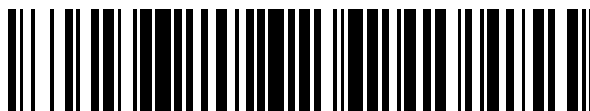


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 690 583**

51 Int. Cl.:

C12Q 1/683 (2008.01)

C12Q 1/6858 (2008.01)

C12Q 1/6886 (2008.01)

C12Q 1/6883 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **31.07.2015 PCT/IB2015/055799**

87 Fecha y número de publicación internacional: **18.02.2016 WO16024182**

96 Fecha de presentación y número de la solicitud europea: **31.07.2015 E 15747564 (1)**

97 Fecha y número de publicación de la concesión europea: **26.09.2018 EP 3180441**

54 Título: **Método para estimar la cantidad de un locus metilado en una muestra**

30 Prioridad:

13.08.2014 US 201462037057 P

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

21.11.2018

73 Titular/es:

**VANADIS DIAGNOSTICS (100.0%)
Vetenskapsvägen 10
19138 Sollentuna, SE**

72 Inventor/es:

**DAHL, CARL OSCAR FREDRIK;
ERICSSON, OLOF JOHN y
BANÉR, JOHAN**

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 690 583 T3

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

DESCRIPCIÓN

Método para estimar la cantidad de un locus metilado en una muestra

5 Antecedentes

5-Metilcitosina es una forma metilada de la base de ADN citosina que se cree que está implicada en la regulación transcripcional. Cuando la citosina está metilada, el ADN mantiene la misma secuencia, pero la expresión de genes metilados puede alterarse.

10 La función de este químico varía significativamente entre las especies: en las bacterias, la 5-metilcitosina se puede encontrar en una variedad de sitios, y se usa a menudo como un marcador para proteger el ADN de ser cortado por enzimas de restricción nativas sensibles a la metilación; en las plantas, la 5-metilcitosina se produce en las secuencias CpG, CpHpG y CpHpH (donde H = A, C o T); y, en hongos y animales, la 5-metilcitosina se produce predominantemente en los dinucleótidos CpG. La mayoría de los eucariotas metilan solo un pequeño porcentaje de estos sitios, pero 70-80% de las citosinas CpG están metiladas en los vertebrados.

20 La metilación de citosina en vertebrados ocurre típicamente en sitios CpG (sitios de citosina-fosfato-guanina, es decir, donde una citosina es seguida directamente por una guanina en la secuencia de ADN). La formación de Me-CpG está catalizada por la enzima ADN metiltransferasa. Aproximadamente el 80-90% de los sitios CpG están metilados en el ADN humano, pero hay ciertas áreas, conocidas como islas CpG, en las que ninguno de los dinucleótidos CpG está metilado. Estos están asociados con los promotores del 56% de los genes de mamíferos, incluidos todos los genes expresados de forma ubicua. De uno a dos por ciento del genoma humano son grupos CpG, y existe una relación inversa entre la metilación de CpG y la actividad transcripcional.

25 El método descrito en este documento proporciona una forma de estimar la cantidad de un locus metilado en una muestra.

30 Resumen

Se proporciona un método para estimar la cantidad de un locus metilado en una muestra. En ciertas realizaciones, el método puede comprender: (a) digerir una muestra de ácido nucleico que contiene copias metiladas y no metiladas de un locus genómico con un miembro de la familia MspJI para producir una población de fragmentos que están en el intervalo de 20-40 pares de bases de longitud y tienen una citosina metilada central; (b) ligar la secuencia adaptadora A y la secuencia adaptadora B a los extremos respectivos de un fragmento diana de la secuencia X mediante: (i) hibridar un oligonucleótido puente de fórmula B'-X'-A' a los fragmentos de (a) en la presencia de las secuencias adaptadoras A y B, en las que X', A' y B' son complementarias a X, A y B, respectivamente, y (ii) ligando la secuencia adaptadora A, la secuencia X y la secuencia adaptadora B entre sí para producir un producto de fórmula A-X-B; y (c) cuantificar la cantidad de productos de ligación de fórmula A-X-B, proporcionando de ese modo una estimación de la cantidad del locus metilado en la muestra de ácido nucleico.

45 La cuantificación se puede hacer usando cualquier método conveniente, incluso mediante qPCR o mediante secuenciación. En algunas realizaciones, las secuencias adaptadoras A y B pueden estar presentes en la misma molécula de oligonucleótido y el producto de la etapa (b) es una molécula circular de ácido nucleico. En estas realizaciones, la cuantificación puede realizarse al (i) amplificar la molécula circular de ácido nucleico mediante amplificación por círculo rodante (RCA), (ii) hibridando el producto RCA con una población de oligonucleótidos marcados que se hibridan con múltiples posiciones en el producto RCA; y (iii) contar individualmente el número de complejos RCA marcados.

50 Como se discutirá con mayor detalle a continuación, en ciertos casos, el método puede usarse para estimar la cantidad de ADN fetal en una muestra de ADN libre de células de una hembra embarazada.

También se proporciona un kit para realizar el método.

55 Breve descripción de las figuras

El experto en la materia entenderá que los dibujos, que se describen a continuación, solo tienen fines ilustrativos. Los dibujos no pretenden limitar el alcance de las presentes enseñanzas de ninguna manera.

60 La fig. 1 ilustra esquemáticamente algunas de las características de una realización del método objeto.

La fig. 2 ilustra esquemáticamente algunas de las características de una implementación del método objeto.

La fig. 3 ilustra esquemáticamente algunas de las características de otra implementación del método objeto.

65 La fig. 4 ilustra esquemáticamente un método mediante el cual pueden contarse los productos de ligación.

Definiciones

5 Antes de describir las realizaciones a modo de ejemplo con mayor detalle, se establecen las siguientes definiciones para ilustrar y definir el significado y alcance de los términos usados en la descripción.

10 Los rangos numéricos incluyen los números que definen el rango. A menos que se indique lo contrario, los ácidos nucleicos se escriben de izquierda a derecha en una orientación de 5' a 3'; las secuencias de aminoácidos se escriben de izquierda a derecha en la orientación de amino a carboxi, respectivamente.

15 A menos que se defina lo contrario, todos los términos técnicos y científicos usados en la presente memoria tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta invención. Singleton, et al., *DICTIONARY OF MICROBIOLOGY AND MOLECULAR BIOLOGY*, 2D ED., John Wiley and Sons, New York (1994), and Hale & Markham, *THE HARPER COLLINS DICTIONARY OF BIOLOGY*, Harper Perennial, N.Y. (1991) le proporciona a un experto el significado general de muchos de los términos utilizados en este documento. Aun así, ciertos términos se definen a continuación por razones de claridad y facilidad de referencia.

20 Debe observarse que tal como se usa en el presente documento y en las reivindicaciones adjuntas, las formas singulares “un”, “una” y “el” incluyen referentes en plural a menos que el contexto indique claramente lo contrario. Por ejemplo, el término “un cebador” se refiere a uno o más cebadores, es decir, un cebador único y cebadores múltiples. Se observa además que las reivindicaciones pueden redactarse para excluir cualquier elemento opcional. Como tal, esta declaración tiene la intención de servir como base previa para el uso de terminología exclusiva como “únicamente”, “solo” y similares en relación con la recitación de elementos de reclamo, o el uso de una limitación “negativa”.

25 El documento WO2010/113088 divulga el uso de la detección por ligación para detectar sitios CpG metilados en ubicaciones genómicas específicas. Si el sitio CpG no está metilado, se usa una enzima de restricción sensible a la metilación para cortar en el sitio no metilado, mientras que los sitios metilados permanecen intactos.

30 Como se usa en este documento, el término “locus genómico” se refiere a una región definida en un genoma. Un locus genómico existe en la misma ubicación en los genomas de diferentes células del mismo individuo o en diferentes individuos. Un locus genómico en una célula o individuo tiene una secuencia de nucleótidos que es idéntica o muy similar (es decir, más del 99% idéntica) al mismo locus genómico en una célula o individuo diferente. La diferencia en la secuencia de nucleótidos entre el mismo locus en diferentes células o individuos puede deberse a una o más sustituciones de nucleótidos. Un locus genómico puede definirse por coordenadas genómicas, por nombre o usando un símbolo.

35 Como se usa en el presente documento, el término “estado de metilación” se refiere a la presencia o ausencia de un grupo metilo en un residuo de citosina en un sitio de metilación. Para mayor claridad, una citosina que no está metilada se denominará “citosina no metilada” o “C no metilada”, y una citosina que está metilada (es decir, 5-metilcitosina) se denominará “citosina metilada”, “C” metilada o “C metilo”.

40 Como se usa en el presente documento, un “sitio de metilación” se refiere a la posición de un nucleótido de citosina que se sabe que al menos a veces está metilado en un locus genómico. La citosina en un sitio de metilación puede ser una citosina no metilada o una citosina metilada. En otras palabras, el término “sitio de metilación” se refiere a una citosina específica en un locus genómico que puede estar en un estado metilado. El sitio de metilación puede definirse mediante coordenadas genómicas, o coordenadas relativas al codón de inicio de un gen, por ejemplo.

45 El término “corresponde a” y equivalentes gramaticales, por ejemplo, “correspondiente”, tal como se usa en el presente documento, se refiere a una relación específica entre los elementos a los que se refiere el término. Por ejemplo, un oligonucleótido que corresponde a una secuencia en un ácido nucleico más largo contiene la misma secuencia de nucleótidos que o es complementaria a una secuencia de nucleótidos en el ácido nucleico.

50 En el contexto de un nucleótido en un oligonucleótido que corresponde a un sitio de metilación o un nucleótido en un oligonucleótido que corresponde a una citosina metilada, el término “corresponde a” y equivalentes gramaticales de los mismos están destinados a identificar el nucleótido que se coloca correspondientemente en relación a (es decir, situado frente a) un sitio de metilación cuando los dos ácidos nucleicos (por ejemplo, un oligonucleótido y ADN genómico que contiene una citosina metilada) están alineados o emparejados por bases. De nuevo, a menos que se indique lo contrario (por ejemplo, en el caso de un nucleótido que “no se parea por base” o “pares base” con un residuo particular) un nucleótido que “corresponde a” un sitio de pares de bases de metilación con un sitio metilado o un sitio no metilado. Para mayor claridad, en un oligonucleótido, un nucleótido G o C en una posición que corresponde a una citosina metilada en una secuencia, por ejemplo, un locus genómico, puede: a) un par base con una citosina metilada en la secuencia, b) un par base con una citosina que corresponde posicionalmente a la citosina metilada en una versión amplificada de la secuencia, o c) un par base con un residuo G que es complementario de dicha citosina en una secuencia amplificada.

Como se usa en este documento, una “secuencia que está metilada” es una secuencia de nucleótidos que contiene un sitio de metilación, es decir, un nucleótido de citosina que se sabe que al menos a veces está metilado.

5 Como se usa en el presente documento, el término “no metilado”, con referencia a una secuencia de nucleótidos, se refiere a las copias de una secuencia que no están metiladas.

10 Como se usa en el presente documento, el término “metilado”, con referencia a una secuencia de nucleótidos, se refiere a copias de una secuencia que contiene 5-metilcitosina. La metilación de un locus genómico puede, por ejemplo, alterar la expresión de una proteína, lo que provoca un cambio fenotípico (por ejemplo, un fenotipo relacionado con el cáncer) en las células que tienen dicho locus metilado. Alternativamente, la metilación de un locus genómico puede ser silenciosa.

15 Una muestra que comprende “copias no metiladas y metiladas de un locus genómico” y equivalentes gramaticales de las mismas, se refiere a una muestra que contiene múltiples moléculas de ADN del mismo locus genómico, donde la muestra contiene copias no metiladas del locus genómico y copias metiladas del mismo locus. En este contexto, el término “copias” no pretende significar que las secuencias fueron copiadas entre sí. Más bien, el término “copias” está destinado a indicar que las secuencias son del mismo locus en diferentes células o individuos. En otras palabras, una muestra contiene una mezcla de moléculas de ácido nucleico que tienen la misma secuencia de nucleótidos, excepto que algunas de las moléculas contienen residuos de citosina metilados.

20 Como se usa en este documento, el término “grado de metilación” se refiere al número relativo, porcentaje o fracción de miembros de una especie de nucleótido diana particular dentro de una muestra que están metilados en comparación con aquellos miembros de esa especie de nucleótido diana particular que no están metilados.

25 El término “mezcla”, como se usa en el presente documento, se refiere a una combinación de elementos, que están intercalados y no en ningún orden particular. Una mezcla es heterogénea y no separable espacialmente en sus diferentes constituyentes. Los ejemplos de mezclas de elementos incluyen una serie de elementos diferentes que se disuelven en la misma solución acuosa y una serie de elementos diferentes unidos a un soporte sólido en posiciones aleatorias (es decir, sin ningún orden particular). Una mezcla no es direccionable. Para ilustrar con un ejemplo, una matriz de polinucleótidos unidos a la superficie espacialmente separados, como se conoce comúnmente en la técnica, no es una mezcla de polinucleótidos unidos a la superficie porque las especies de polinucleótidos unidos a la superficie son espacialmente distintas y la matriz es direccionable.

35 El término “nucleótido” pretende incluir aquellos restos que contienen no solo las bases conocidas de purina y pirimidina, sino también otras bases heterocíclicas que se han modificado. Tales modificaciones incluyen purinas y pirimidinas metiladas, purinas o pirimidinas aciladas, ribosas alquiladas u otros heterociclos. Además, el término “nucleótido” incluye aquellos restos que contienen marcadores de hapteno o fluorescentes y puede contener no solo azúcares convencionales de ribosa y desoxirribosa, sino también otros azúcares. Los nucleósidos o nucleótidos modificados también incluyen modificaciones en el resto de azúcar, por ejemplo, en donde uno o más de los grupos hidroxilo están reemplazados con átomos de halógeno o grupos alifáticos, están funcionalizados como éteres, aminas o similares.

45 El término “ácido nucleico” y “polinucleótido” se usan indistintamente en este documento para describir un polímero de cualquier longitud, por ejemplo, mayor que aproximadamente 2 bases, mayor que aproximadamente 10 bases, mayor que aproximadamente 100 bases, mayor que aproximadamente 500 bases, mayor que 1000 bases, hasta aproximadamente 10,000 bases o más compuestas de nucleótidos, por ejemplo, desoxirribonucleótidos o ribonucleótidos, y pueden producirse enzimáticamente o sintéticamente (por ejemplo, PNA como se describe en la Patente de Estados Unidos N° 5,948.902 y las referencias citadas en este documento) que pueden hibridarse con los ácidos nucleicos de origen natural de una manera específica de secuencia análoga a la de dos ácidos nucleicos naturales, por ejemplo, pueden participar en interacciones de emparejamiento de bases Watson-Crick. Los nucleótidos naturales incluyen guanina, citosina, adenina, timina, uracilo (G, C, A, T y U, respectivamente). El ADN y el ARN tienen una cadena principal de desoxirribosa y azúcar de ribosa, respectivamente, mientras que la columna del PNA está compuesta de unidades repetitivas de N-(2-aminoetil)-glicina unidas por enlaces peptídicos. En PNA, diversas bases de purina y pirimidina están ligadas a la columna principal por enlaces carbonilo de metileno. Un ácido nucleico bloqueado (LNA), a menudo denominado RNA inaccesible, es un nucleótido de RNA modificado. El resto de ribosa de un nucleótido de LNA se modifica con un puente adicional que conecta el oxígeno 2' y el carbono 4'. El puente “bloquea” la ribosa en la conformación 3'-endo (norte), que a menudo se encuentra en los dúplex de forma A. Los nucleótidos de LNA se pueden mezclar con residuos de ADN o ARN en el oligonucleótido siempre que se desee. El término “ácido nucleico no estructurado”, o “UNA”, es un ácido nucleico que contiene nucleótidos no naturales que se unen entre sí con estabilidad reducida. Por ejemplo, un ácido nucleico no estructurado puede contener un residuo G' y un residuo C', donde estos residuos corresponden a formas no naturales, es decir, análogos, de G y C que emparejan bases entre sí con estabilidad reducida, pero conservan una capacidad de par base con residuos C y G naturales, respectivamente. El ácido nucleico no estructurado se describe en el documento US20050233340.

El término “polinucleótido diana”, como se usa en el presente documento, se refiere a un polinucleótido de interés en estudio. En ciertas realizaciones, un polinucleótido diana contiene una o más secuencias que son de interés y en estudio.

5 El término “oligonucleótido”, como se usa en el presente documento, denota un multímero de cadena sencilla de nucleótidos de aproximadamente 2 a 200 nucleótidos, hasta 500 nucleótidos de longitud. Los oligonucleótidos pueden ser sintéticos o pueden prepararse enzimáticamente, y, en algunas realizaciones, tienen una longitud de 30 a 150 nucleótidos. Los oligonucleótidos pueden contener monómeros de ribonucleótidos (es decir, pueden ser oligorribonucleótidos) o monómeros de desoxirribonucleótidos. Un oligonucleótido puede ser de 10 a 20, de 21 a 30,
10 de 31 a 40, de 41 a 50, de 51 a 60, de 61 a 70, de 71 a 80, de 80 a 100, de 100 a 150 o de 150 a 200 nucleótidos de longitud, por ejemplo.

El término “cebador” como se usa en el presente documento se refiere a un oligonucleótido que es capaz de actuar como punto de iniciación de la síntesis cuando se coloca en condiciones en las que se induce síntesis de un producto de extensión de cebador, que es complementario a una cadena de ácido nucleico, es decir, en presencia de nucleótidos y un agente inductor tal como una ADN polimerasa y a una temperatura y pH adecuados. El cebador puede ser de cadena sencilla y debe ser suficientemente largo para cebar la síntesis del producto de extensión deseado en presencia del agente inductor. La longitud exacta del cebador dependerá de muchos factores, incluida la temperatura, la fuente del cebador y el uso del método. Por ejemplo, para aplicaciones de diagnóstico, dependiendo de la complejidad de la secuencia diana, el cebador oligonucleotídico típicamente contiene 15-25 o más nucleótidos, aunque puede contener menos nucleótidos. Los cebadores en este documento se seleccionan para que sean sustancialmente complementarios a diferentes cadenas de una secuencia de ADN diana particular. Esto significa que los cebadores deben ser suficientemente complementarios para hibridar con sus cadenas respectivas. Por lo tanto, la secuencia del cebador no necesita reflejar la secuencia exacta de la plantilla. Por ejemplo, un fragmento de nucleótido no complementario se puede unir al extremo 5' del cebador, siendo el resto de la secuencia del cebador el complementario de la cadena. Alternativamente, las bases no complementarias o las secuencias más largas pueden intercalarse en el cebador, con la condición de que la secuencia del cebador tenga una complementariedad suficiente con la secuencia de la cadena para hibridar con la misma y formar así la plantilla para la síntesis del producto de extensión.
15
20
25
30

El término “hibridación” o “hibridiza” se refiere a un proceso en el que una cadena de ácido nucleico se hibrida y forma un dúplex estable, un homodúplex o un heterodúplex, en condiciones de hibridación normales con una segunda cadena complementaria de ácido nucleico, y no forma un dúplex estable con moléculas de ácido nucleico no relacionadas en las mismas condiciones de hibridación normales. La formación de un dúplex se logra hibridando dos cadenas complementarias de ácido nucleico en una reacción de hibridación. La reacción de hibridación puede hacerse altamente específica mediante el ajuste de las condiciones de hibridación (a menudo denominadas rigurosidad de hibridación) en la que tiene lugar la reacción de hibridación, de modo que la hibridación entre dos cadenas de ácido nucleico no formará un dúplex estable, por ejemplo, dúplex que retiene una región de doble cadena en condiciones de rigurosidad normal, a menos que las dos cadenas de ácido nucleico contengan un cierto número de nucleótidos en secuencias específicas que son sustancial o completamente complementarias. “Hibridación normal o condiciones de rigurosidad normales” se determinan fácilmente para cualquier reacción de hibridación dada. Véase, por ejemplo, Ausubel et al., *Current Protocols in Molecular Biology*, John Wiley & Sons, Inc., Nueva York, o Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory Press. Como se usa en el presente documento, el término “hibridizar” o “hibridación” se refiere a cualquier proceso mediante el cual una cadena de ácido nucleico se une con una cadena complementaria a través del emparejamiento de la base.
35
40
45

Se considera que un ácido nucleico es “hibridable selectivamente” a una secuencia de ácido nucleico de referencia si las dos secuencias hibridan específicamente entre sí en condiciones de hibridación y lavado moderadas a alta rigurosidad. Se conocen condiciones de hibridación de rigurosidad moderada y alta (véase, por ejemplo, Ausubel, et al., *Short Protocols in Molecular Biology*, 3ª edición, Wiley & Sons 1995 y Sambrook et al., *Molecular Cloning: A Laboratory Manual*, Third Edition, 2001 Cold Spring Harbor, N.Y.). Un ejemplo de condiciones de alta rigurosidad incluye hibridación a aproximadamente 42°C en formamida al 50%, SSC 5X, solución de Denhardt 5X, SDS al 0.5% y ADN portador desnaturizado de 100 ug/ml seguido de lavado dos veces en 2xSSC y SDS al 0.5% a temperatura ambiente y dos veces adicionales en 0.1 × SSC y 0.5% SDS a 42°C.
50
55

El término “dúplex” o “duplexado”, como se usa en el presente documento, describe dos polinucleótidos complementarios que están emparejados por bases, es decir, hibridados juntos.

El término “amplificar” como se usa en la presente memoria se refiere al proceso de síntesis de moléculas de ácido nucleico que son complementarias a una o ambas cadenas de un ácido nucleico molde. La amplificación de una molécula de ácido nucleico incluye típicamente desnaturizar el ácido nucleico molde, hibridar cebadores al ácido nucleico molde a una temperatura que está por debajo de las temperaturas de fusión de los cebadores, y elongación enzimática de los cebadores para generar un producto de amplificación. Los pasos de desnaturización, recocido y alargamiento se pueden realizar una vez. Generalmente, sin embargo, las etapas de desnaturización, hibridación y alargamiento se realizan múltiples veces de manera que la cantidad de producto de amplificación aumenta, a menudo de manera exponencial, aunque la amplificación exponencial no es requerida por los presentes métodos. La
60
65

amplificación requiere típicamente la presencia de desoxirribonucleósido trifosfatos, una enzima ADN polimerasa y un tampón apropiado y/o cofactores para la actividad óptima de la enzima polimerasa. El término “producto de amplificación” se refiere a las secuencias de ácido nucleico, que se producen a partir del proceso de amplificación como se define en este documento. Una reacción de amplificación puede ser isotérmica (por ejemplo, en el caso de amplificación de círculo rodante) o puede requerir termociclado (en el caso de PCR).

Los términos “determinar”, “medir”, “evaluar”, “valorar”, “ensayar” y “analizar” se usan indistintamente en este documento para referirse a cualquier forma de medición, e incluyen determinar si un elemento está presente o no. Estos términos incluyen determinaciones cuantitativas y/o cualitativas. La evaluación puede ser relativa o absoluta. “Evaluar la presencia de” incluye determinar la cantidad de algo presente, así como determinar si está presente o ausente.

El término “particionar”, con respecto a un genoma, se refiere a la separación de una parte del genoma del resto del genoma para producir un producto que está aislado del resto del genoma. El término “particionar” abarca enriquecer.

El término “región genómica”, como se usa en este documento, se refiere a una región de un genoma, por ejemplo, un genoma de animal o planta tal como el genoma de un ser humano, mono, rata, pez o insecto o planta. En ciertos casos, un oligonucleótido utilizado en el método descrito aquí puede diseñarse utilizando una región genómica de referencia, es decir, una región genómica de secuencia de nucleótidos conocida, por ejemplo, una región cromosómica cuya secuencia se deposita en la base de datos Genbank del NCBI u otras bases de datos, por ejemplo. Tal oligonucleótido puede emplearse en un ensayo que usa una muestra que contiene un genoma de prueba, donde el genoma de prueba contiene un sitio de unión para el oligonucleótido.

El término “secuencia genómica”, como se usa en el presente documento, se refiere a una secuencia que se produce en un genoma.

El término “fragmento genómico”, como se usa en el presente documento, se refiere a una región de un genoma, por ejemplo, un genoma animal o vegetal tal como el genoma de un ser humano, mono, rata, pez o insecto o planta. Un fragmento genómico puede ser un cromosoma completo o un fragmento de un cromosoma.

El término “etiqueta de afinidad”, como se usa en el presente documento, se refiere a un resto que se puede usar para separar una molécula a la que está unida la etiqueta de afinidad de otras moléculas que no contienen la etiqueta de afinidad. Una “etiqueta de afinidad” es un miembro de un par de unión específica, es decir, dos moléculas donde una de las moléculas a través de medios químicos o físicos se une específicamente a la otra molécula. El miembro complementario del par de unión específico, al que se hace referencia aquí como “agente de captura” puede inmovilizarse (por ejemplo, a un soporte de cromatografía, un cordón o una superficie plana) para producir un soporte de cromatografía de afinidad que se une específicamente a la etiqueta de afinidad. En otras palabras, una “etiqueta de afinidad” puede unirse a un “agente de captura”, donde la etiqueta de afinidad se une específicamente al agente de captura, facilitando así la separación de la molécula a la que se une la etiqueta de afinidad de otras moléculas que no contienen la etiqueta de afinidad.

Como se usa en el presente documento, el término “resto de biotina” se refiere a un agente de afinidad que incluye biotina o un análogo de biotina tal como destiobiotina, oxibiotina, 2'-iminobiotina, diaminobiotina, sulfóxido de biotina, biocitina, etc. Los restos de biotina se unen a la estreptavidina con una afinidad de al menos 10^{-8} M. Un agente de afinidad de biotina también puede incluir un ligador, por ejemplo, -LC-biotina, -LC-LC-Biotina, -SLC-Biotina o -PEG_n-Biotina donde n es 3-12.

El término “nucleótido terminal”, como se usa en el presente documento, se refiere al nucleótido en el extremo 5' o el extremo 3' de una molécula de ácido nucleico. La molécula de ácido nucleico puede estar en forma de cadena doble (es decir, duplexada) o en forma de cadena sencilla.

El término “ligar”, como se usa en el presente documento, se refiere a la unión enzimáticamente catalizada del nucleótido terminal en el extremo 5' de una primera molécula de ADN al nucleótido terminal en el extremo 3' de una segunda molécula de ADN.

Una “pluralidad” contiene al menos 2 miembros. En ciertos casos, una pluralidad puede tener al menos 10, al menos 100, al menos 100,000, al menos 100,000, al menos 10^6 , al menos 10^7 , al menos 10^8 o a al menos 10^9 o más miembros.

Si dos ácidos nucleicos son “complementarios”, se hibridan entre sí en condiciones de alta rigurosidad. El término “perfectamente complementario” se usa para describir un dúplex en el que cada base de uno de los ácidos nucleicos se empareja con un nucleótido complementario en el otro ácido nucleico. En muchos casos, dos secuencias que son complementarias tienen al menos 10, por ejemplo, al menos 12 o 15 nucleótidos de complementariedad y en ciertos casos pueden tener una, dos o tres bases no complementarias.

- 5 El término “digerir” pretende indicar un proceso mediante el cual un ácido nucleico es escindido por una enzima de restricción. Para digerir un ácido nucleico, se ponen en contacto una enzima de restricción y un ácido nucleico que contiene un sitio de reconocimiento para la enzima de restricción en condiciones adecuadas para que funcione la enzima de restricción. Se conocen las condiciones adecuadas para la actividad de las enzimas de restricción comercialmente disponibles, y se suministran con esas enzimas después de la compra.
- 10 Un “sitio de unión de oligonucleótidos” se refiere a un sitio al que se hibrida un oligonucleótido en un polinucleótido diana. Si un oligonucleótido “proporciona” un sitio de unión para un cebador, entonces el cebador puede hibridarse con ese oligonucleótido o su complemento.
- 15 El término “separar”, como se usa en el presente documento, se refiere a la separación física de dos elementos (por ejemplo, por tamaño o afinidad, etc.) así como a la degradación de un elemento, dejando intacto al otro.
- 20 El término “región cromosómica de referencia”, como se usa en el presente documento, se refiere a una región cromosómica de secuencia de nucleótidos conocida, por ejemplo, una región cromosómica cuya secuencia se deposita en la base de datos Genbank del NCBI u otras bases de datos, por ejemplo.
- 25 En una célula, el ADN generalmente existe en una forma de cadena doble, y como tal, tiene dos cadenas complementarias de ácido nucleico a las que se hace referencia en este documento como las cadenas “superior” e “inferior”. En ciertos casos, las cadenas complementarias de una región cromosómica se pueden denominar cadenas “más” y “menos”, las cadenas “primera” y “segunda”, las cadenas “codificante” y “no codificante”, y las cadenas “Watson” y “Crick” o las cadenas “sentido” y “antisentido”. La asignación de una cadena por ser una cadena superior o inferior es arbitraria y no implica ninguna orientación, función o estructura particular. Las secuencias de nucleótidos de la primera cadena de varias regiones cromosómicas de mamífero de ejemplo (por ejemplo, BAC, conjuntos, cromosomas, etc.) son conocidas, y pueden encontrarse en la base de datos de Genbank de NCBI, por ejemplo.
- 30 El término “cadena superior”, como se usa en este documento, se refiere a cualquiera de las cadenas de un ácido nucleico, pero no a ambas cadenas de un ácido nucleico. Cuando un oligonucleótido o un cebador se une o hibrida “solo a una cadena superior”, se une a una sola cadena, pero no a la otra. El término “cadena inferior”, como se usa en el presente documento, se refiere a la cadena que es complementaria a la “cadena superior”. Cuando un oligonucleótido se une o hibrida “solo a una cadena”, se une a una sola cadena, por ejemplo, la primera o segunda
- 35 cadena, pero no la otra cadena.
- 40 El término “enlace covalente” se refiere a la producción de un enlace covalente entre dos moléculas separadas, por ejemplo, las cadenas superior e inferior de un ácido nucleico de cadena doble. Ligar es un tipo de enlace covalente.
- 45 El término “desnaturalización”, como se usa en este documento, se refiere a la separación de al menos una porción de los pares de bases de un dúplex de ácido nucleico colocando el dúplex en condiciones de desnaturalización adecuadas. Las condiciones de desnaturalización son bien conocidas en la técnica. En una realización, con el fin de desnaturalizar un dúplex de ácido nucleico, el dúplex puede exponerse a una temperatura que está por encima de la temperatura de fusión del dúplex, liberando de ese modo una cadena del dúplex de la otra. En ciertas realizaciones, un ácido nucleico puede desnaturalizarse exponiéndolo a una temperatura de al menos 90°C durante un período de tiempo adecuado (por ejemplo, al menos 30 segundos, hasta 30 minutos). Los ácidos nucleicos también pueden desnaturalizarse químicamente (por ejemplo, usando urea o NaOH).
- 50 Como se usa en el presente documento, el término “marcador” se refiere a cualquier átomo o molécula que se puede usar para proporcionar un efecto detectable (preferiblemente cuantificable), y que se puede unir a un ácido nucleico o proteína. Los marcadores incluyen, pero no se limitan a tintes y radiomarcadores como ³²P; restos de unión tales como biotina; haptenos tales como digoxigenina; restos luminogénicos, fosforescentes o fluorogénicos; y colorantes fluorescentes solos o en combinación con restos que pueden suprimir o cambiar espectros de emisión por transferencia de energía de resonancia de fluorescencia (FRET). Los marcadores pueden proporcionar señales detectables por fluorescencia, radioactividad, colorimetría, gravimetría, difracción o absorción de rayos X, magnetismo, actividad enzimática y similares. Un marcador puede ser un resto cargado (carga positiva o negativa) o, alternativamente, puede ser neutral a la carga. Los marcadores pueden incluir o consistir en un ácido nucleico o una secuencia de proteína, siempre que la secuencia que comprende el marcador sea detectable.
- 55 El término “oligonucleótido marcado”, como se usa en el presente documento, se refiere a un oligonucleótido que tiene una etiqueta de afinidad (por ejemplo, un resto de biotina), un oligonucleótido modificado con átomos o grupos que permite la separación o detección (por ejemplo, bromo-desoxiuridina o partículas de oro coloidal que le confieren diferente densidad), y un oligonucleótido modificado con un marcador ópticamente detectable (por ejemplo, una fluorescencia u otro tipo de marcador emisor de luz). Los oligonucleótidos que contienen solo nucleótidos naturales
- 60 no son oligonucleótidos marcados.
- 65

El término “secuenciación”, como se usa en este documento, se refiere a un método por el cual la identidad de al menos 10 nucleótidos consecutivos (por ejemplo, la identidad de al menos 20, al menos 50, al menos 100 o al menos 200 o más nucleótidos consecutivos) de un polinucleótido se obtiene.

5 El término “secuenciación de próxima generación” se refiere a las llamadas plataformas paralelas secuenciadas por síntesis o secuenciación por ligamiento actualmente empleadas por Illumina, Life Technologies y Roche, etc. Los métodos de secuenciación de próxima generación también pueden incluir métodos de secuenciación de nanoporos o métodos basados en detección electrónica como la tecnología Ion Torrent comercializada por Life Technologies.

10 El término “extender”, como se usa en el presente documento, se refiere a la extensión de un cebador mediante la adición de nucleótidos usando una polimerasa. Si se extiende un cebador que se hibrida con un ácido nucleico, el ácido nucleico actúa como molde para una reacción de extensión.

15 El término “PCR clonal” es una técnica de PCR en la que cada reacción se realiza en una única molécula patrón, y las reacciones de PCR se mantienen espacialmente separadas entre sí. La PCR puente y la PCR en emulsión, comúnmente utilizadas en aplicaciones de secuenciación de próxima generación, son ejemplos de PCR clonal.

20 El término “PCR puente” se refiere a una reacción en cadena de la polimerasa en fase sólida en la que los cebadores que se extienden en la reacción están unidos a un sustrato por sus extremos 5'. Durante la amplificación, los amplicones forman un puente entre los cebadores unidos. PCR puente (que también se conoce como “grupo PCR”) se utiliza en la plataforma Solexa de Illumina. PCR puente y la plataforma Solexa de Illumina se describen generalmente en una variedad de publicaciones, por ejemplo, Gudmundsson et al (Nat. Genet. 2009 41:1122-6), Out et al (Hum. Mutat. 2009 30: 1703-12) y Turner (Nat. Methods 2009 6: 315-6), la patente de EE.UU. N° 7,115,400, y la publicación de solicitud de EE.UU. Nos. US20080160580 y US20080286795. El PCR puente es un tipo de “PCR clonal”, es decir, es una técnica de PCR en la que cada reacción se inicia en una única molécula patrón, y las reacciones de PCR se mantienen espacialmente separadas entre sí.

30 El término “secuencia de código de barras” o “código de barras molecular”, como se usa en el presente documento, se refiere a una secuencia única de nucleótidos utilizada para a) identificar y/o rastrear la fuente de un polinucleótido en una reacción y/o b) contar cuántas veces la molécula inicial se secuencia (por ejemplo, en los casos en que sustancialmente cada molécula en una muestra se etiqueta con una secuencia diferente, y luego la muestra se amplifica). Una secuencia de código de barras puede estar en el extremo 5', el extremo 3' o en el medio de un oligonucleótido. Las secuencias de códigos de barras pueden variar ampliamente en tamaño y composición; las siguientes referencias proporcionan una guía para seleccionar conjuntos de secuencias de códigos de barras apropiadas para realizaciones particulares: Brenner, patente de los EE.UU. No. 5,635,400; Brenner et al, Proc. Natl. Acad. Sci., 97: 1665-1670 (2000); Shoemaker et al., Nature Genetics, 14: 450-456 (1996); Morris et al, publicación de patente europea 0799897A1; Wallace, Pat. U.S. No. 5,981,179; y similares. En realizaciones particulares, una secuencia de código de barras puede tener una longitud en el intervalo de 4 a 36 nucleótidos, o de 6 a 30 nucleótidos, o de 8 a 20 nucleótidos.

40 Como se usa en el presente documento, el término “reactivos de PCR” se refiere a todos los reactivos que se requieren para realizar una reacción en cadena de la polimerasa (PCR) en un molde. Como es sabido en la técnica, los reactivos de PCR incluyen esencialmente un primer cebador, un segundo cebador, una polimerasa termoestable y nucleótidos. Dependiendo de la polimerasa utilizada, los iones (por ejemplo, Mg²⁺) también pueden estar presentes. Los reactivos de PCR pueden contener opcionalmente una plantilla a partir de la cual se puede amplificar una secuencia diana.

El término “ligar intramolecularmente” se refiere a una ligadura en la que el extremo 5' y el extremo 3' de una cadena de ácido nucleico se ligan entre sí para producir una molécula circular de ADN.

50 Como se usa en el presente documento, el término “endonucleasa de restricción de la familia MspJI” se refiere a una familia de endonucleasas de restricción que reconocen citosina metilada y generan una ruptura bicatenaria en la molécula de ADN en un sitio que está cadena arriba o cadena abajo de la citosina metilada. Estas enzimas no escinden el ADN no metilado. Si ambas cadenas del ADN están metiladas (lo que es comúnmente el caso en la metilación de CpG), la enzima cortará el ADN para producir un fragmento de 20-40 pares de bases de longitud, dependiendo de la enzima utilizada. MspJI, por ejemplo, reconoce cada sitio hemi-metilado individualmente y se escinde bidireccionalmente para generar fragmentos de 32 bases o 31 bases, respectivamente. Estos fragmentos contienen el sitio metilado central y tienen proyecciones de 5' de 4 bases en cada extremo. La familia de endonucleasas MspJI se describe en, por ejemplo, Zheng et al. (Nucleic Acids Res. 2010. Una familia única de endonucleasas de restricción dependientes de la modificación de tipo Mrr. 38: 5527-34) y Cohen-Karni (Proc. Natl. Acad. Sci. 2011. La familia MspJI de endonucleasas de restricción dependientes de la modificación para estudios epigenéticos 108: 11040-5) y US20100167942.

65 Los ejemplos de endonucleasas de restricción de la familia MspJI incluyen FspEI, LpnPI, AspBHI, RlaI y SgrTI. La referencia a un miembro de la familia MspJI, genéricamente o por nombre, pretende referirse a una endonucleasa de restricción de tipo silvestre, así como a variantes que tienen una secuencia de aminoácidos que es al menos 90% (por ejemplo, al menos 95%) idéntica a la endonucleasa de restricción de tipo silvestre.

Como se usa en el presente documento, el término "población de fragmentos que están en el intervalo de 20-40 pares de bases de longitud" se refiere a una mezcla de fragmentos de digestión. Los fragmentos tienen un residuo de citosina central (que es parte del sitio de reconocimiento de la endonucleasa de la familia MspJI) y tienen proyecciones de 3' o 5', dependiendo de qué enzima se use.

Como se usa en la presente memoria, el término "extremos respectivos", en la frase "secuencia adaptadora de ligamiento A y secuencia adaptadora B a los extremos respectivos de un fragmento diana" pretende significar que la secuencia A se agrega a un extremo del fragmento diana y la secuencia B se agrega al otro extremo del fragmento objetivo.

Ciertos polinucleótidos descritos en la presente memoria pueden referirse mediante una fórmula (por ejemplo, "B'-X'-A" y "A-X-B"). Tales fórmulas siguen la convención establecida en que describen un polinucleótido que está orientado en la dirección 5' a 3'. Los componentes de la fórmula, por ejemplo, "A", "X" y "B" se refieren a secuencias de nucleótidos definibles por separado dentro de un polinucleótido, donde las secuencias se unen entre sí covalentemente de manera que un polinucleótido descrito por una fórmula es una molécula individual. En muchos casos, los componentes de la fórmula son inmediatamente adyacentes entre sí en la molécula individual. Siguiendo la convención, el complemento de una secuencia que se muestra en una fórmula se indicará con una prima (') tal que el complemento de la secuencia "A" será "A'". Además, a menos que se indique lo contrario o implícito a partir del contexto, un polinucleótido definido por una fórmula puede tener una secuencia adicional, un sitio de unión de cebador, un código de barras molecular, un promotor o un espaciador, etc., en su extremo 3', su extremo 5' o ambos extremos 3' y 5'.

Como se usa en el presente documento, el término "secuencias adaptadoras A y B" se refiere a diferentes secuencias.

Como se usa en el presente documento, el término "ligablemente adyacente" en el contexto de dos secuencias de oligonucleótidos que están ligablemente adyacentes entre sí, significa que no hay nucleótidos intermedios entre dos oligonucleótidos y que se pueden ligar entre sí.

Como se usa en el presente documento, el término "oligonucleótido puente", como se usa en la presente memoria, se refiere a un oligonucleótido que, cuando se hibrida con dos o más polinucleótidos diferentes, actúa como una "puente" para colocar los polinucleótidos uno al lado del otro para que puedan ligarse juntos, como se ilustra en la FIG. 1.

Como se usa en el presente documento, el término "molécula circular de ácido nucleico" se refiere a una cadena que está en forma de un círculo cerrado que no tiene extremos 3' o 5' libres.

Como se usa en este documento, el término "fracción fetal" se refiere al porcentaje de ADN libre de células de un feto en desarrollo en el flujo sanguíneo materno de una hembra embarazada.

Otras definiciones de términos pueden aparecer a lo largo de la especificación.

Descripción de modalidades de ejemplo

Se proporciona un método para estimar la cantidad de un locus metilado. En ciertas realizaciones, el método comprende: digerir una muestra de ácido nucleico que contiene copias metiladas y no metiladas de un locus genómico con un miembro de familia MspJI para producir una población de fragmentos que están en el intervalo de 20-40 nucleótidos de longitud, ligar la secuencia adaptadora A y la secuencia adaptadora B a los extremos respectivos de un fragmento diana de la secuencia X, y cuantificar la cantidad de productos de ligación de fórmula A-X-B. También se proporciona un kit para realizar el método.

Antes de que se describan las diversas realizaciones, debe entenderse que las enseñanzas de esta divulgación no están limitadas a las realizaciones particulares descritas, y como tales pueden, por supuesto, variar. También debe entenderse que la terminología utilizada en este documento tiene el propósito de describir realizaciones particulares solamente, y no pretende ser limitativa, ya que el alcance de las presentes enseñanzas estará limitado solo por las reivindicaciones adjuntas.

Los títulos de las secciones utilizados en este documento son solo para fines de organización y no deben interpretarse como una limitación del tema descrito de ninguna manera. Si bien las presentes enseñanzas se describen junto con diversas realizaciones, no se pretende que las presentes enseñanzas se limiten a dichas realizaciones. Por el contrario, las presentes enseñanzas abarcan diversas alternativas, modificaciones y equivalentes, como apreciarán los expertos en la materia.

A menos que se defina lo contrario, todos los términos técnicos y científicos usados en la presente memoria tienen el mismo significado que entiende comúnmente un experto habitual en la técnica a la que pertenece esta divulgación. Aunque cualquier método y material similar o equivalente a los descritos en este documento también se puede usar en la práctica o prueba de las presentes enseñanzas, se describirán ahora algunos métodos y materiales de ejemplo.

La cita de cualquier publicación es para su divulgación antes de la fecha de presentación y no debe interpretarse como una admisión de que las presentes reivindicaciones no tienen derecho a anteceder a dicha publicación en virtud de la invención anterior. Además, las fechas de publicación proporcionadas pueden ser diferentes de las fechas de publicación reales que pueden necesitar confirmación independiente.

Como será evidente para los expertos en la materia tras la lectura de esta descripción, cada una de las realizaciones individuales descritas e ilustradas en este documento tiene componentes discretos y características que pueden separarse fácilmente o combinarse con las características de cualquiera de las otras varias realizaciones sin salir del alcance o el espíritu de las presentes enseñanzas. Cualquier método citado puede llevarse a cabo en el orden de los eventos citados o en cualquier otro orden que sea lógicamente posible.

Con referencia a la FIG. 1, algunas realizaciones del método comprenden digerir una muestra 2 de ácido nucleico que contiene tanto copias no metiladas (4 y 6) como copias metiladas (8) de un locus genómico usando una endonucleasa de restricción de la familia MspJI. Como se ilustra, si la secuencia de reconocimiento está metilada en ambas cadenas (indicada por el círculo 10 lleno), entonces la enzima se escindirán en ambos lados de la secuencia de reconocimiento, como se indica por las x en el fragmento 8. Esta digestión produce la muestra 12 digerida que comprende una población de fragmentos que están en el intervalo de 20-40 pares de bases de longitud y tienen una citosina metilada central. Como sería evidente, la muestra inicial que se digiere no es una muestra amplificada. Si se digiere ADN genómico, la muestra digerida debe contener al menos varios miles, si no al menos 10,000 o 100,000 o más, fragmentos que tienen una longitud de 20-40 pares de bases y corresponden a diferentes loci metilados, así como fragmentos más largos de los mismos loci que no han sido digeridos en un pequeño fragmento porque no están metilados. Uno de estos fragmentos que está en el intervalo de 20-40 pares de bases de longitud corresponde al fragmento 14 objetivo de la secuencia X. En este método, el fragmento 14 objetivo se selecciona ligando la secuencia adaptadora A 16 y la secuencia adaptadora B 18 a sus extremos. Esto se hace al (i) hibridar un oligonucleótido 20 de puente de fórmula B'-X'-A' a los fragmentos de 12 en presencia de las secuencias adaptadoras A 16 y B 18, en donde X', A' y B' son complementario a X, A y B, respectivamente, para hacer que el dúplex 21 y (ii) la secuencia adaptadora de ligadura A 16, el fragmento de secuencia X 14 y la secuencia adaptadora B 18 el uno al otro para producir un producto 22 de fórmula A-X-B. Como sería evidente, el oligonucleótido 20 puente y las secuencias adaptadoras A 16 y B 18 están diseñados de manera que los extremos de las secuencias A 16 y B 18 están ligablemente adyacentes a los extremos del fragmento diana de la secuencia X cuando se hibridan con el oligonucleótido puente 20 en dúplex 21. El oligonucleótido 20 puente y las secuencias adaptadoras A 16 y B 18 se muestran como dúplex 19 en la FIG. 1 únicamente para mostrar qué secuencias son complementarias entre sí. En la práctica, el oligonucleótido 20 puente y las secuencias adaptadoras A 16 y B 18 generalmente no se hibridan entre sí antes de la hibridación con la muestra. En la práctica, los diversos oligonucleótidos pueden combinarse con la muestra digerida en un único recipiente, y la mezcla puede calentarse y luego enfriarse y desnaturalizando e hibridando de este modo las diversas secuencias entre sí antes de la ligación. Después de ligar la secuencia adaptadora A 16, el fragmento de secuencia X 14 y la secuencia adaptadora B 12 entre sí para producir un producto 22 de fórmula A-X-B, la cantidad de producto 22 puede cuantificarse, proporcionando así una estimación de la cantidad del locus metilado en la muestra 2 de ácido nucleico. En realizaciones particulares, la ligasa usada puede ser una ligasa termoestable, y la etapa de ligación puede realizarse ciclando la reacción a través de múltiples rondas de desnaturalización y renaturalización, conduciendo así la reacción a completarse. La cuantificación del producto 22 puede realizarse de diversas maneras, por ejemplo, mediante PCR cuantitativa, mediante secuenciación y mediante conteo digital, ejemplos de los cuales se describen con mayor detalle a continuación. En ciertas realizaciones, las sondas de puente y la secuencia adaptadora se pueden diseñar y/o ligar usando un método descrito en la solicitud de patente del Reino Unido número de serie. 1321191.7, presentada el 2 de diciembre de 2013.

En la realización ilustrada en la FIG. 2, las secuencias adaptadoras A 16 y B 18 pueden estar en moléculas de oligonucleótidos separadas que comprenden secuencias 40 y 42 no complementarias, que proporcionan sitios de unión de cebador en el producto 22. En esta realización, el producto 22 puede amplificarse usando los cebadores 44 y 46. En esta realización, la cantidad de producto 22 puede cuantificarse mediante cualquier ensayo de qPCR adecuado, por ejemplo, un ensayo TaqMan o similar. En otra realización, el producto 22 puede secuenciarse (con o sin amplificación) usando los cebadores 44 y 46, o cebadores que se hibridan a colas que se han añadido a los cebadores 44 y 46, por ejemplo. En estas realizaciones, la cantidad de producto 22 puede estimarse contando el número de lecturas de secuencia correspondientes a la secuencia X. En realizaciones alternativas, las secuencias 40 y/o 42 no complementarias pueden contener un código de barras molecular y cada molécula de producto 22 contiene una secuencia de código de barras diferente. En estas realizaciones, la cantidad de producto 22 puede estimarse contando el número de secuencias de códigos de barras únicas que están asociadas con las lecturas de secuencia correspondientes a la secuencia X.

En otra realización ilustrada en la FIG. 3, las secuencias adaptadoras A y B están presentes en una sola molécula 50 de oligonucleótido y el producto 22 es una molécula circular de ácido nucleico. En estas realizaciones, el producto 22 puede cuantificarse amplificando el producto mediante amplificación por círculo rodante (RCA) (por ejemplo, usando el cebador 52, que puede ser complementario a una secuencia en cualquier parte del producto) y luego estimando el número de productos RCA producidos. Los productos de amplificación de círculo rodante contendrán múltiples copias de cada una de las secuencias en el producto 22. En estas realizaciones, la cuantificación puede realizarse hibridando

el producto RCA con una población de oligonucleótidos marcados que se hibridan con múltiples posiciones en el producto RCA; e individualmente contando el número de complejos marcados RCA.

La fig. 4 ilustra una forma en la que puede implementarse esta realización. En esta implementación, el producto 22 está compuesto por cuatro moléculas 22a, 22b, 22c y 22d de producto. En la práctica, el producto 22 puede estar compuesto de al menos 100, al menos 1,000 o al menos 10,000 o más moléculas de producto de fórmula A-X-B. En esta realización, los productos se amplifican mediante amplificación de círculo rodante usando el cebador 52 (que puede ser complementario a una secuencia en cualquier parte del producto) para producir una pluralidad de productos RCA. El número de productos de amplificación de círculo rodante se puede estimar distribuyendo los productos RCA en la superficie de un soporte (un portaobjetos), hibridando los productos RCA usando oligonucleótidos marcados (por ejemplo, oligonucleótidos marcados fluorescentemente) y contando el número de señales discretas en un área del soporte por microscopía, por ejemplo, microscopía de fluorescencia. El etiquetado puede realizarse antes o después de que los productos se hayan distribuido en el soporte y, dado que cada producto RCA contiene miles de copias de las mismas secuencias, debe haber miles de sitios de unión para los oligonucleótidos marcados, aumentando así la señal. En ciertas realizaciones, los métodos de amplificación y/ o detección pueden implementarse usando un método descrito en la solicitud de patente del Reino Unido número de serie 1321196.6, presentada el 2 de diciembre de 2013.

La cantidad del locus metilado en la muestra se puede expresar de diferentes maneras. Por ejemplo, en algunas realizaciones, los resultados obtenidos en la etapa de cuantificación pueden normalizarse a la cantidad de ADN de entrada y expresarse como una medida por cantidad de ADN de entrada (por ejemplo, x número de moléculas en y cantidad de ADN de entrada). Este número puede proporcionar una métrica útil cuando se compara con la cantidad de otras secuencias en la misma muestra. Por ejemplo, la cantidad relativa (por ejemplo, el porcentaje) del locus metilado en la muestra puede calcularse comparando los resultados obtenidos mediante el presente método con los resultados obtenidos para un locus de referencia que se sabe que siempre está metilado en la muestra. En este ejemplo, la cantidad del locus de referencia en la muestra se puede cuantificar usando un método similar (donde se puede usar la misma enzima u otra enzima de la familia MspJI y las secuencias específicas diana en el oligonucleótido puente se hibridan con un fragmento correspondiente a locus de referencia, no al locus X). En otro ejemplo, la cantidad relativa (por ejemplo, el porcentaje) del locus metilado en la muestra puede calcularse comparando los resultados obtenidos mediante el presente método con los resultados obtenidos para un locus de referencia que se sabe que está presente en la muestra. En este ejemplo, la cantidad del locus de referencia en la muestra puede cuantificarse usando un método similar (donde se puede usar una enzima distinta de una enzima de la familia MspJI, por ejemplo, una enzima insensible a la metilación y las secuencias específicas diana en el oligonucleótido puente hibridan a un fragmento correspondiente al locus de referencia, no a X). En algunas realizaciones, la cantidad absoluta de locus metilado en la muestra se puede comparar con una curva estándar (por ejemplo, una curva estándar generada usando muestras de control que contienen cantidades conocidas del locus metilado) para proporcionar una estimación del número absoluto de moléculas del locus metilado en la muestra.

Las longitudes de las diversas regiones de los adaptadores y los oligonucleótidos puente pueden variar mucho dependiendo de la aplicación deseada y de la cantidad de carga (es decir, cuántos sitios de unión del cebador, códigos de barras, etc.) llevan los adaptadores. En la práctica, la secuencia X' del oligonucleótido puente se diseñará para que coincida con la secuencia de un fragmento que se espera que se genere digiriendo una muestra genómica con la endonucleasa de restricción de la familia MspJI que se usa en el método. Por ejemplo, si se usa MspJI, entonces la secuencia X' tendrá aproximadamente 32 nucleótidos de longitud y los extremos de X corresponderán a los sitios de escisión para MspJI en el ADN genómico. Otros miembros de la familia MspJI crean fragmentos de diferentes longitudes y, como tal, la longitud de X' no necesita ser de aproximadamente 32 bases si se usa otra enzima, que incluye FspEI, LpnPI, AspBHI, RlaI y SgrTI. En algunos casos, X' puede estar en el rango de 25-35 nucleótidos de longitud. Debido a que los sitios de reconocimiento para estas enzimas son conocidos y se conoce la secuencia de varios genomas, incluido el genoma humano, el diseño de los oligonucleótidos puente puede realizarse a mano o computacionalmente. Dependiendo de la aplicación deseada, las secuencias adaptadoras A y B pueden ser de 15 a 100 bases (por ejemplo, de 18 a 30 bases) de longitud. Como debería ser fácilmente evidente, la secuencia de nucleótidos de cualquier secuencia adicional que esté unida a las secuencias adaptadoras, por ejemplo, sitios de unión de cebadores/códigos de barras, etc., debería diseñarse de modo que no se hibriden con el genoma en estudio.

En realizaciones particulares, una de las secuencias adaptadoras puede estar unida a una etiqueta de afinidad, por ejemplo, una fracción de biotina, de modo que, después de la ligación, los productos de ligación pueden separarse de las moléculas no ligadas usando un soporte sólido que comprende un agente de captura unido a la superficie para la etiqueta de afinidad, uniéndose así los productos de ligación al soporte sólido y aislando esas moléculas de otros ácidos nucleicos en la muestra.

Como sería evidente, los cebadores usados en algunas realizaciones pueden contener secuencias que son compatibles con el uso en, por ejemplo, el método de terminación reversible de Illumina, el método de pirosecuenciación de Roche (454), la secuenciación de Life Technologies por ligadura (plataforma SOLiD) o Life Technologies Plataforma Ion Torrent. Los ejemplos de tales métodos se describen en las siguientes referencias: Margulies et al (Nature 2005 437: 376-80); Ronaghi et al (Analytical Biochemistry 1996 242: 84-9); Shendure (Science 2005 309: 1728); Imelfort et al (Breve Bioinform. 2009 10:609-18); Fox et al. (Methods Mol Biol. 2009; 553:79-108); Appleby et al (Methods Mol Biol. 2009; 513: 19-39) y Morozova (Genomics, 2008 92: 255-64).

En algunas realizaciones, el método puede multiplexarse para cuantificar varios loci metilados en la muestra. Esto se puede hacer usando varios oligonucleótidos de puente, donde la secuencia correspondiente a X es complementaria a los loci investigados. El método puede usarse para analizar al menos 2, al menos 5, al menos 10, al menos 50 o al menos 100 loci diferentes en la misma reacción.

El método descrito anteriormente puede emplearse para analizar ADN genómico de prácticamente cualquier organismo, que incluye, entre otros, plantas, animales (por ejemplo, reptiles, mamíferos, insectos, gusanos, peces, etc.), muestras de tejidos, bacterias, hongos (ejemplo, levadura), fagos, virus, tejido cadavérico, muestras arqueológicas/antiguas, etc. En ciertas realizaciones, el ADN genómico usado en el método puede derivarse de un mamífero, donde en ciertas realizaciones el mamífero es un humano. En realizaciones de ejemplo, la muestra genómica puede contener ADN genómico de una célula de mamífero, tal como una célula humana, de ratón, de rata o de mono. La muestra puede prepararse a partir de células cultivadas o células de una muestra clínica, por ejemplo, una biopsia de tejido, raspado o lavado o células de una muestra forense (es decir, células de una muestra recogida en una escena del crimen). En realizaciones particulares, la muestra de ácido nucleico se puede obtener a partir de una muestra biológica tal como células, tejidos, fluidos corporales y heces. Los fluidos corporales de interés incluyen, pero no están limitados a, sangre, suero, plasma, saliva, mucosa, flema, líquido cefalorraquídeo, líquido pleural, lágrimas, líquido del conducto lactial, linfa, esputo, líquido cefalorraquídeo, líquido sinovial, orina, líquido amniótico y semen. En realizaciones particulares, se puede obtener una muestra de un sujeto, por ejemplo, un humano. En algunas realizaciones, la muestra analizada puede ser una muestra de ADN libre de células obtenido de sangre, por ejemplo, de la sangre de una hembra embarazada.

En ciertas realizaciones, el presente método puede usarse para determinar el estado de metilación de un locus diana que está asociado con una enfermedad, por ejemplo, cáncer o una enfermedad hepática tal como hepatitis crónica o cirrosis, donde el estado de metilación del locus puede usarse para diagnosticar la enfermedad. En estas realizaciones, la fuente del ADN genómico puede ser una biopsia de tejido o un fluido corporal que contiene ADN derivado de tejido enfermo.

En ciertas realizaciones, el presente método puede usarse en pruebas prenatales no invasivas (NIPT) para el diagnóstico de anomalías genéticas y epigenéticas en el feto, que incluyen aneuploidías fetales tales como trisomía 21, trisomía 18 y trisomía 13; trastornos de la impresión como el síndrome de Beckwith-Wiedemann, los síndromes de Prader-Willi y Angelman y la osteodistrofia hereditaria de Albright; y otros defectos genéticos, como el síndrome X frágil. Por ejemplo, la hipermetilación de los siguientes loci puede ser diagnóstica: islas CpG alrededor de CRYAA, HLCS, C21orf63, OLIG2, CBR1, SIM2, DSCAM, TRPM2, C21orf29, MIR155HG, ICOS tipo ligando, SIM2, DSCR6, chr21 grup00165, PRMT2 y COL18A1 para la trisomía 21 (Lim et al., BMC Med Genomics, 2014 7:1; Tong et al., Clin Chem, 2009 56:90; Tong et al., PLoS One, 2010 5:e15244; Chim et al., Clin. Chem, 2008 54:500; Patente de Estados Unidos N° 8,476,013); sitios CpG hipermetilados en la secuencia promotora de SERPINB5 (maspin), entre los genes VAPA y APCDD1, y sitios CpG en CIDEA, chr18 group00091 chr18, group00094, KLHL14, ST8SIA3, ONECUT2, RAX, chr18 group00277, NETO1, MBP y NFATC1 para la trisomía 18 (Tong et al., Clin Chem, 2006 52: 2194; Tsui et al., PLoS One, 2010 5:e15069; Patente de Estados Unidos N° 8,476,013); sitios de hipermetilación en chr13 group00016, ATP8A2, GSI1, PDX1, MAB21L1, RB1, PCDH17, KLHL1, POU4F1, GPC6, SOX21, ZIC2, chr13 grup00385, chr13 grup00390, chr13 grup00391, chr13 grup00395, chr13 grup00399 y PROZ para trisomía 13 (Patente de EE. UU. No. 8,476,013); sitios de hipermetilación de CpG a 11p15.5 para el síndrome de Beckwith-Wiedemann (Coffee et al., Genet Med, 2006 8: 628); 15q11.2-q13 para los síndromes de Prader-Willi y Angelman (Procter et al., Clin Chem, 2006 52:1276); el locus XL asociado con la Osteodistrofia Hereditaria de Albright (Izzi et al., PLoS ONE, 2012 7:e38579); y los sitios CpG en el gen FMR1 para el síndrome X frágil (Hansen et al., Hum Mol Genet, 1992 1: 571; Alisch et al., BMC Med Genet, 2013 14:18). Por lo tanto, un oligonucleótido puente del presente método se puede diseñar para que sea complementario a los sitios hipermetilados de estos loci para diagnosticar anomalías genéticas y epigenéticas prenatales.

En ciertas realizaciones, el presente método se puede usar para NIPT. Por ejemplo, las secuencias hipermetiladas en la región promotora del gen RASSF1A se pueden usar para mejorar el genotipado de RhD prenatal y la detección de susceptibilidad a preeclampsia (Chan et al., Clin Chem, 2006 52: 2211; Tsui et al., Prenat Diagn, 2007 27:1212). Como tal, un oligonucleótido puente del presente método se puede diseñar para que sea complementario a las secuencias metiladas en la región promotora del gen RASSF1A. Se encontró que la familia de dominios RASSF1 (asociación Ras (RalGDS/AF-6) 1) estaba hipermetilada en la placenta pero completamente no metilada en las células sanguíneas de la madre.

En ciertas realizaciones, el presente método puede usarse para detectar la hipermetilación en ADN fetal mediante el diseño de un oligonucleótido puente complementario de secuencias metiladas en o asociado con los siguientes genes: SOX14, TBX3, SIX2, TLX3, FOXP4, NPY, SHH, OSR2, GLIS3, PRMT8, PAX9, SIX1, ISL2, DLX4, CBX4 y EDG6 (Nygren et al., Clin Chme, 2010, 56: 1627; Patente de los Estados Unidos No. 8,476,013).

En otras formas de realización más, el presente método puede usarse para diagnosticar la susceptibilidad a embarazos patológicos, tales como preeclampsia. Por ejemplo, la hipermetilación en sitios CpG en la región promotora de c-myc y el exón 1 de H19 de ADN fetal están asociados con preeclampsia (Rahat et al., Mol Hum Reprod, 2014; Lu et al., Int

J Mol Med, 2014). De este modo, en ciertas realizaciones, se puede usar un oligonucleótido puente del presente método diseñado para ser complementario a las secuencias metiladas en estas regiones para determinar la predisposición a la preeclampsia.

5 En ciertas realizaciones, el presente método se puede emplear como parte de un diagnóstico de cáncer. Por ejemplo, el cáncer colorrectal se asocia con una mayor metilación de islas CpG en o alrededor de varios genes, incluidos BMP3
 10 NDRG4, Septin 9, TFPI2, p14, EYA2, ALX4, IGFBP7, GATA4/5, MGMT, TBX5, ID4, BTG4, miRNA- 34b/c, CDH13, TPEF/HPP1, NPY, PENK, WIF, EN1, SCTR, INHBB y Vimentin (Zou et al., Cancer Epidemiol Biomarkers Prev, 2007 16: 2686; Melotte et al., J Natl Cancer Inst, 2009 101:916; Grützmann et al., PLoS One, 2008 3:e3759; DeVos et al.,
 Clin Chem, 2009, 55:1337; Wasserkort et al., BMC Cancer, 2013 13:398; Hibi et al., Cancer. Lett, 2011 311:96; Esteller et al., Cancer Res, 2000 60:129; Zou et al., Cancer Epidemiol Biomarkers Prev, 2007 16:2686; Hinoue et al., PLoS
 15 One 2009 4:e8357; Hellebrekers et al., Clin Cáncer Res, 2009 15:3990; Shen et al., J Natl Cancer Inst, 2005 97:1330; Yu et al., Oncogene, 2010 29:6464; Clin Cancer res, 2004 10:7475; Toyota et al., Cancer Res, 2008 68: 4123; Toyooka et al., Cancer Res, 2002 62:3382; Ebert et al., Neoplasia, 2005, 7:771; Chen et al. I., J Natl Cancer Inst, 2005 97:112, Cancer Epidemiol Biomarkers Prev, 2007 16:2686; Roperch et al., BMC Cancer, 2013 13:566; Mayor et al., Br J Cancer, 2009 100:1534). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el presente método se puede diseñar para que sea complementario a una secuencia metilada en esos genes.

20 El carcinoma hepatocelular y/o la cirrosis hepática y/o la infección por el virus de la hepatitis C se correlacionan con una mayor metilación de islas CpG asociadas con p16, p15, RASSF1A, SSBP2, B4GALT1, CASP8, SOCS1, el locus D17S5 de HIC-1, APC, WIF- 1, RUNX-3, DLC-1, SFRP-1, DKK, CDH1, KLK10, OCGR1, DUSP4, NPR1, CYP24A1, CDKN2A, CCNA1, GSTP1, p14, p73, RAR-β, AR, DBCCR1, IRF7 y OCT6 (Wong et al., Cancer Res, 1999 59:71; Chu et al., J Korean Med Sci, 2004 19:83; Wong et al., Clin Cancer Res, 2000 6:3516; Michailidi et al., Gastroenterol Res Pract. , 2014 2014: 597164; Yu et al., BMC Cancer, 2002 2:29; Yoshikawa et al., Nat Genet, 2001 28:29; Kanai et al., Hepatology, 1999 29:703; Liu et al., World J Gastroenterol, 2011 17:4718; Mah et al., Biomark Res, 2014 2:5). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el presente método se puede diseñar para que sea complementario a una secuencia metilada en estas regiones.

30 Del mismo modo, el mieloma múltiple está asociado con una metilación aumentada de las islas CpG en las regiones promotoras de p16 y SOCS-1 (Merlo et al., Nat Med, 1995 1:686; Lo et al., Cancer Res, 1999 59:3899; Wong et al., Clin Cancer Res, 2003 9:1047; Galm et al., Blood, 2003, 101:2784). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario a una secuencia promotora metilada en esos genes. Además, múltiples formas de leucemia se asocian con hipermetilación en p15, p16, p14, p53, DAPK, NES-1, ADAMTS5, WIF-1, SFRP-1, MYOD1, PTPRZ1, PPARG, FOXE3, FBXO39, PKDREJ, TCF3, EGR4, BTG4, PAX5, IKZF1, TLX3, RAG1, POU2AF1, COBL, COL6A2, CPVL, DFNB31, EYA4, FAM24B, FAT1, FUCA2, INADL, MYO3A, PCDHGA12, PON3, ROR1, SYN1, TNIK, ZNF502, CDKN2A, PTPRO, CSMD1, ABI3, SCGB2A1, VHL, GPX3, IGSF4, SERPIND5, ADORA3, AIRE, CARD15, LOC340061, UNC5CL, LDOC1, PRF1, FABP7, SOX11, DLX1, FAM62C, SOX14, RSPO1, ADCY5, HAND2, SPOCK, MLL, ING1, PRIMA1, BCL11B, LTBP2, BNC1, NR2F2, SALL1, GALGT2, LHX1, DLX4, KLK10, TFAP2, APP, FLJ21062, BNIP3, MGMT, RBP1, GATA4, CRABP1, LANCL1, KCNK12, SORL1, CXorf57, SOX9, KIAA0746, ASPHD2, ARHGAP17, PMM2, IL12A, JDP2, PAK1, GALNS, FGD2, LYAR, HOXA9, AHR, ROBO1, NPTX2, CDH1, CDKN2B, HOXD8, MLF-1, PCDH8, CD44, GADD45, ZMAT3, IRF7, KLF6 y p73 (Bodoor et al., Asian Pac J Cancer Prev, 2014 15:75; Zhao et al., Biomark Res, 2013 1:24; Martínez-Delgado et al, Int J Cancer, 2002 102: 15). Por lo tanto, en ciertas realizaciones, el oligonucleótido puente usado en el presente método se puede diseñar para que sea complementario a una secuencia metilada en esos genes.

45 Del mismo modo, el colangiocarcinoma se asocia con hipermetilación de islas CpG asociadas con OPCML, SFRP1, HIC1, PTEN y DcR1 (Sriraksa et al., Br J Cancer, 2011 104: 1313). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario a las islas CpG metiladas de estos genes.

50 Además, el cáncer de pulmón está asociado con una mayor metilación de las islas CpG en las regiones promotoras de p16, DAP quinasa, MGMT, SRBC, GSTP1 (Belinski et al., Proc Natl Acad Sci, 1998 95: 11891; Esteller et al., Cancer Res, 1999 59:67; Esteller et al., Cancer Res, 1999 59:67; Esteller et al., Proc Am Assoc Cancer Res, 1998 39:92; Zöchbauer-Müller et al., Oncogene, 2005 24:6249; Esteller et al., Cancer Res, 1999 59:67; Jain et al., PLoS One, 2012 7: e35789). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario a una secuencia metilada asociada con estos loci.

60 Además, el cáncer de mama está asociado con hipermetilación de islas CpG en SFRP1, SFRP2, SFRP5, BRCA1, LKB1, ER, PR, SYK, RIZ1, GSTP1 (Veeck et al., Oncogene, 2006 25:3479; Veeck et al. , Mol Cancer, 2008 7:83; Veeck et al., Carcinogenesis, 2008 29:991; Esteller et al., J Natl Cancer Inst, 2000 92:564; Esteller et al., Cancer Res, 2001 61:3225; Esteller et al., Oncogene, 2000 19:164; Ottaviano et al., Cancer Res, 1994 54:2552; Lapidus et al., Clin Cancer Res, 1996 2:805; Yuan et al., Cancer Res, 2001 61:5558; Du et al., Cancer Res, 2001 61:8094; Esteller et al., Cancer Res, 1998 58: 4515). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario de una secuencia metilada en una secuencia metilada asociada a estos genes.

65

Asimismo, el carcinoma renal está asociado con una metilación incrementada de islas CpG en las regiones promotoras de GSTP1 y VHL (Esteller et al., *Cancer Res*, 1998 58:4515; Herman et al., *Proc Natl Acad Sci*, 1994 91:9700); el carcinoma endometrial está asociado con una metilación incrementada en las regiones promotoras de hMLH1 (Esteller et al., *Am J Pathol*, 1999 155:1767); y el adenocarcinoma esofágico está asociado con hipermetilación en las regiones promotoras de APC (Kawakami et al., *J Natl Cancer Inst*, 2000 92:1805). Por lo tanto, en ciertas realizaciones, el oligonucleótido puente usado en el método objeto puede diseñarse para que sea complementario a una secuencia metilada en o a una secuencia metilada en las regiones promotoras de estos genes.

Además, el carcinoma de células escamosas oral se asocia con hipermetilación en KIF1A, HOXA9, NID2, EDNRB, p16, RAR β , CDH-1, CYGB y CYCA1 (Guerrero-Preston et al., *Cancer Prev Res (Phila)*, 2011 4: 1061; Shaw et al., *Br J Cancer*, 2006 94:561); mientras que la hipermetilación en EDNRB y DCC están asociadas con lesiones orales premalignas o malignas (Pattani et al., *Cancer Prev Res (Phila)*, 2010 3:1093; Schussel et al., *Clin Cancer Res*, 2013 19:3268). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario a una secuencia metilada en las regiones promotoras de estos genes. Además, el carcinoma de células escamosas del esófago se asocia con una mayor metilación de las regiones reguladoras 5' de los miARN, incluyendo miR-34a, miR-34b/c y miR-129-2 (Chen et al., *Int J Cancer*, 2012 130:1607) Por lo tanto, en ciertos aspectos, el oligonucleótido puente usado en el presente método puede diseñarse para que sea complementario a una secuencia metilada en las regiones reguladoras 5' de estos miARN. El carcinoma de células escamosas esofágico también se asocia con metilación incrementada en GPX3 (He et al., *Dig Dis Sci*, 2011 56:681); y ECRG4 (Yue et al., *World J Gastroenterol*, 2003 9:1174). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario a una secuencia metilada en las regiones promotoras de estos genes.

Del mismo modo, el carcinoma de células escamosas vulvar se asocia con hipermetilación en MGMT, RASSF2A y TSP-1 (Guerrero et al., *Int J Cancer*, 2011 128:2853). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el presente método se puede diseñar para que sea complementario a una secuencia metilada en las regiones promotoras de estos genes. Los carcinomas gástricos asociados al virus de Epstein-Barr se asocian con hipermetilación en loci tales como MINT2, MINT31, p14, p16, p73 y RUNX3 (Saito et al., *J Med Virol*, 2013 85:121). Por lo tanto, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario a una secuencia metilada en esos loci.

Además, el cáncer de próstata está asociado con la hipermetilación en GSTP1, AR, TIMP2 (Lee et al., *Proc Natl Acad Sci*, 1994 85:11733; Jarrard et al., *Cancer Res*, 1998 58:5310; Pulukuri et al., *Oncogene*, 2007 26:5229). Como tal, en ciertas realizaciones, el oligonucleótido puente usado en el presente método se puede diseñar para que sea complementario a una secuencia metilada en estos genes. El retinoblastoma se asocia con secuencias promotoras hipermetiladas de genes, tales como Rb (Stirzaker et al., *Cancer Res*, 1997 57:2229). Por lo tanto, el oligonucleótido puente usado en el presente método se puede diseñar para que sea complementario a una secuencia metilada en las regiones promotoras de estos genes. Asimismo, el glioblastoma multiforme está asociado con hipermetilación de sitios CpG en la región reguladora 5' de THBS1 (Li et al., *Oncogene*, 1999 18:3284). En ciertas realizaciones, el oligonucleótido puente usado en el método objeto se puede diseñar para que sea complementario a una secuencia metilada en estos genes.

En algunas realizaciones, el método objeto se puede emplear para diagnosticar trastornos neurodegenerativos. Por ejemplo, la hipermetilación de los sitios CpG en C9orf72 está asociada con la esclerosis lateral amiotrófica y la degeneración lobular frontotemporal (Xi et al., *Am J Hum Genet*, 2013 92:981; Xi et al., *Am J Hum Genet*, 2014). Por lo tanto, en ciertas realizaciones, un oligonucleótido puente del método objeto se puede diseñar para que sea complementario a una secuencia metilada en los sitios CpG de C9orf72. Del mismo modo, el aumento de la metilación en sitios CpG en los siguientes genes se asocia con la enfermedad de Parkinson: KCTD5, VAV2, MOG, TRIM10, HLA-DQA1, ARHGEF10, GFPT2, HLA-DRB5, TMEM9, MRI1, MAPT, HLA-DRB6, LASS3, GSTTP2, GSTTP1 (Masliah et al., *Epigenetics* 2013 8:1030; Coupland et al., *Mov Disord*, 2013). Como tal, un oligonucleótido puente del método objeto se puede diseñar para que sea complementario a una secuencia metilada en los sitios CpG de esos genes.

Como se indicó anteriormente, en algunos casos la muestra analizada puede ser una muestra de ADN libre de células obtenido de sangre, por ejemplo, de la sangre de una hembra embarazada. En estas realizaciones, el método puede usarse para detectar anomalías cromosómicas en el feto en desarrollo (como se describió anteriormente) o para calcular la fracción de ADN fetal en la muestra, por ejemplo. Estas realizaciones proporcionan la detección y cuantificación de ácido nucleico fetal en una muestra materna en función del estado de metilación del ácido nucleico en la muestra. En algunos casos, la cantidad de ácido nucleico fetal de una muestra materna puede determinarse con relación a la cantidad total de ácido nucleico presente, proporcionando así el porcentaje de ácido nucleico fetal en la muestra. En algunos casos, el número de copias de ácido nucleico fetal se puede determinar en una muestra materna. En algunos casos, la cantidad de ácido nucleico fetal puede determinarse de una manera específica de locus y algunas veces con suficiente sensibilidad para permitir un análisis de dosis cromosómico preciso (por ejemplo, para detectar la presencia o ausencia de una aneuploidía fetal). En algunos casos, el método puede usarse para determinar la concentración de ADN fetal en una muestra materna, por ejemplo, mediante el siguiente método: a) determinar la cantidad total de ADN presente en una muestra materna; determinar la cantidad de un marcador fetal metilado en la muestra de material usando el presente método; y comparando la cantidad de ADN total con la cantidad de ADN del

fabricante fetal. La concentración de ADN fetal en la muestra materna se puede extrapolar de estos resultados. En algunos casos, se puede determinar el número absoluto de copias de ácido nucleico fetal en una muestra materna.

5 En realizaciones particulares, se proporciona un método para estimar la fracción fetal en una muestra de ADN libre de células obtenido de la sangre de una hembra embarazada. En algunos casos, este método comprende; (a) digerir dicha muestra con una endonucleasa de restricción de la familia MspJI para producir una población de fragmentos que están en el intervalo de 20-40 pares de bases de longitud y tienen una citosina metilada central; (b) ligar la secuencia adaptadora A y la secuencia adaptadora B a los extremos respectivos de una pluralidad de fragmentos diana que están solo metilados en ADN fetal, al: (i) hibridar una pluralidad de oligonucleótidos puente de fórmula B'-X'-A' a los fragmentos de (a) en presencia de las secuencias adaptadoras A y B, en donde A' y B' son complementarios a A y B, respectivamente, y la secuencia X' varía entre los diferentes oligonucleótidos puente y cada secuencia X' es complementaria a un fragmento diana que solo está metilado en el ADN fetal; y (ii) ligar la secuencia adaptadora A, la secuencia X y la secuencia adaptadora B entre sí para producir productos de fórmula A-X-B; y (c) cuantificar la cantidad de productos de fórmula AX-B y (d) normalizar la cantidad obtenida en (c) a la cantidad de locus de control en la muestra, donde dicha normalización proporciona una estimación de la fracción fetal en la muestra. En algunos casos, este método puede implementarse usando al menos 100, al menos 200, al menos 300, al menos 500 o al menos 1,000 o más oligonucleótidos de puente diferentes, donde cada oligonucleótido de puente es complementario de una secuencia que solo está metilada en el ADN fetal.

20 En estos casos, la fracción fetal se puede estimar mediante, por ejemplo, la comparación de la cantidad normalizada obtenida en (d) con una curva estándar. En algunos casos, el locus de control puede estar en uno o más de los cromosomas 21, 18 y 13. En algunos casos, la cantidad de producto de ligación cuantificado en (c) se puede comparar con la cantidad de un locus de control que no está metilado diferencialmente en la muestra.

25 En ciertas realizaciones, el ADN de doble cadena que se analiza puede derivarse de una única fuente (por ejemplo, un único sujeto, etc.), mientras que, en otras realizaciones, la muestra de ácido nucleico puede ser un conjunto de ácidos nucleicos extraídos de una pluralidad de fuentes (por ejemplo, un conjunto de ácidos nucleicos de una pluralidad de sujetos, etc.), donde por "pluralidad" se quiere decir dos o más. Como tal, en ciertas realizaciones, una muestra de ácido nucleico puede contener ácidos nucleicos de 2 o más fuentes, 3 o más fuentes, 5 o más fuentes, 10 o más fuentes, 50 o más fuentes, 100 o más fuentes, 500 o más fuentes, 1000 o más fuentes, 5000 o más fuentes, hasta e incluyendo aproximadamente 10,000 o más fuentes. Los códigos de barras moleculares pueden permitir que las secuencias de diferentes fuentes se distingan después de que se analicen.

Kits

35 También se proporcionan en esta divulgación kits para practicar los presentes métodos, como se describió anteriormente. En algunas realizaciones, un kit puede contener al menos: (a) una endonucleasa de restricción de la familia MspJI; (b) secuencia adaptadora A y secuencia adaptadora B; (c) un oligonucleótido puente de fórmula B'-X'-A', donde X', A' y B' son complementarios a X, A y B, respectivamente, y X es una secuencia genómica de 20-40 pares base en longitud y tiene una citosina central que está metilada diferencialmente; y (d) una ligasa. En algunas realizaciones, las secuencias adaptadoras A y B pueden estar presentes en la misma molécula de oligonucleótido. Los diversos componentes del kit pueden estar presentes en contenedores separados o ciertos componentes compatibles pueden precombinarse en un único contenedor, según se desee.

45 Además de los componentes mencionados anteriormente, los kits objeto pueden incluir instrucciones para usar los componentes del kit para practicar los métodos objeto, es decir, instrucciones para el análisis de muestras. Las instrucciones para practicar los métodos objeto generalmente se graban en un medio de grabación adecuado. Por ejemplo, las instrucciones se pueden imprimir en un sustrato, como papel o plástico, etc. Como tal, las instrucciones pueden estar presentes en los kits como un prospecto, en el etiquetado del contenedor del kit o componentes del mismo (es decir, asociado con el empaque o subempaque), etc. En otras realizaciones, las instrucciones están presentes como un archivo de datos de almacenamiento electrónico presente en un medio de almacenamiento legible por ordenador adecuado, por ejemplo, CD-ROM, disquete, etc. En otras realizaciones más, las instrucciones reales no están presentes en el kit, pero se proporcionan medios para obtener las instrucciones de una fuente remota, por ejemplo, a través de Internet. Un ejemplo de esta realización es un kit que incluye una dirección web en la que se pueden ver las instrucciones y/o desde la que se pueden descargar las instrucciones. Al igual que con las instrucciones, este medio para obtener las instrucciones se registra en un sustrato adecuado.

REIVINDICACIONES

1. Un método para estimar la cantidad de un locus metilado en una muestra, que comprende:
 - 5 (a) digerir una muestra de ácido nucleico que contiene copias metiladas y no metiladas de un locus genómico con una endonucleasa de restricción de la familia MspJI para producir una población de fragmentos que están en el intervalo de 20-40 pares base de longitud y tienen una citosina metilada central;
 - 10 (b) ligar la secuencia adaptadora A y la secuencia adaptadora B a los extremos respectivos de un fragmento diana de la secuencia X mediante:
 - (i) hibridar un oligonucleótido puente de fórmula B'-X'-A' a los fragmentos de (a) en presencia de las secuencias adaptadoras A y B, en donde X', A' y B' son complementarios a X, A y B, respectivamente, y
 - 15 (ii) ligar la secuencia A del adaptador, la secuencia X y la secuencia adaptadora B entre sí para producir un producto de fórmula A-X-B; y
 - (c) cuantificar la cantidad de productos de ligación de fórmula A-X-B, proporcionando de ese modo una estimación de la cantidad del locus metilado en la muestra de ácido nucleico.
- 20 2. El método de la reivindicación 1, en el que la endonucleasa de restricción de la familia MspJI es MspJI.
3. El método de cualquier reivindicación anterior, en el que la cuantificación se realiza por secuenciación.
- 25 4. El método de cualquiera de las reivindicaciones 1-2 en el que la cuantificación se realiza mediante qPCR.
5. El método de cualquiera de las reivindicaciones anteriores, en el que las secuencias adaptadoras A y B están presentes en la misma molécula de oligonucleótido y el producto de la etapa (b) es una molécula circular de ácido nucleico.
- 30 6. El método de la reivindicación 5, en el que la cuantificación se realiza mediante:
 - (i) amplificar la molécula de ácido nucleico circular mediante la amplificación de círculo rodante (RCA),
 - 35 (ii) hibridar el producto RCA con una población de oligonucleótidos marcados que se hibridan con múltiples posiciones en el producto RCA; y
 - (iii) contar individualmente el número de complejos RCA marcados.
- 40 7. El método de cualquiera de las reivindicaciones anteriores, en el que la metilación de dicho locus está asociada con una enfermedad.
8. El método de la reivindicación 7, en el que la metilación de dicho locus está asociada con cáncer.
- 45 9. El método de la reivindicación 8, en el que dicho locus es el de BMP3, TFPI1, NDRG4, Septin 9, TFPI2 o Vimentin.
10. El método de la reivindicación 7, en el que la metilación de dicho locus está asociada con una enfermedad hepática.
- 50 11. El método de cualquiera de las reivindicaciones anteriores, en el que la muestra es una muestra de ADN libre de células obtenido de sangre.
12. El método de cualquiera de las reivindicaciones anteriores, en el que la muestra es una muestra de ADN libre de células obtenido de la sangre de una hembra embarazada y el locus metilado solo está metilado en el ADN fetal.
- 55 13. El método de la reivindicación 12, en el que la cantidad del locus metilado en la muestra de ácido nucleico se usa para calcular la fracción de ADN fetal en la muestra.
14. Un método para estimar la fracción fetal en una muestra de ADN libre de células obtenida de la sangre de una hembra embarazada, que comprende:
 - 60 (a) digerir dicha muestra con una endonucleasa de restricción de la familia MspJI para producir una población de fragmentos que están en el intervalo de 20-40 pares base de longitud y tienen una citosina metilada central;
 - (b) ligar la secuencia adaptadora A y la secuencia adaptadora B a los extremos respectivos de una pluralidad de fragmentos diana que solo están metilados en el ADN fetal, mediante:
 - 65

- 5 (i) hibridar una pluralidad de oligonucleótidos puente de fórmula B'-X'-A' a los fragmentos de (a) en presencia de las secuencias adaptadoras A y B, en donde A' y B' son complementarios a A y B, respectivamente, y la secuencia X' varía entre los diferentes oligonucleótidos puente y cada secuencia X' es complementaria a un fragmento diana que solo está metilado en el ADN fetal; y
- 5 (ii) ligar la secuencia adaptadora A, la secuencia X y la secuencia adaptadora B entre sí para producir productos de fórmula A-X-B; y
- 10 (c) cuantificar la cantidad de productos de fórmula A-X-B, y
- 10 (d) normalizar la cantidad obtenida en (c) a la cantidad de uno o más loci de control en la muestra, donde dicha normalización proporciona una estimación de la fracción fetal en la muestra.
- 15 15. El método de la reivindicación 14, en el que la fracción fetal se estima comparando la cantidad normalizada obtenida en (d) con una curva estándar.
- 15 16. El método de cualquiera de las reivindicaciones 14-15, en el que uno o más loci de control están en uno o más de los cromosomas 21, 18 y 13.
- 20 17. El método de cualquiera de las reivindicaciones 14-16, en el que la cantidad de producto de ligación cuantificado en (c) se compara con la cantidad de un locus de control que no está metilado diferencialmente en la muestra.
- 25 18. Un kit que comprende:
- 25 (a) un miembro de la familia MspJI;
- (b) secuencia adaptadora A y secuencia adaptadora B,
- 30 (c) un oligonucleótido puente de fórmula B'-X'-A', donde X', A' y B' son complementarios a X, A y B, respectivamente, y X es una secuencia genómica de 20-40 pares base de longitud y tiene una citosina central que está metilada diferencialmente; y
- (d) una ligasa.
- 35 19. El kit de la reivindicación 18, en el que las secuencias adaptadoras A y B están presentes en la misma molécula de oligonucleótido.

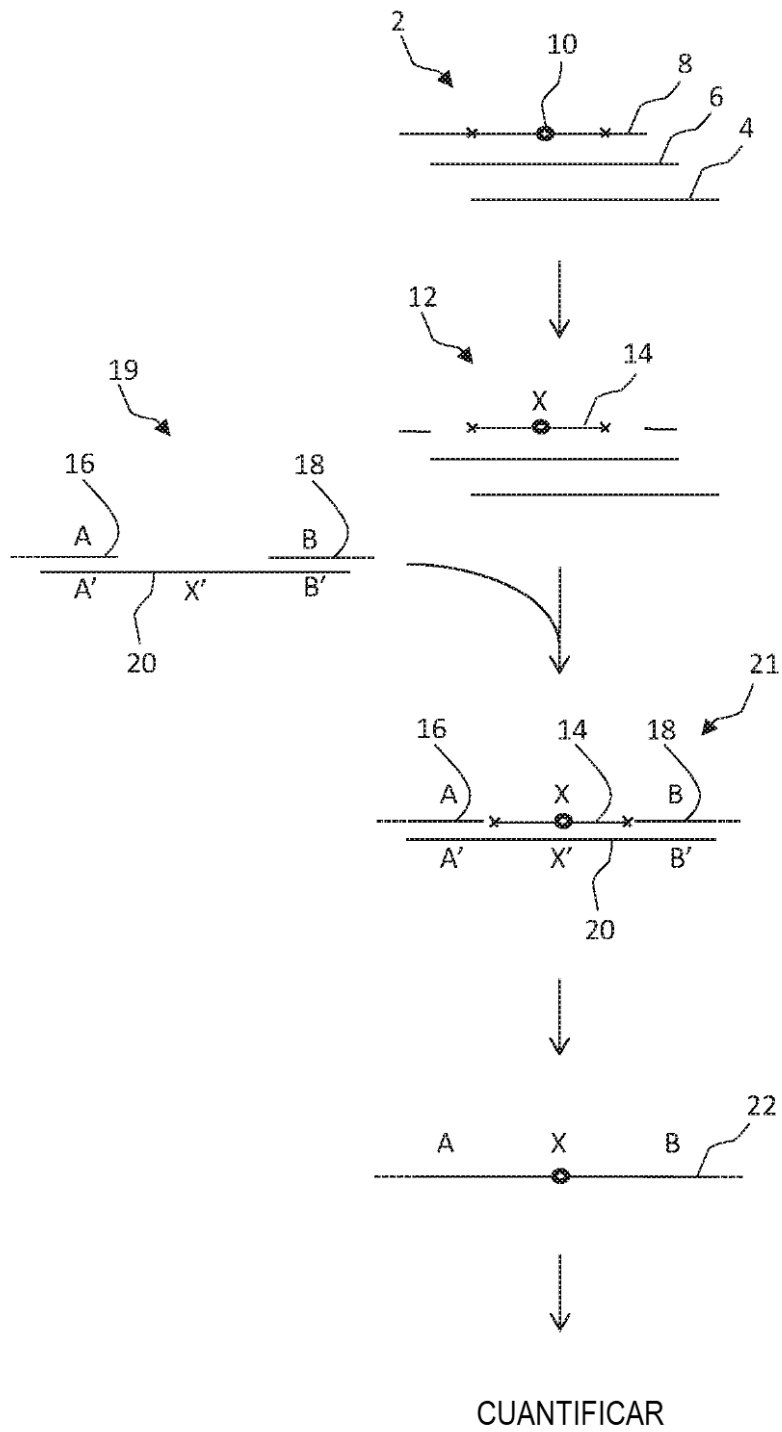


FIG. 1

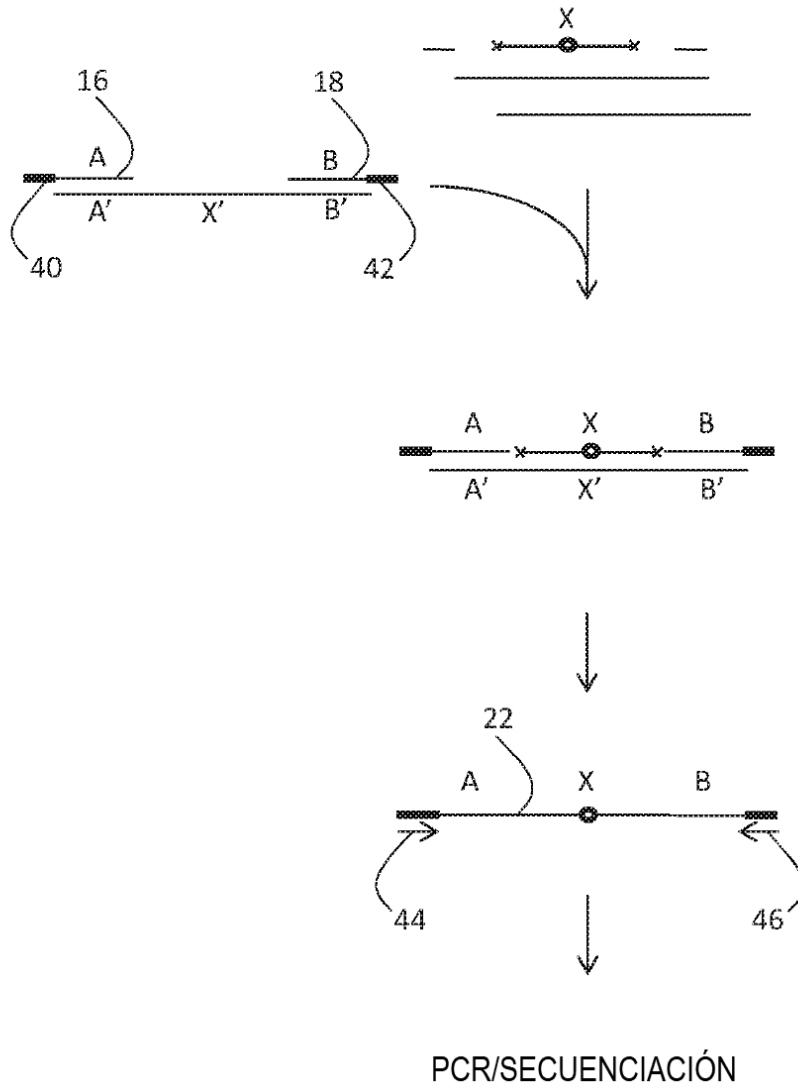


FIG. 2

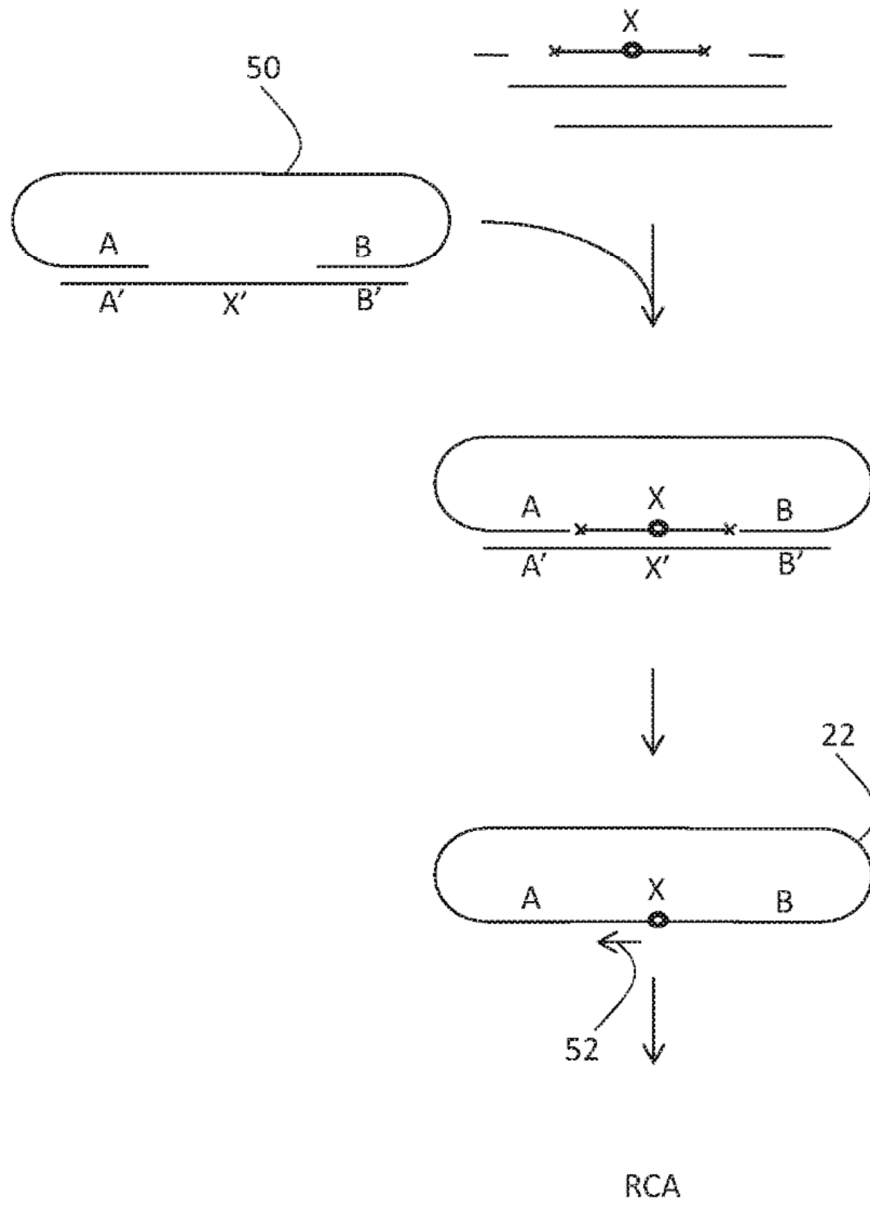


FIG. 3

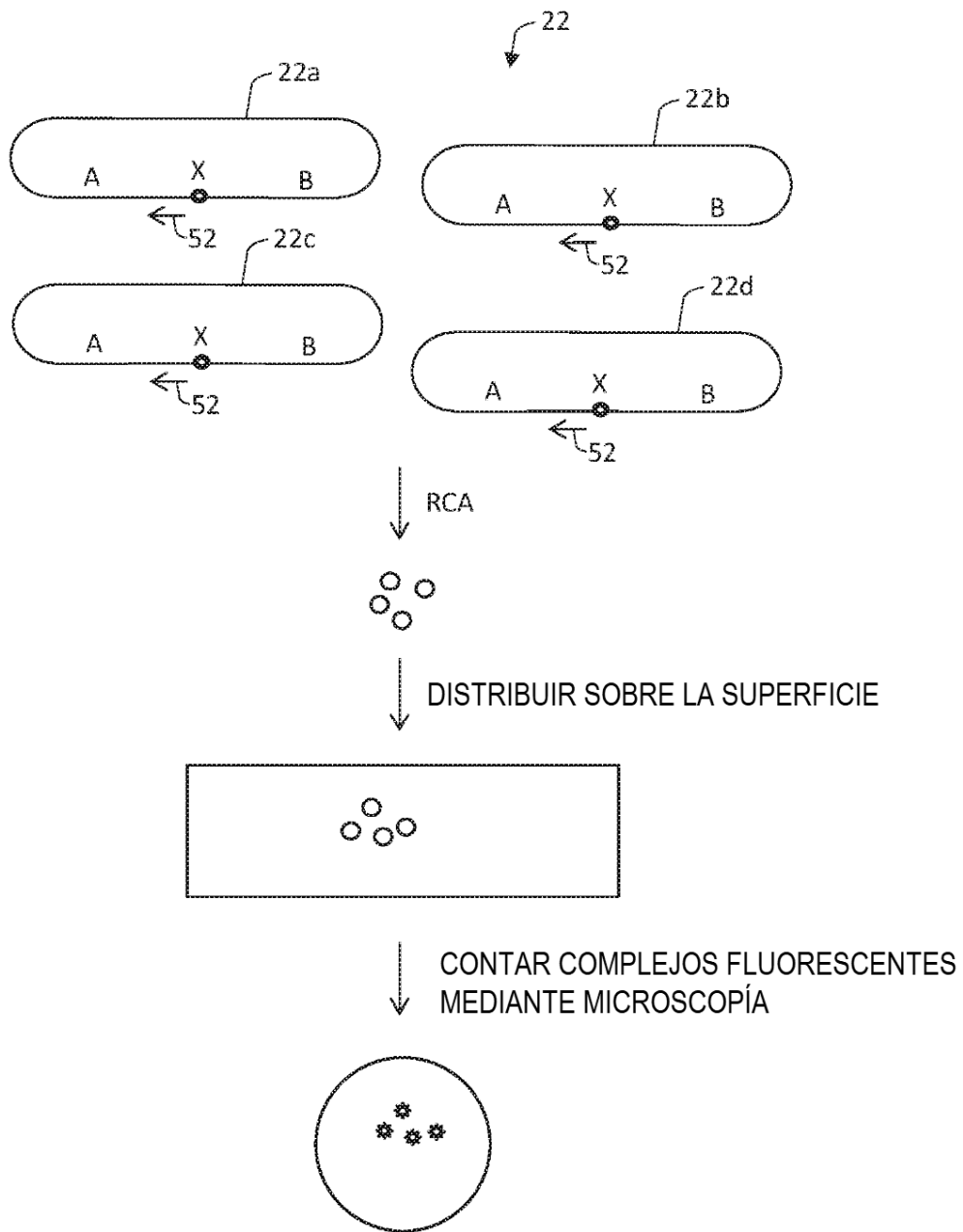


FIG. 4