(24)
 <223>

35 <220>
 <221> horquilla (stem_loop)
 <222> (25)..(30)
 <223>

40 <400> 1
 gcgagccaac cagtcccacc tgtcgctcgc 30
 <210> 2
 <211> 30
 45 <212> DNA
 <213> Secuencia artificial
 <220>
 <223> sonda de horquilla

50 <220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>

55 <220>
 <221> unión a sonda (probe_bind)
 <222> (7)..(24)
 <223>

60 <220>
 <221> horquilla (stem_loop)
 <222> (25)..(30)

<223>
 <400> 2
 5 gcgagccaac cagttccacc tgcgctcgc 30
 <210> 3
 <211> 30
 <212> DNA
 <213> Secuencia artificial
 10 <220>
 <223> sonda de horquilla
 <220>
 15 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 <220>
 20 <221> unión a sonda (probe_bind)
 <222> (7)..(24)
 <223>
 <220>
 25 <221> horquilla (stem_loop)
 <222> (25)..(30)
 <223>
 <400> 3
 30 gcgagccaac cagtgccacc tgcgctcgc 30
 <210> 4
 <211> 30
 <212> DNA
 35 <213> Secuencia artificial
 <220>
 <223> sonda de horquilla
 <220>
 40 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 <220>
 45 <221> unión a sonda (probe_bind)
 <222> (7)..(24)
 <223>
 <220>
 50 <221> horquilla (stem_loop)
 <222> (25)..(30)
 <223>
 <400> 4
 55 gcgagccaac cagtaccacc tgcgctcgc 30
 <210> 5
 <211> 42
 60 <212> DNA
 <213> Secuencia artificial
 <220>
 65 <223> diana del oligonucleótido

<400> 5
 aggtgtggg acaggtggga ctggtgagg ccagtggat cg 42

5 <210> 6
 <211> 42
 <212> DNA
 <213> Secuencia artificial

10 <220>
 <223> diana del oligonucleótido

<400> 6
 aggtgtggg acaggtggaa ctggtgagg ccagtggat cg 42

15 <210> 7
 <211> 120
 <212> DNA
 <213> Homo sapiens

20 <400> 7
atcaagggtt acatgcagca acacaacatc cccagaggg aggtggtcga tgcaccggc 60
ctgaaccagt cgcacctctc ccagcatctc aacaaggga cccctatgaa gaccagaag 120

25 <210> 8
 <211> 50
 <212> DNA
 <213> Homo sapiens

30 <400> 8
 atcaagggtt acatgcagca acacaacatc cccaggacc cagaagcgtg 50

35 <210> 9
 <211> 24
 <212> DNA
 <213> Secuencia artificial

<220>
 <223> diana del oligonucleótido

40 <400> 9
 cacgcttctg ggtcctgggg gatg 24

45 <210> 10
 <211> 30
 <212> DNA
 <213> Secuencia artificial

<220>
 <223> sonda de horquilla

50 <220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>

55 <220>
 <221> horquilla (stem_loop)
 <222> (25)..(30)
 <223>

60 <220>
 <221> unión a sonda (probe_bind)
 <222> (7)..(24)

<223>
 <400> 10
 5 gcgagccccc caggaccag aagcgctcgc 30
 <210> 11
 <211> 6
 <212> DNA
 <213> Secuencia artificial
 10 <220>
 <223> sonda
 <400> 11
 15 gcgagc 6
 <210> 12
 <211> 35
 <212> DNA
 20 <213> Secuencia artificial
 <220>
 <223> sonda de horquilla
 25 <220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 30 <220>
 <221> horquilla (stem_loop)
 <222> (30)..(35)
 <223>
 35 <220>
 <221> unión a sonda (probe_bind)
 <222> (7)..C29
 <223>
 40 <400> 12
 gcgagcgaat aagaagtagg ctggtgagcg ctgc 35
 <210> 13
 <211> 35
 45 <212> DNA
 <213> Secuencia artificial
 <220>
 <223> sonda de horquilla
 50 <220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 55 <220>
 <221> horquilla (stem_loop)
 <222> (30)..(35)
 <223>
 60 <220>
 <221> unión a sonda (probe_bind)
 <222> (7)..(29)
 65 <223>

	<400> 13 gcgagcgaat aagaagtagg ctggtgagcg ctcgc	35
5	<210> 14 <211> 30 <212> DNA <213> Secuencia artificial	
10	<220> <223> sonda de horquilla	
	<220> <221> horquilla (stem_loop)	
15	<222> (1)..(6) <223>	
	<220> <221> horquilla (stem_loop)	
20	<222> (25)..(30) <223>	
	<220> <221> unión a sonda (probe_bind)	
25	<222> (7)..(24) <223>	
	<400> 14 gcgagccacc agagggtccc ctgggctcgc	30
30	<210> 15 <211> 30 <212> DNA <213> Secuencia artificial	
35	<220> <223> sonda de horquilla	
	<220> <221> horquilla (stem_loop)	
40	<222> (1)..(6) <223>	
	<220> <221> horquilla (stem_loop)	
45	<222> (25)..(30) <223>	
	<220> <221> unión a sonda (probe_bind)	
50	<222> (7)..(24) <223>	
	<400> 15 gcgagccacc agagagtccc ctgggctcgc	30
55	<210> 16 <211> 31 <212> DNA <213> Secuencia artificial	
	<220> <223> sonda de horquilla	
65		

<220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 5
 <220>
 <221> horquilla (stem_loop)
 <222> (26)..(31)
 <223>
 10
 <220>
 <221> unión a sonda (probe_bind)
 <222> (7)..(25)
 <223>
 15
 <400> 16
 gcgagctttt gacaggggt gaaggctcg c 31
 20
 <210> 17
 <211> 31
 <212> DNA
 <213> Secuencia artificial
 25
 <220>
 <223> sonda de horquilla
 <220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 30
 <220>
 <221> horquilla (stem_loop)
 <222> (26)..(31)
 <223>
 35
 <220>
 <221> unión a sonda (probe_bind)
 <222> (7)..(25)
 <223>
 40
 <400> 17
 gcgagctttt gacaggggt gaaggctcg c 31
 45
 <210> 18
 <211> 30
 <212> DNA
 <213> Secuencia artificial
 50
 <220>
 <223> sonda de horquilla
 <220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 55
 <220>
 <221> horquilla (stem_loop) _
 <222> (25)..(30)
 <223>
 60
 <220>
 <221> unión a sonda (probe_bind)
 65

	<222> (7)..(24)	
	<223>	
5	<400> 18 gcgagcccga gtccgtccag agctgctcgc	30
	<210> 19	
	<211> 30	
	<212> DNA	
10	<213> Secuencia artificial	
	<220>	
	<223> sonda de horquilla	
15	<220>	
	<221> horquilla (stem_loop)	
	<222> (1)..(6)	
	<223>	
20	<220>	
	<221> horquilla (stem_loop)	
	<222> (25)..(30)	
	<223>	
25	<220>	
	<221> unión a sonda (probe_bind)	
	<222> (7)..(24)	
	<223>	
30	<400> 19 gcgagcccga gtccatccag agctgctcgc	30
	<210> 20	
	<211> 29	
35	<212> DNA	
	<213> Secuencia artificial	
	<220>	
	<223> sonda de horquilla	
40	<220>	
	<221> horquilla (stem_loop)	
	<222> (1)..(6)	
	<223>	
45	<220>	
	<221> horquilla (stem_loop)	
	<222> (24)..(29)	
	<223>	
50	<220>	
	<221> unión a sonda (probe_bind)	
	<222> (7)..(23)	
	<223>	
55	<400> 20 gcgagccaac cactccacct gtcgctcgc	29
	<210> 21	
60	<211> 29	
	<212> DNA	
	<213> Secuencia artificial	
65	<220>	

<223> sonda de horquilla
 <220>
 <221> horquilla (stem_loop)
 5 <222> (1)..(6)
 <223>
 <220>
 <221> horquilla (stem_loop)
 10 <222> (24)..(29)
 <223>
 <220>
 <221> unión a sonda (probe_bind)
 15 <222> (7)..(23)
 <223>
 <400> 21
 20 gcgagccaac cattccacct gtcgctcgc 29
 <210> 22
 <211> 29
 <212> DNA
 <213> Secuencia artificial
 25 <220>
 <223> sonda de horquilla
 <220>
 30 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 <220>
 35 <221> horquilla (stem_loop)
 <222> (24)..(29)
 <223>
 <220>
 40 <221> unión a sonda (probe_bind)
 <222> (7)..(23)
 <223>
 <400> 22
 45 gcgagccaac cagtccacct gtcgctcgc 29
 <210> 23
 <211> 29
 <212> DNA
 50 <213> Secuencia artificial
 <220>
 <223> sonda de horquilla
 55 <220>
 <221> horquilla (stem_loop)
 <222> (1)..(6)
 <223>
 60 <220>
 <221> horquilla (stem_loop)
 <222> (24)..(29)
 <223>
 65

	<220>	
	<221> unión a sonda (probe_bind)	
	<222> (7)..(23)	
	<223>	
5	<400> 23	
	gcgagccaac caatccacct gtcgctcgc	29
	<210> 24	
10	<211> 30	
	<212> DNA	
	<213> Secuencia artificial	
	<220>	
15	<223> sonda de horquilla	
	<400> 24	
	30 gcgagcagta tggctctcgg tcatgctcgc	
20	<210> 25	
	<211> 30	
	<212> DNA	
	<213> Secuencia artificial	
25	<220>	
	<223> sonda de horquilla	
	<400> 25	
30	gcgagcagta tggctcttgg tcatgctcgc	30
	<210> 26	
	<211> 31	
	<212> DNA	
	<213> Secuencia artificial	
35	<220>	
	<223> sonda de horquilla	
	<400> 26	
40	gcgagcgctc cattgtcata cacaagctcg c	31
	<210> 27	
	<211> 31	
	<212> DNA	
45	<213> Secuencia artificial	
	<220>	
	<223> sonda de horquilla	
50	<400> 27	
	gcgagcgctc cattttcata cacaagctcg c	31
	<210> 28	
	<211> 29	
55	<212> DNA	
	<213> Secuencia artificial	
	<220>	
	<223> sonda de horquilla	
60	<400> 28	
	gcgagccaat ggcacgttca cttgctcgc	29
	<210> 29	
65		

	<211> 27	
	<212> DNA	
	<213> Secuencia artificial	
5	<220>	
	<223> sonda de horquilla	
	<400> 29	
10	gcgagccaga actgtagcac cgctcgc	27
	<210> 30	
	<211> 27	
	<212> DNA	
	<213> Secuencia artificial	
15	<220>	
	<223> sonda de horquilla	
	<400> 30	
20	gcgagccaga accgtagcac cgctcgc	27
	<210> 31	
	<211> 26	
	<212> DNA	
25	<213> Secuencia artificial	
	<220>	
	<223> sonda de horquilla	
30	<400> 31	
	gcgacgtcct ggcctgggtg cgctcgc	26
	<210> 32	
	<211> 26	
35	<212> DNA	
	<213> Secuencia artificial	
	<220>	
	<223> sonda de horquilla	
40	<400> 32	
	gcgacgtcct ggcccgggtg cgctcgc	26
	<210> 33	
45	<211> 30	
	<212> DNA	
	<213> Secuencia artificial	
	<220>	
50	<223> diana del oligonucleótido	
	<400> 33	
	atgaagatga ccgagagcca tactcggcca	30
55	<210> 34	
	<211> 30	
	<212> DNA	
	<213> Secuencia artificial	
60	<220>	
	<223> diana del oligonucleótido	
	<400> 34	
65		

	tctaggtgt gtatgacaat ggagccatcc	30
	<210> 35	
	<211> 30	
5	<212> DNA	
	<213> Secuencia artificial	
	<220>	
	<223> diana del oligonucleótido	
10	<400> 35	
	cgtcacaagt gaacgtgcc ttgtcactgt	30
	<210> 36	
15	<211> 26	
	<212> DNA	
	<213> Secuencia artificial	
	<220>	
20	<223> diana del oligonucleótido	
	<400> 36	
	gaagaggtgc tacagttctg ccagag	26
25	<210> 37	
	<211> 26	
	<212> DNA	
	<213> Secuencia artificial	
30	<220>	
	<223> diana del oligonucleótido	
	<400> 37	
35	gaaggccacc caggccagga tgtagg	26
	<210> 38	
	<211> 30	
	<212> DNA	
	<213> Secuencia artificial	
40	<220>	
	<223> diana del oligonucleótido	
	<400> 38	
45	atgaagatga ccaagagcca tactcgcca	30
	<210> 39	
	<211> 30	
	<212> DNA	
50	<213> Secuencia artificial	
	<220>	
	<223> diana del oligonucleótido	
55	<400> 39	
	tctaggtgt gtatgaaaat ggagccatcc	30
	<210> 40	
	<211> 26	
60	<212> DNA	
	<213> Secuencia artificial	
	<220>	
65		

	<223> diana del oligonucleótido	
	<400> 40	
5	gaagaggtgc tacggtctg ccagag	26
	<210> 41	
	<211> 26	
	<212> DNA	
10	<213> Secuencia artificial	
	<220>	
	<223> diana del oligonucleótido	
	<400> 41	
15	gaaggccacc cgggccagga tgtagg	26
20		

REIVINDICACIONES

1. Kit que comprende:
 - (i) un biochip de sondas que comprende un soporte y al menos dos sondas sin marcar, fijadas al soporte, que comprenden
 - 5 (a) una secuencia específica de una diana, de 6 a 30 nucleótidos de longitud;
 - (b) un primer brazo de menos de 10 nucleótidos de longitud, situado en 5' de la secuencia específica de una diana; y
 - (c) un segundo brazo de menos de 10 nucleótidos de longitud, situado en 3' de la secuencia específica de una diana,
 - 10 en donde la secuencia específica de la diana no es complementaria de ninguna otra parte de la sonda; en donde el primer brazo y el segundo brazo son perfectamente complementarios entre sí; en donde los primeros brazos de dichas al menos dos sondas poseen una secuencia idéntica y las secuencias específicas de la diana de dichas al menos dos sondas, son diferentes;
 - 15 (ii) una molécula "informadora" que comprende un marcador detectable y menos de 10 nucleótidos perfectamente complementarios con la secuencia de ácido nucleico del primer o del segundo brazo de la sonda, en donde dicha molécula "informadora" es capaz de hibridarse con el primer o el segundo brazo de la sonda cuando una molécula diana se hibrida con la secuencia específica de la diana y la sonda adopta una conformación "abierta"; y
 - (iii) uno o varios reactivos adicionales.
- 20 2. Kit de acuerdo con la reivindicación 1, caracterizado porque la secuencia específica de la diana tiene una longitud de 10-25 nucleótidos, preferentemente de 15-20 nucleótidos.
3. Kit de acuerdo con una de las reivindicaciones 1 a 2, caracterizado porque la Dtm (diferencia entre la temperatura de fusión (Tm) del híbrido perfecto formado entre la secuencia específica de la diana de la sonda y la molécula diana, y la temperatura de fusión (Tm) del híbrido formado entre el primer brazo y el segundo brazo de la sonda), es superior a 10, preferentemente igual a 15.
- 25 4. Kit de acuerdo con una de las reivindicaciones 1 a 2, caracterizado porque la Dtm (diferencia entre la temperatura de fusión (Tm) del híbrido perfecto formado entre la secuencia específica de la diana de la sonda y la molécula diana, y la temperatura de fusión (Tm) del híbrido formado entre el primer brazo y el segundo brazo de la sonda), es inferior a 10.
- 30 5. Kit de acuerdo con una de las reivindicaciones 1 a 4, caracterizado porque dicha sonda comprende además un espaciador y está fijada a dicho soporte con dicho espaciador.
6. Kit de acuerdo con una de las reivindicaciones 1 a 5, caracterizado porque el soporte consiste en una superficie de vidrio funcionalizada, una superficie de plástico funcionalizada, una superficie metálica funcionalizada, una superficie conductora de metal, una superficie conductora de plástico, un soporte poroso, un metal poroso, una fibra óptica, un soporte obtenido a partir de fibra de vidrio, dióxido de silicio, una membrana lipídica funcional, un liposoma, una membrana de filtración o perlas o partículas de tamaño inferior a una micra.
- 35 7. Kit de acuerdo con una de las reivindicaciones 1 a 6, caracterizado porque una misma molécula "informadora" se puede hibridar con cada sonda.
8. Kit de acuerdo con una de las reivindicaciones 1 a 7, caracterizado porque todos los híbridos perfectos formados entre las secuencias específicas de la diana de las sondas y las moléculas diana tienen una temperatura de fusión igual, dentro de un intervalo de 4°C, preferentemente dentro de un intervalo de 1°C.
- 40 9. Kit de acuerdo con una de las reivindicaciones 1 a 8, caracterizado porque la diferencia entre la temperatura de fusión del híbrido formado entre la secuencia específica de la diana de cada sonda y la molécula diana, y la temperatura de fusión para la asociación entre la secuencia específica de la diana y una molécula para la que no se ha diseñado la secuencia específica de la diana de la sonda, es superior o igual a 5°C, preferentemente superior o igual a 8°C.
- 45 10. Kit de acuerdo con una de las reivindicaciones 1 a 9, caracterizado porque las Dtm de al menos dos sondas (diferencia entre la temperatura de fusión (Tm) del híbrido perfecto formado entre la secuencia específica de la diana de la sonda y la molécula diana, y la temperatura de fusión (Tm) del híbrido formado entre el primer brazo y el segundo brazo de la sonda), son iguales, dentro de un intervalo de 1°C.
- 50

11. Kit de acuerdo con una de las reivindicaciones 1 a 10, que comprende además un conjunto de sondas sin marcar y sin inmovilizar que comprende:

- (a) una secuencia específica de una diana, de 6 a 30 nucleótidos de longitud;
- (b) un primer brazo de menos de 10 nucleótidos de longitud, situado en 5' de la secuencia específica de una diana; y
- (c) un segundo brazo de menos de 10 nucleótidos de longitud, situado en 3' de la secuencia específica de una diana,

en donde la secuencia específica de la diana no es complementaria de ninguna otra parte de la sonda; el primer brazo y el segundo brazo son completamente complementarios entre sí.

12. Kit de acuerdo con una de las reivindicaciones 1 a 11, caracterizado porque el marcador detectable es un análogo de nucleótido, un marcador fluorescente, biotina, iminobiotina, un antígeno, un cofactor, dinitrofenol, ácido lipoico, un compuesto olefínico, un polipéptido, una molécula rica en electrones, una enzima o un radioisótopo.

13. Método de detección de ácido nucleico que comprende:

- (a) poner en contacto *ex vivo* una muestra de ácido nucleico con un biochip tal como el definido en el kit de acuerdo con una de las reivindicaciones 1 a 12; y
- (b) detectar una señal que proviene de al menos una sonda del biochip que ha adoptado una conformación abierta después del contacto establecido durante la etapa (a) con ayuda de una molécula "informadora" tal como se define en el kit de acuerdo con una de las reivindicaciones 1 a 12.

14. Método de detección *ex vivo* de una variante genética en una muestra de ácido nucleico que comprende:

- (a) poner en contacto la muestra con un biochip tal como el definido en el kit de acuerdo con una de las reivindicaciones 1 a 12, caracterizado porque al menos una sonda del biochip es una sonda específica de una variante genética, y
- (b) detectar una señal que proviene de la sonda específica de la variante genética con ayuda de una molécula "informadora" tal como se define en el kit de acuerdo con una de las reivindicaciones 1 a 12, en donde la señal detectada en la etapa (b) indica la presencia de una variante genética en la muestra de ácido nucleico.

15. Método de acuerdo con la reivindicación 14, caracterizado porque la variante genética detectada es un polimorfismo de un solo nucleótido.

16. Método de detección *ex vivo* de cualquier organismo que contenga ácido nucleico o de cualquier resto de esos organismos, que comprende:

- (a) poner en contacto una muestra de ácido nucleico con un biochip tal como el definido en el kit de acuerdo con una de las reivindicaciones 1 a 12, en donde al menos una de dichas sondas es una sonda específica de ácido nucleico de un organismo o de restos de ese organismo; y
- (b) detectar una señal que proviene de la sonda específica de ácido nucleico de un organismo con ayuda de una molécula "informadora" tal como se define en el kit de acuerdo con una de las reivindicaciones 1 a 12,

en donde la señal detectada en la etapa (b) indica la presencia del organismo que contiene el ácido nucleico o restos de ese organismo.

17. Método de acuerdo con la reivindicación 16, caracterizado porque el organismo que contiene los ácido nucleico es un virus o una bacteria.

18. Método de detección *ex vivo* de un producto del corte y empalme alternativo de un gen en una muestra de ácido nucleico, que comprende:

- (a) poner en contacto la muestra con un biochip tal como se ha definido en el kit de acuerdo con una de las reivindicaciones 1 a 12, caracterizado porque al menos una sonda del biochip es una sonda de horquilla específica de un exón de un gen o de la unión de dos exones; y
- (b) detectar una señal que proviene de la sonda específica del exón o de la unión de dos exones, con ayuda de una molécula "informadora" tal como se ha definido en el kit de acuerdo con una de las reivindicaciones 1 a 12, en donde la señal detectada en la etapa (b) indica la presencia del producto de

corte y empalme alternativo de un gen en la muestra de ácido nucleico.

19. Método de acuerdo con la reivindicación 18, caracterizado porque la muestra de ácido nucleico comprende ARNm o ADNc.

20. Un método de secuenciación ex vivo de un oligonucleótido que comprende:

- 5 (a) poner en contacto una muestra que contiene el oligonucleótido con un biochip tal como se ha definido en el kit de acuerdo con una de las reivindicaciones 1 a 12; y
- (b) detectar una señal que proviene de al menos una sonda del biochip con ayuda de una molécula "informadora" tal como se ha definido en el kit de acuerdo con una de las reivindicaciones 1 a 12;

en donde la señal detectada en la etapa (b) se utiliza para determinar la secuencia del oligonucleótido.

10 21. Método de acuerdo con una de las reivindicaciones 13 a 20, en donde la muestra de la etapa (a) está marcada previamente con un marcador detectable.

22. Método de detección ex vivo de desequilibrios alélicos y pérdidas de heterocigosidad en una muestra de ácido nucleico que comprende:

- 15 (a) amplificar al menos una región de ADN cromosómico de tipo microsatélite mediante una pareja de cebadores a partir de al menos dos muestras de ácido nucleico procedentes de tejidos o fluidos biológicos, en donde al menos una de las muestras procede de células o de tejidos que no presentan ningún desequilibrio alélico o pérdida de heterocigosidad, y en donde cada tejido o fluido está marcado de manera diferente durante la amplificación;
- (b) eliminar esos cebadores después de la amplificación;
- 20 (c) poner en contacto dichos productos de la amplificación con un biochip tal como se ha definido en el kit de acuerdo con una de las reivindicaciones 1 a 12, en donde al menos una sonda del biochip es complementaria de un cebador utilizado para amplificar dicha región de ADN cromosómico; y
- (d) detectar señales procedentes de al menos una sonda de dicho biochip con ayuda de una molécula "informadora" tal como la que se ha definido en el kit de acuerdo con una de las reivindicaciones 1 a 12;

25 en donde las señales detectadas en la etapa (d) se utilizan para determinar la presencia de un desequilibrio alélico o una pérdida de heterocigosidad en una de las muestras de ácido nucleico.

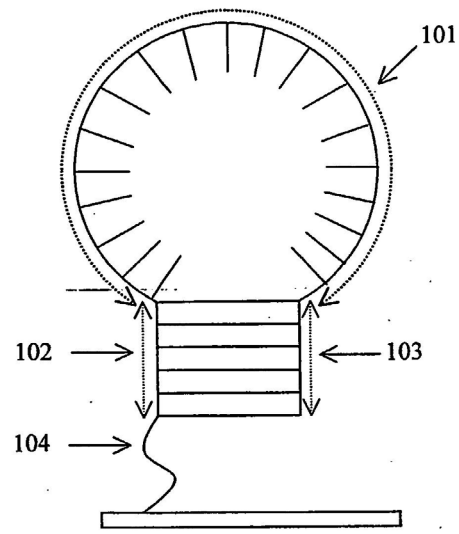


FIG. 1A

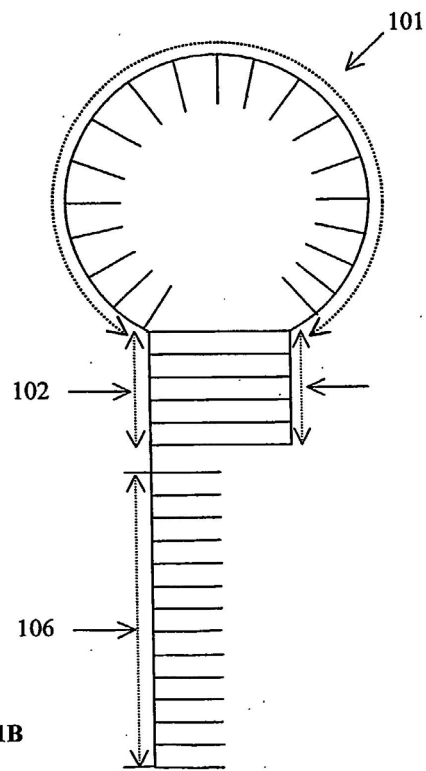


FIG. 1B

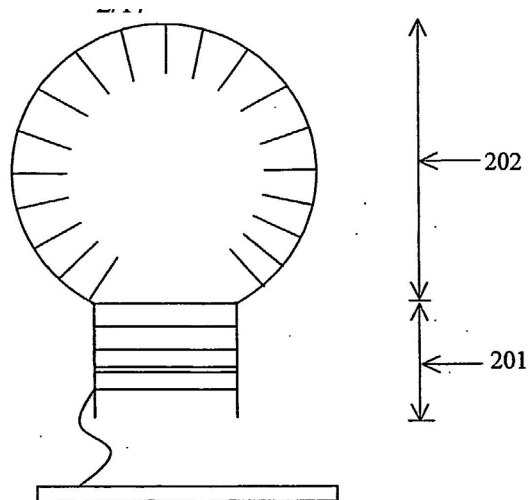


FIG. 2A

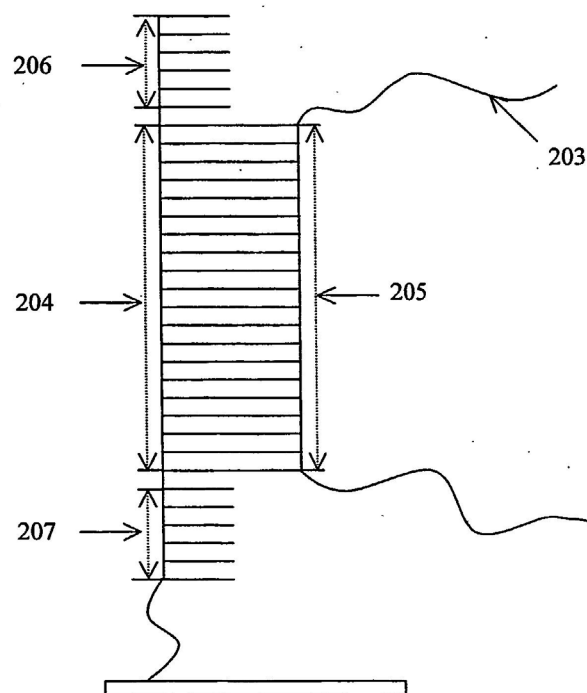


FIG. 2B

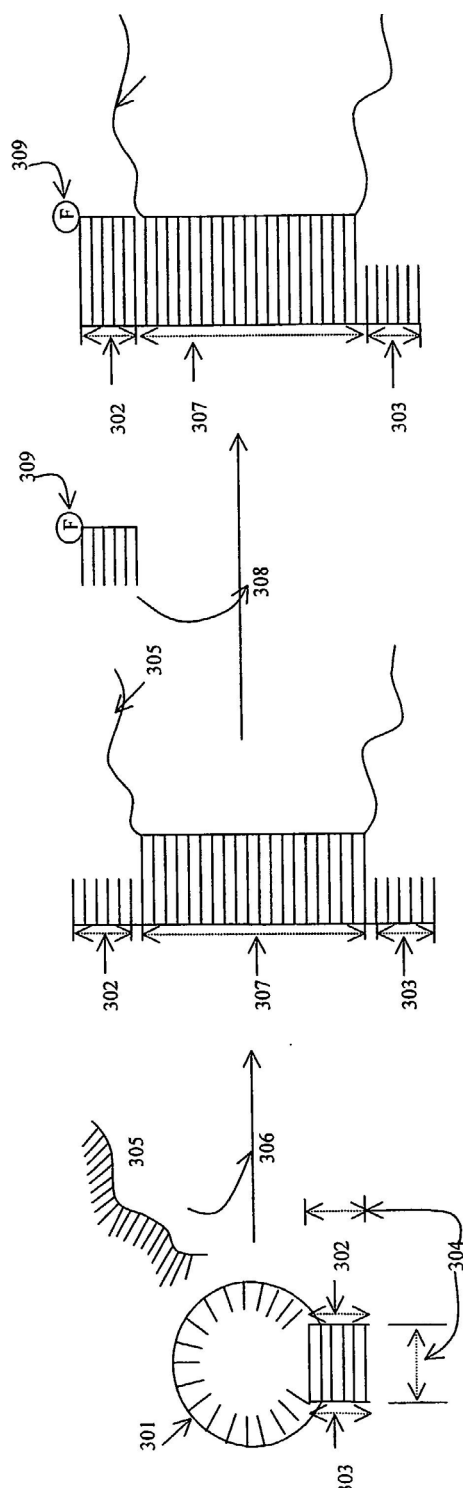


FIG. 3A

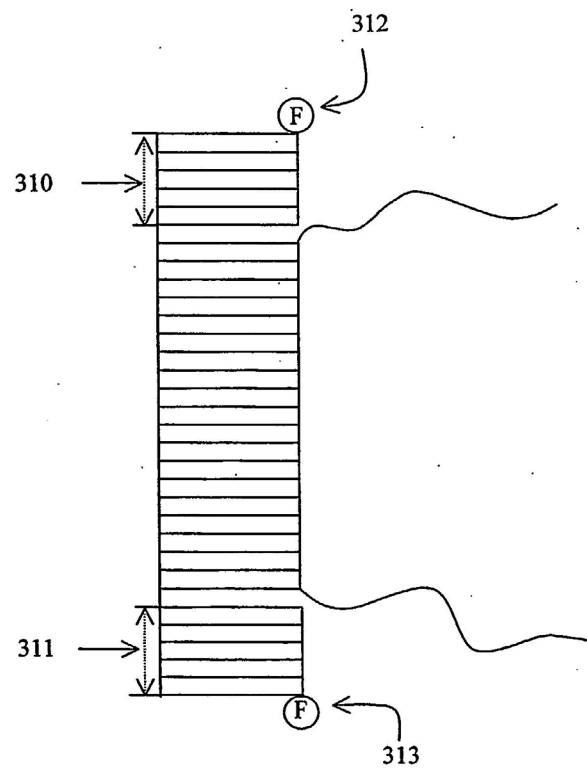


FIG. 3B

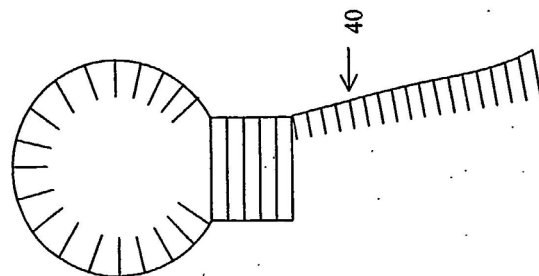


FIG. 4A

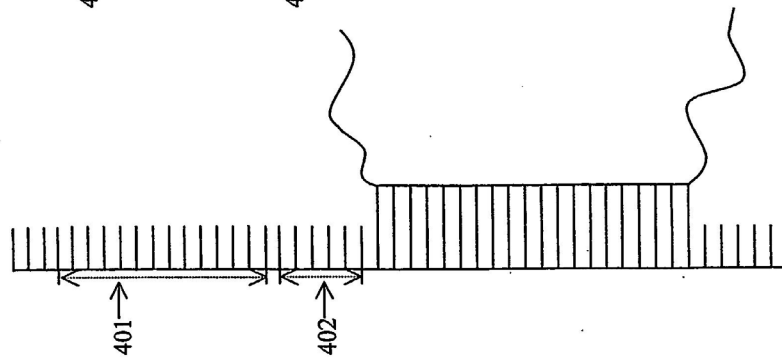


FIG. 4B

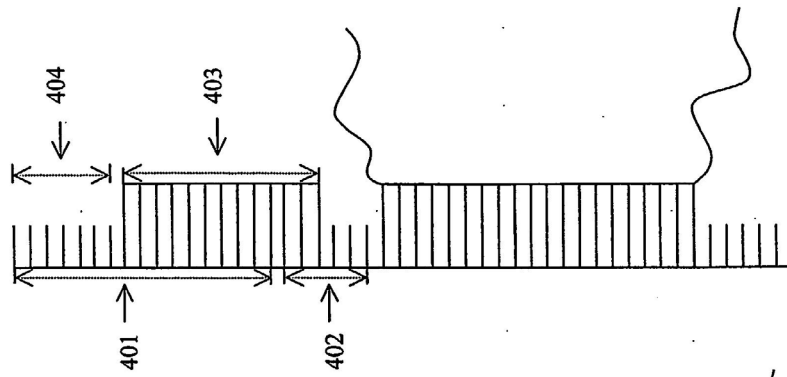


FIG. 4C

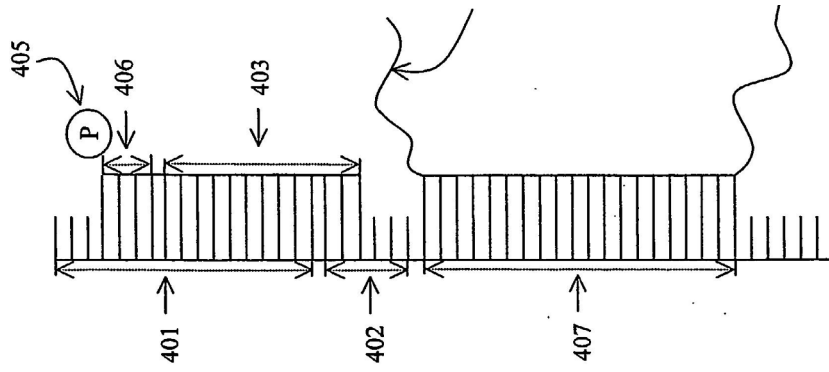
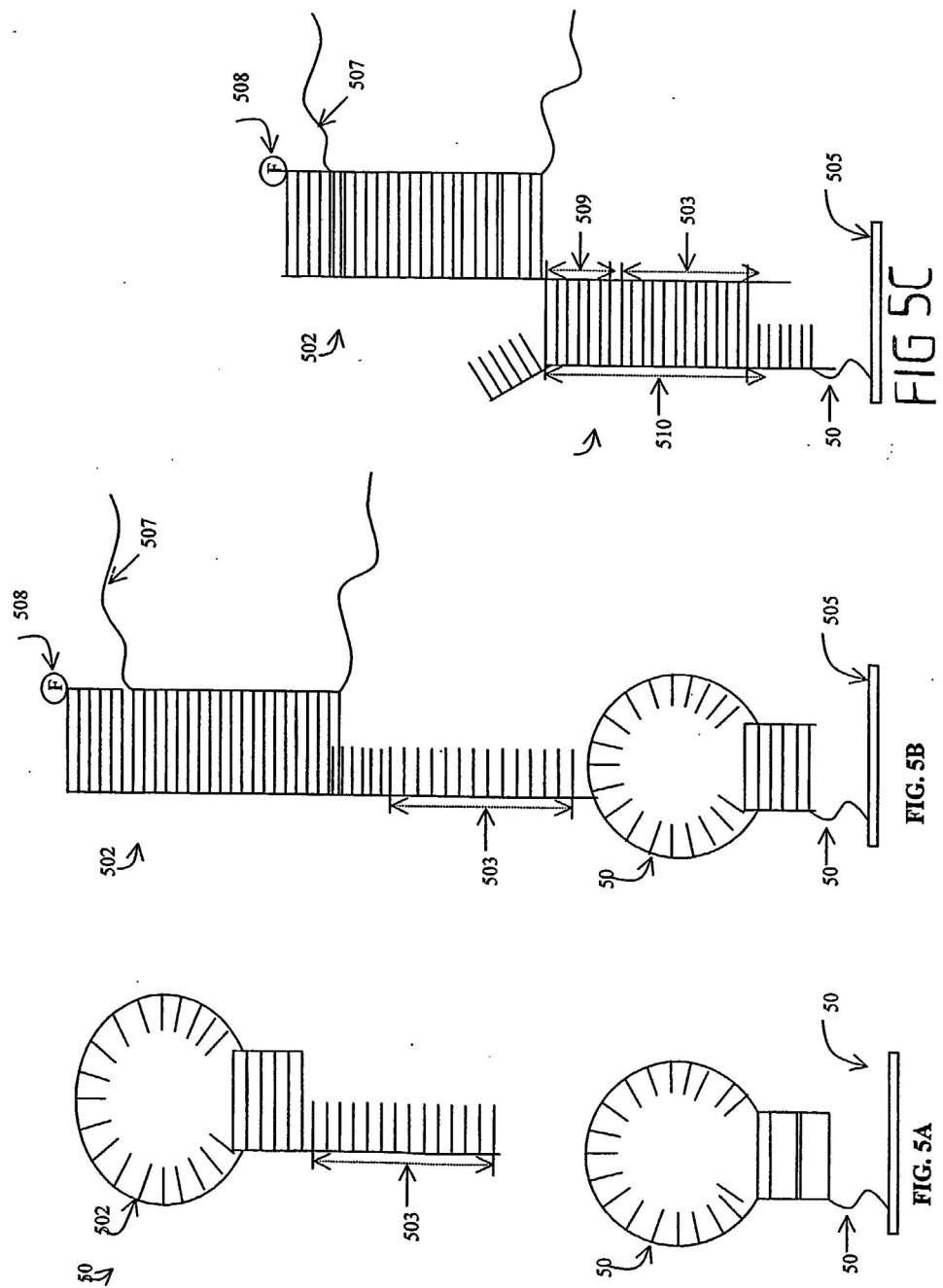


FIG. 4D



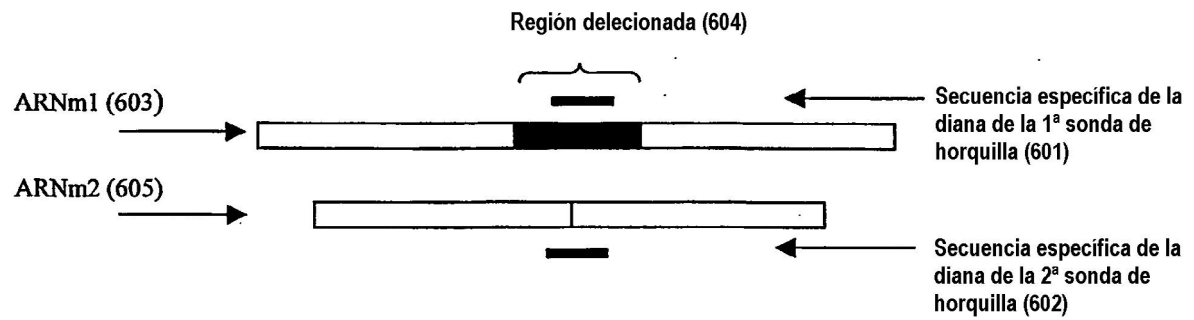


FIG. 6

- A) 241 atcaagggtt acatgcagca acacaacatc cccag [aggg aggtgtgtga tgtacaggcc tgaaccagt cgcaccttc cagcatctc aacaaggga cccctatgaa] gaccagaaag 360 (SEQ. NO.7) secuencia entre corchetes en negrita
- B) 241 atcaagggtt acatgcagca acacaacatc cccag | gacc cagaagcgtg 290 (SEQ. NO.8)
- C) Tg2X2C137M1: Cy5 -5' -cac gct tct ggg tcc tgg ggg atg- 3' (SEQ. NO.9)
- D) 2X2C137M1: 5' -gcg agc ccc cca gga ccc aga agc gct cgc- 3' (SEQ. NO.10)
- E) Molécula informadora Cy3 -5' -gcgagc -3' (SEQ. NO.11)

FIG. 7

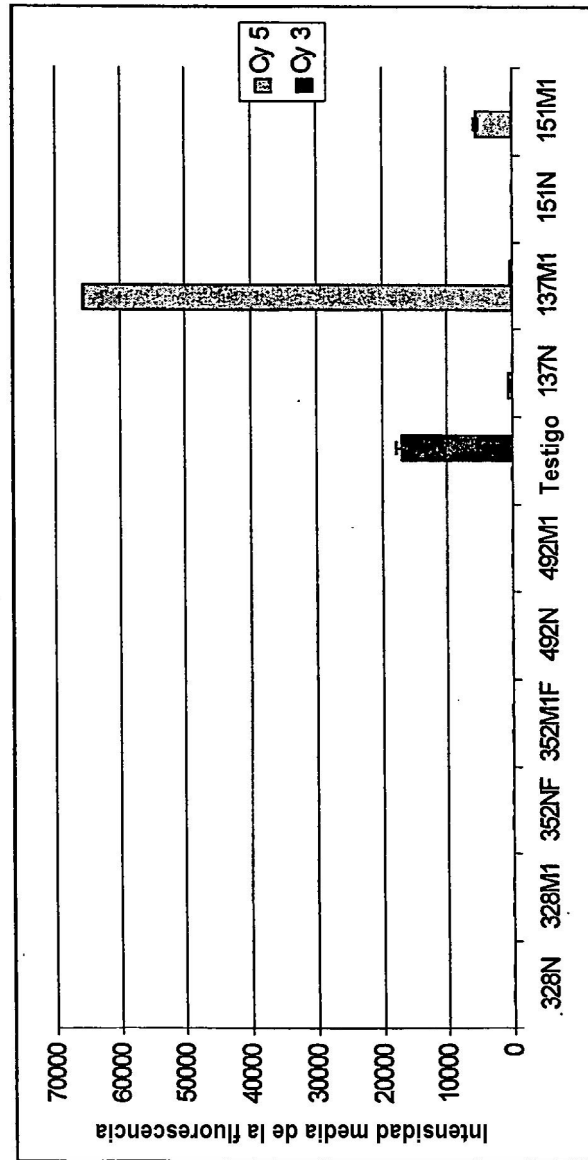


FIG. 8

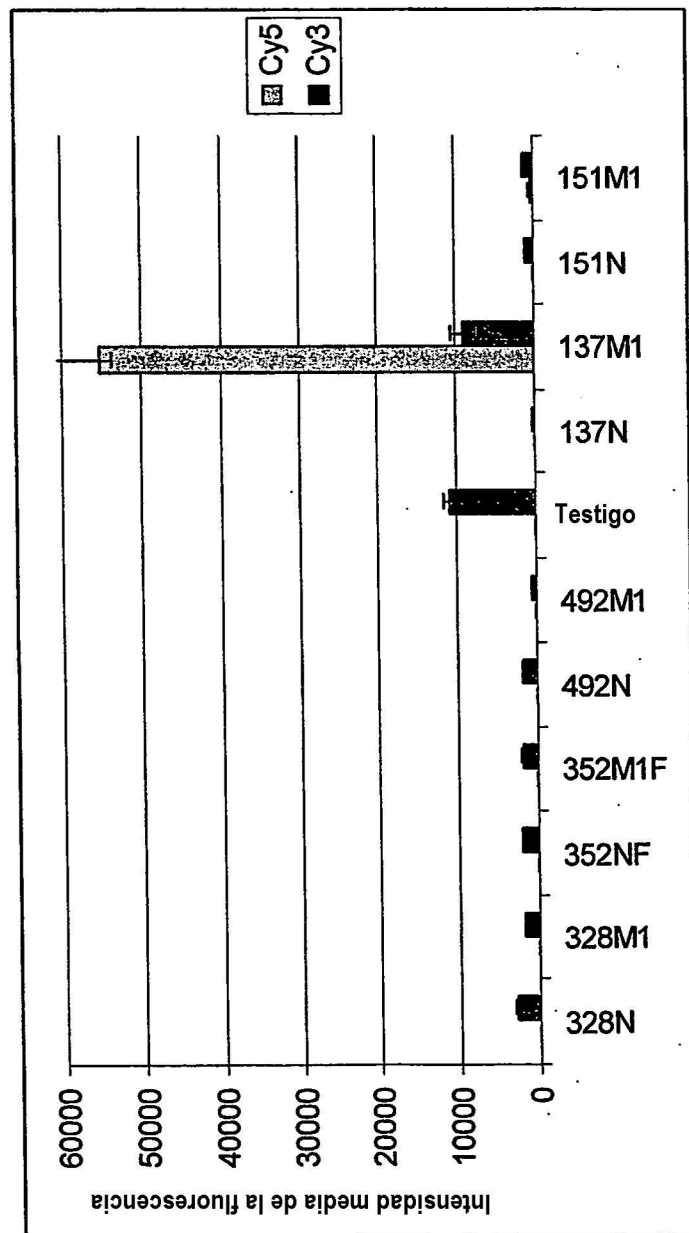


FIG. 9

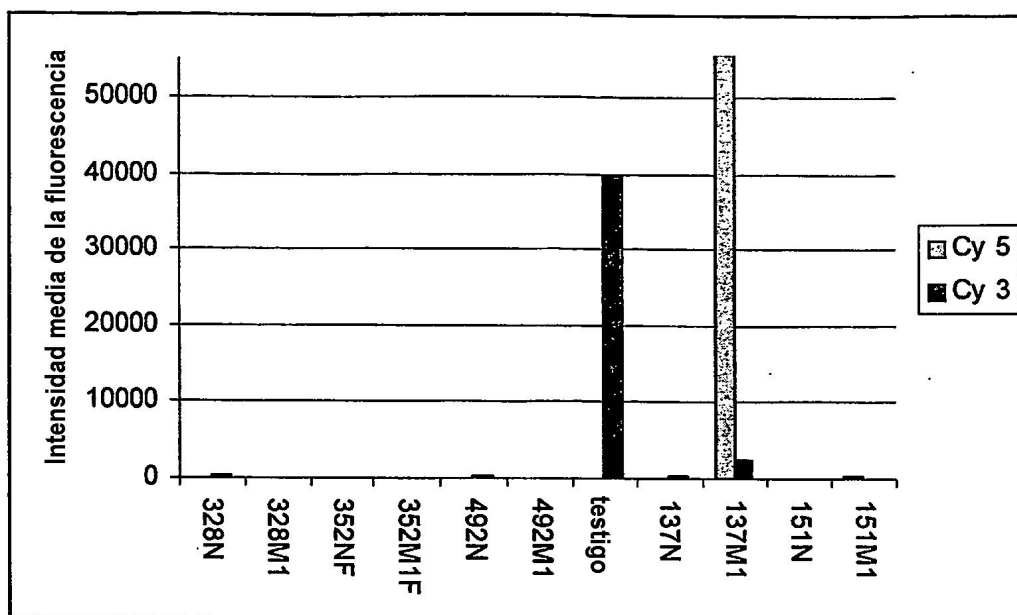


FIG. 10

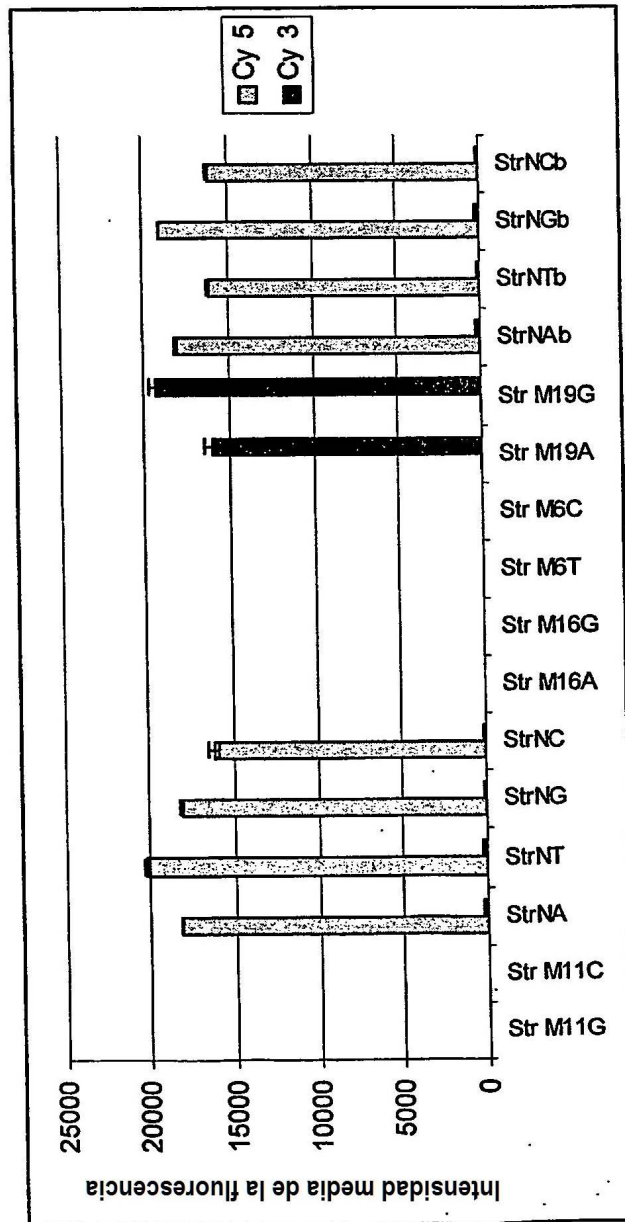


FIG. 11

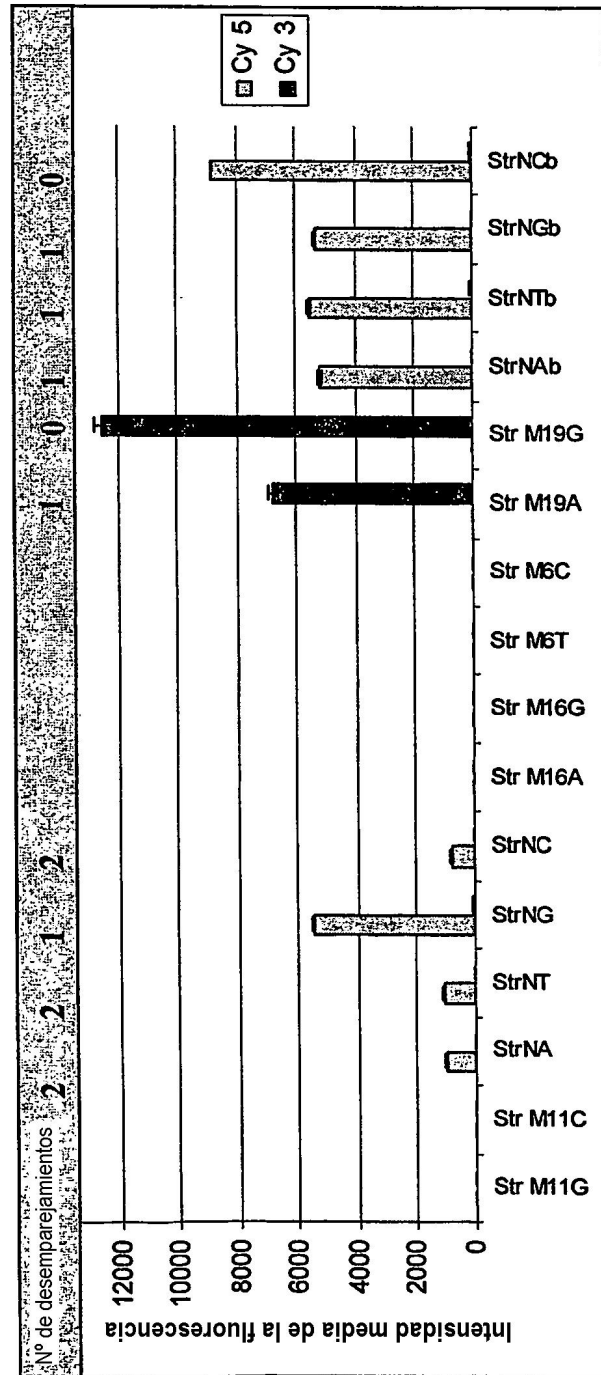


FIG. 12

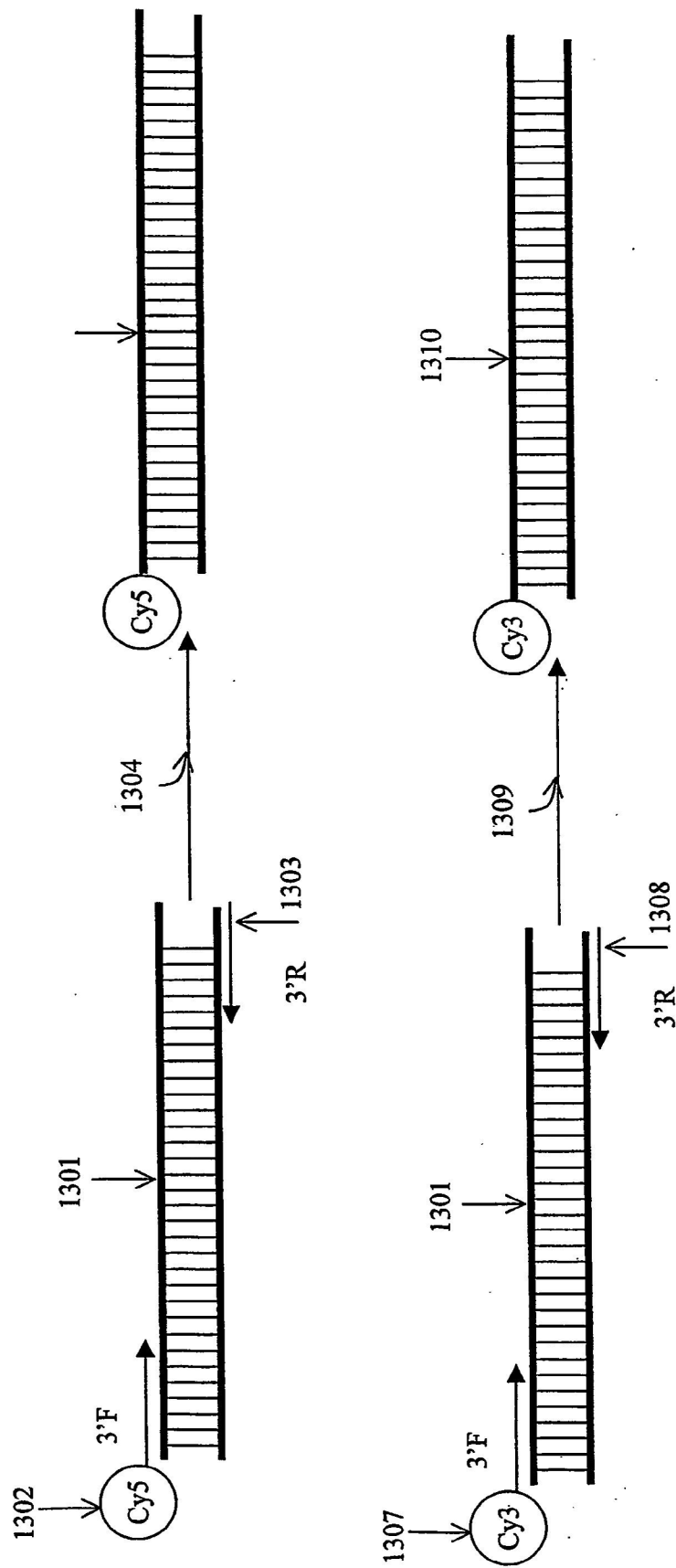


FIG. 13A

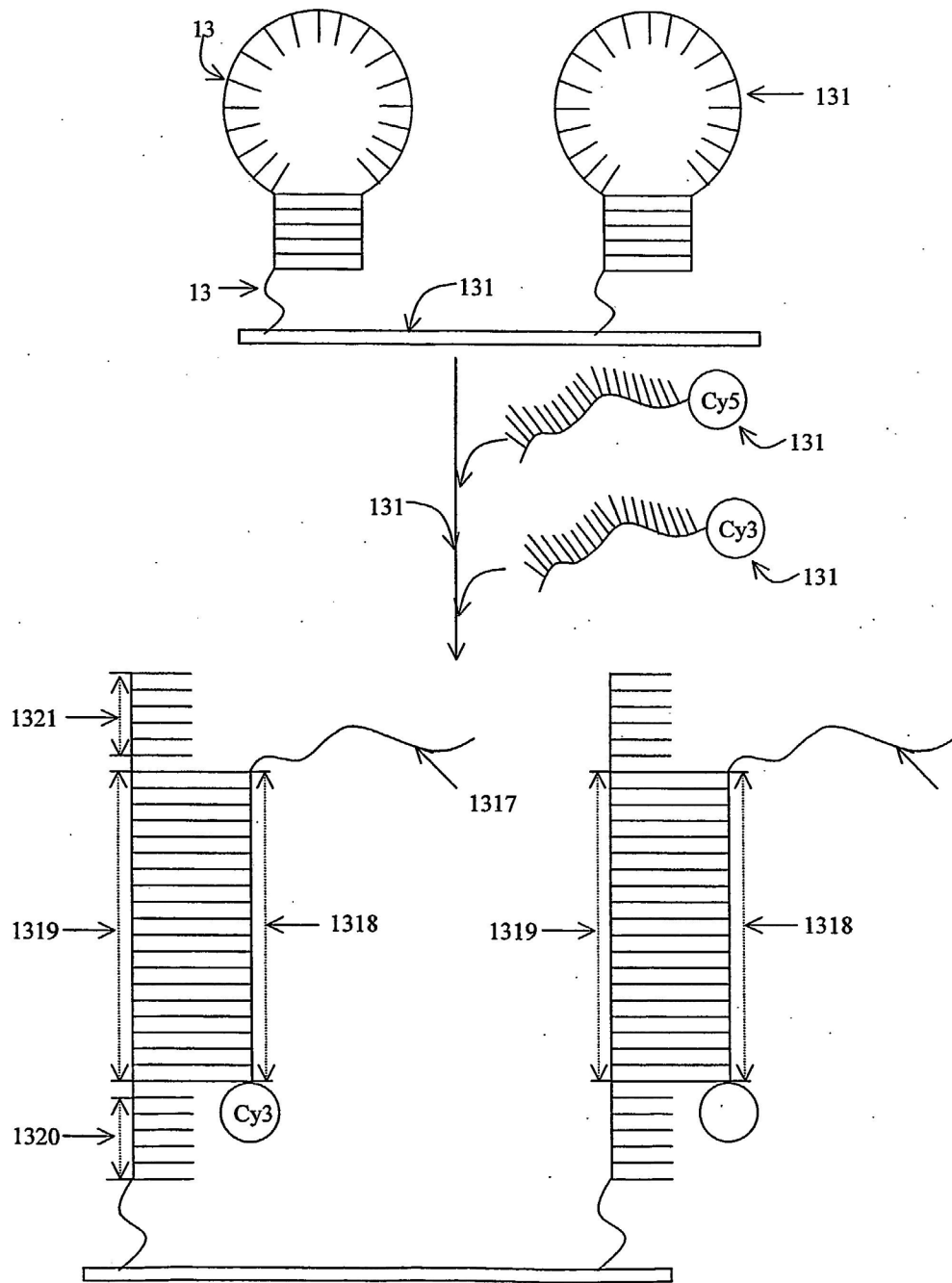


FIG. 13B

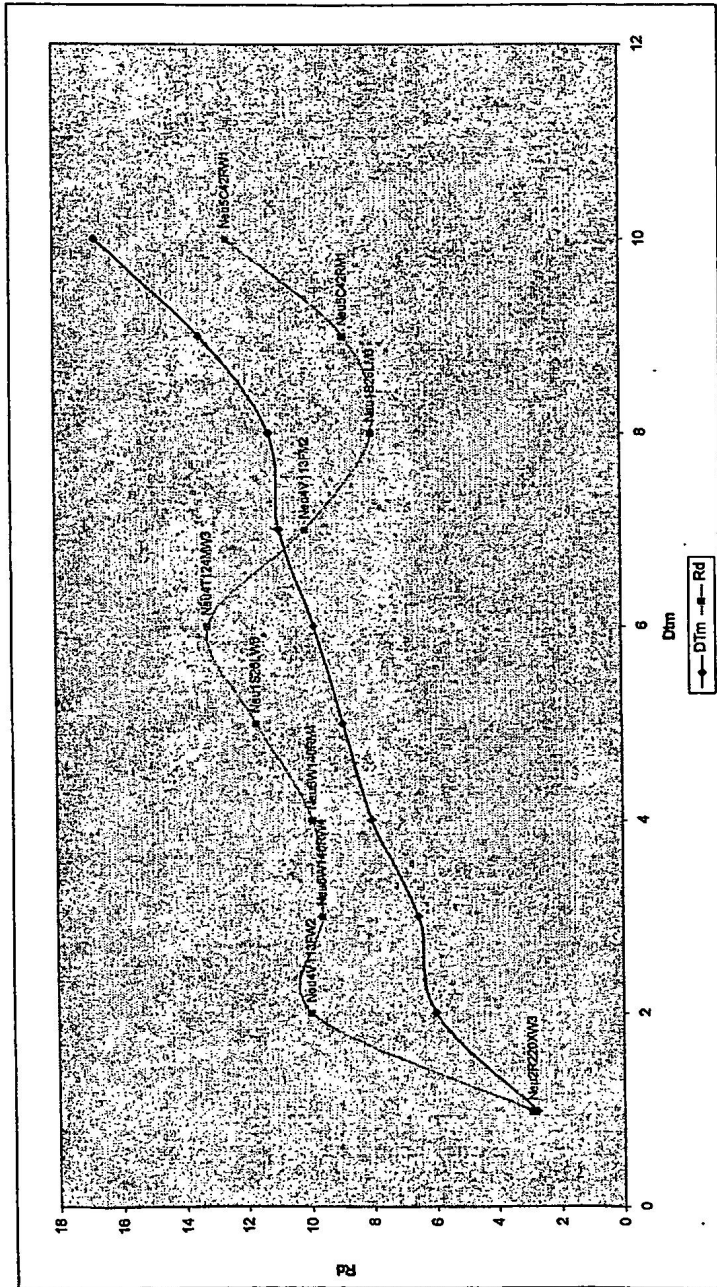


FIG. 14

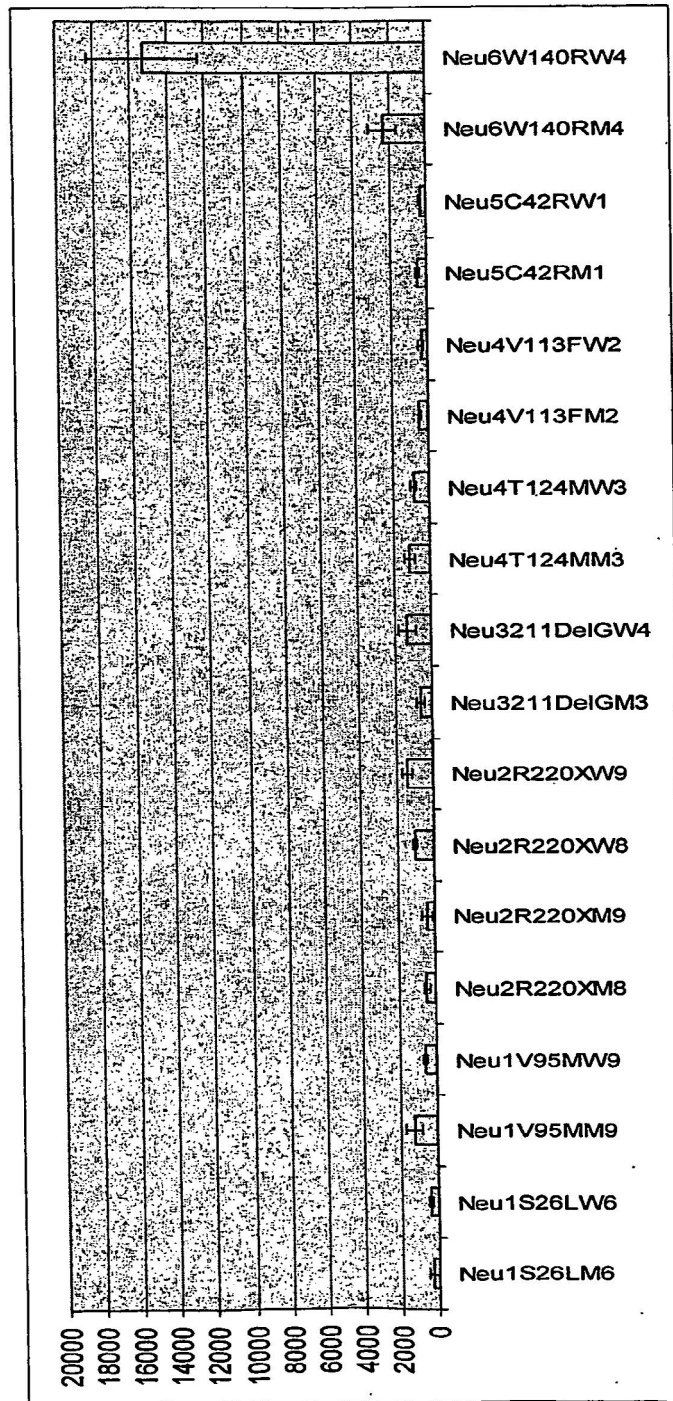


FIG. 15