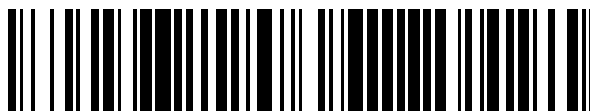


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 788 737**

51 Int. Cl.:

C12Q 1/6858 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.09.2016 PCT/IB2016/055558**

87 Fecha y número de publicación internacional: **23.03.2017 WO17046775**

96 Fecha de presentación y número de la solicitud europea: **16.09.2016 E 16770570 (6)**

97 Fecha y número de publicación de la concesión europea: **19.02.2020 EP 3350342**

54 Título: **Conjunto de sondas para analizar una muestra de ADN y método para usar el mismo**

30 Prioridad:

18.09.2015 US 201562220746 P

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

22.10.2020

73 Titular/es:

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

72 Inventor/es:

**DAHL, CARL OSCAR FREDRIK;
ERICSSON, OLOF JOHN;
KARLSSON, FILIP y
ROOS, FREDRIK**

74 Agente/Representante:

ARIAS SANZ, Juan

ES 2 788 737 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

Conjunto de sondas para analizar una muestra de ADN y método para usar el mismo

Antecedentes

5 El ADN libre ("ADNI") puede analizarse para proporcionar un pronóstico, un diagnóstico o una predicción de una respuesta a un tratamiento para una variedad de enfermedades y estados, incluyendo diversos tipos de cáncer, fracaso o éxito de trasplante, enfermedades inflamatorias, enfermedad infecciosa y aneuploidía fetal.

10 El ADN fetal libre (ADNfl) está presente en la sangre de una mujer embarazada. Este descubrimiento condujo a la posibilidad de realizar pruebas prenatales no invasivas (PPNI) de un feto usando una muestra de sangre de la mujer embarazada. Las pruebas prenatales invasivas (por ejemplo, la amniocentesis o la biopsia de vellosidades coriónicas (BVC)) pueden resultar estresantes para la madre y algunos creen que tales procedimientos pueden aumentar el riesgo de aborto espontáneo. Las PPNI pueden proporcionar información relacionada con una variedad de defectos genéticos, incluyendo el síndrome de Down (trisomía del cromosoma 21), el síndrome de Patau (trisomía 13) y el síndrome de Edwards (trisomía 18). Dichos métodos deben ser muy sólidos, ya que un falso positivo puede conducir a procedimientos médicos innecesarios, y un falso negativo puede privar a la futura madre de comprender las opciones médicas disponibles.

20 Existen muchos obstáculos técnicos asociados con la implementación de una prueba prenatal no invasiva a escala clínica. Por ejemplo, muchos esfuerzos de las PPNI se han centrado en el análisis de ADNfl para identificar cambios en el número de copias en secuencias particulares (por ejemplo, secuencias del cromosoma 21). Sin embargo, tales métodos son difíciles de implementar de manera sólida porque, en parte, la inmensa mayoría del ADNI en una muestra de sangre es de origen materno y en muchos casos sólo una cantidad muy pequeña (por ejemplo, en promedio ~ el 10% y hasta alrededor del 3%) es del feto. Por ejemplo, la presencia o ausencia de una copia adicional de un cromosoma (tal como el cromosoma 21) en el feto puede determinarse comparando el número de copias de secuencias correspondientes al cromosoma 21 con el número de copias de secuencias correspondientes a un cromosoma autosómico. Aunque estos métodos parecen atractivos, representan de hecho un reto, porque la concentración muy pequeña de ADN fetal en relación con el ADN materno en la sangre materna puede ser de tan solo el 3%. Como tal, por cada 1000 secuencias correspondientes al cromosoma 21 que se encuentran en el torrente sanguíneo materno, sólo un pequeño porcentaje de esas secuencias (por ejemplo, 30 secuencias si la fracción fetal es del 3%) proceden del feto. Por tanto, una copia adicional de un cromosoma en el feto sólo conducirá a un aumento relativamente pequeño en el número de secuencias correspondientes a ese cromosoma en el torrente sanguíneo materno. Por ejemplo, si la fracción fetal es de 4, la trisomía 21 fetal sólo conducirá a un aumento del 1,5% en el número de fragmentos correspondientes al cromosoma 21 en el torrente sanguíneo materno. Como resultado de este problema, el rigor estadístico sólo puede lograrse contando un gran número de secuencias correspondientes a una región cromosómica que se sospecha que tiene una diferencia en el número de copias (por ejemplo, al menos 1.000 y a veces al menos 5.000 o más secuencias) y comparando ese número con un número similar para otra región cromosómica que no se sospecha que tenga una diferencia en el número de copias. Poder contar fragmentos de manera sistemática y precisa es primordial para el éxito de muchos métodos de PPNI.

40 Algunos métodos de PPNI usan la reacción en cadena de la polimerasa (PCR) para amplificar el ADN. La PCR se usa ampliamente, pero tiene varias limitaciones que pueden afectar negativamente a la precisión de los resultados. La PCR puede introducir artefactos de secuencia y crear un sesgo de amplificación en una muestra. Los artefactos de secuencia de PCR son errores introducidos en la secuencia de ADN del producto amplificado por PCR mediante la reacción de PCR. Los artefactos de secuencia en la PCR pueden producirse por diversos acontecimientos, tales como la formación de moléculas quiméricas (por ejemplo, dos fragmentos diferentes de ADN unidos extremo con extremo), la formación de ADN heterodúplex (por ejemplo, la hibridación de dos moléculas de ADN diferentes entre sí) y por errores cometidos por la enzima de amplificación (por ejemplo, por la ADN polimerasa Taq que coloca un nucleótido erróneo en el molde de ADN). El sesgo de secuencia a partir de la PCR es una desviación de la distribución de los productos de PCR en comparación con la muestra original. El sesgo de secuencia en la PCR puede producirse por diversos acontecimientos, tales como las diferencias intrínsecas en la eficacia de amplificación de los moldes o la inhibición de amplificación debido al auto-apareamiento de los moldes de ADN. Los errores de PCR dan como resultado una amplificación desigual de las diferentes moléculas de ADN, de modo que la muestra amplificada ya no es representativa de la muestra original. La PCR también es notoriamente sensible a la contaminación por ADN exógeno procedente del entorno. Debido a la amplificación exponencial del ADN durante la PCR, incluso cantidades muy pequeñas de contaminación por ADN exógeno en una reacción de PCR pueden conducir a resultados altamente inexactos. La contaminación por ADN exógeno puede introducirse a partir de gotitas en aerosol que flotan en el aire o puede transferirse a una reacción desde un equipo contaminado.

55 El uso de la amplificación de círculo rodante (RCA) para analizar el ADNI en la sangre materna evita muchos de los problemas asociados con la PCR. Sin embargo, los productos de RCA no son muy fáciles de cuantificar de un modo que proporcione solidez estadística. A nivel práctico, aunque el número absoluto de productos en una reacción de RCA puede ser lo suficientemente alto como para proporcionar solidez estadística, diferentes productos de RCA pueden amplificarse y detectarse con diferentes eficacias y, como tal, detectar de manera sistemática decenas o cientos de miles de productos de RCA de manera uniforme ha supuesto un reto.

60

Los siguientes documentos constituyen técnica anterior:

El documento WO2014/165267 proporciona composiciones y métodos para las pruebas y el análisis de alteraciones genéticas de una muestra que comprende polinucleótidos maternos y fetales.

5 El documento WO2012/019200 proporciona sistemas y métodos de ensayo para la detección de la variación en el número de copias en uno o más loci y la detección de polimorfismo en uno o más loci en una muestra mixta de un individuo.

El documento WO97/10364 proporciona un método de amplificación de un ácido nucleico específico sin necesidad de ciclar una reacción. El método produce transcritos de ARN que pueden detectarse posteriormente. También se proporcionan kits de amplificación y detección.

10 Sumario

En el presente documento se describe, entre otras cosas, un sistema de sondas para analizar una muestra de ácido nucleico. Las sondas pueden diseñarse de tal manera que puedan ligarse a fragmentos diana de ADN genómico (también denominados en el presente documento "secuencias diana" o simplemente "fragmentos") de diferentes loci (por ejemplo, diferentes cromosomas) para producir moléculas de ADN circulares. Las moléculas de ADN circulares, aunque contengan fragmentos de diferentes cromosomas, contienen todas ellas la misma secuencia de "estructura principal". Además, en algunas realizaciones, todas las moléculas de ADN circulares que contienen un fragmento del mismo locus contienen la misma secuencia de identificador específica de locus, es decir, un código de barras específico de locus. En estas realizaciones, las moléculas de ADN circulares pueden amplificarse usando un cebador que se hibrida con una secuencia en la estructura principal, y el locus del que se deriva el fragmento clonado puede detectarse hibridando los productos de RCA con un oligonucleótido marcado que se hibrida con la secuencia de identificador específica de locus. Tal como resultará evidente, esta realización del método puede multiplexarse usando múltiples secuencias de identificador específicas de locus y oligonucleótidos marcados de manera distinguible que se hibridan con esas secuencias. Debido a que todos los productos circulares tienen la misma estructura principal y sólo se diferencian entre sí por la secuencia del fragmento clonado y el código de barras específico de locus, los productos de RCA amplificados a partir de esos productos amplificados de manera sistemática, y el locus al que corresponden esos productos de RCA, pueden detectarse con precisión. También se proporciona un método que emplea el sistema de sonda, así como un kit para poner en práctica el mismo.

Tal como se comentará en mayor detalle a continuación, en determinados casos, el método puede usarse para detectar anomalías cromosómicas (por ejemplo, trisomía 21) en un feto usando una muestra de ADN de una mujer embarazada que lleva el feto.

Se proporciona un sistema de sonda para analizar una muestra de ácido nucleico. En algunas realizaciones, el sistema de sonda puede comprender: (a) un conjunto de oligonucleótidos identificadores de secuencia B; (b) un conjunto de oligonucleótidos de fórmula de fórmula $X'-A'-B'-Z'$, en el que: dentro del conjunto: (i) las secuencias A' y B' varían, y (ii) las secuencias X' y Z' son diferentes entre sí y no son variables; y, dentro de cada oligonucleótido de fórmula: (i) la secuencia A' es complementaria a un fragmento genómico de la muestra de ácido nucleico y (ii) la secuencia B' es complementaria a al menos un miembro del conjunto de oligonucleótidos identificadores; en el que cada oligonucleótido identificador o su secuencia B' complementaria en un oligonucleótido de fórmula identifica (i) un locus en un genoma del que se deriva el fragmento genómico, o (ii) el cromosoma del que se deriva el fragmento genómico; y (c) una o más secuencias de sonda que comprenden X y Z, donde las secuencias X y Z no son variables y se hibridan con las secuencias X' y Z'; en el que cada oligonucleótido de fórmula puede hibridarse con: (i) las secuencias de sonda, (ii) un miembro del conjunto de oligonucleótidos identificadores y, (iii) el fragmento genómico, produciendo de ese modo un complejo que puede ligarse de fórmula X-A-B-Z.

En algunas realizaciones, los diferentes oligonucleótidos identificadores y sus secuencias B' complementarias identifican diferentes cromosomas, por ejemplo, los cromosomas 21, 18 y 13.

45 En algunas realizaciones, el conjunto de oligonucleótidos identificadores puede comprender al menos dos (por ejemplo, dos, tres o cuatro o más) oligonucleótidos identificadores de secuencia B diferente y, dentro del conjunto de oligonucleótidos de fórmula hay al menos 100 secuencias A' diferentes y al menos dos secuencias B' diferentes que son complementarias a al menos dos oligonucleótidos identificadores diferentes.

50 En algunas realizaciones, cada oligonucleótido identificador o su secuencia B' complementaria en un oligonucleótido de fórmula puede corresponder al fragmento genómico.

En algunas realizaciones, cada oligonucleótido identificador o su secuencia B' complementaria en un oligonucleótido de fórmula puede identificar uno o más del cromosoma 21, el cromosoma 18 y el cromosoma 13.

En algunas realizaciones, el fragmento genómico puede ser un fragmento de restricción.

55 En algunas realizaciones, la una o más secuencias de sonda de (c) pueden comprender además un oligonucleótido que comprende la secuencia Y, y en el que el complejo que puede ligarse es lineal.

En algunas realizaciones, el sistema de sonda puede comprender además un par de cebadores de PCR que se hibridan con la una o más sondas de (c).

5 En algunas realizaciones, la una o más secuencias de sonda de (c) pueden comprender una sonda de estructura principal de fórmula X-Y-Z, donde Y comprende una secuencia de oligonucleótido, de manera que el complejo que puede ligarse es un complejo que puede ligarse circular de fórmula X-A-B-Z-Y, donde la secuencia Y se une a las secuencias X y Z.

En algunas realizaciones, el sistema de sonda puede comprender además un cebador de amplificación de círculo rodante que se hibrida con una secuencia en la sonda de estructura principal.

10 En algunas realizaciones, el sistema de sonda puede comprender además (A) un cebador de amplificación de círculo rodante que hibrida una secuencia con la sonda de estructura principal; y (B) hasta cuatro oligonucleótidos de detección marcados de manera distinguible, en el que cada uno de los oligonucleótidos de detección marcados distinguibles se hibrida con una secuencia B'.

15 También se proporciona un método de análisis de muestra. En algunas realizaciones, el método puede comprender: (a) hibridar cualquier realización del sistema de sonda resumido anteriormente con una muestra genómica de prueba que comprende fragmentos genómicos para producir complejos que pueden ligarse de fórmula X-A-B-Z; (b) ligar los complejos que pueden ligarse para producir moléculas de ADN producto de fórmula X-A-B-Z; y (c) contar las moléculas de ADN producto correspondientes a cada secuencia B o B'.

20 En algunas realizaciones, las moléculas de ADN producto pueden ser circulares, y el recuento puede comprender amplificar las moléculas de ADN producto mediante amplificación de círculo rodante, y contar el número de productos de amplificación que comprenden cada secuencia de B o complemento de la misma (B'). En estas realizaciones, el método puede comprender marcar los productos de RCA usando sondas marcadas de manera distinguible que se hibridan con la secuencia B', y el recuento se realiza contando el número de productos de RCA para cada marcador distinguible.

25 En algunas realizaciones, el método puede comprender: i. depositar los productos de RCA sobre un soporte plano; y ii. contar el número de los productos de RCA marcados individuales en una zona del soporte. En estas realizaciones, el soporte puede ser un portaobjetos de vidrio o una membrana capilar transparente porosa, por ejemplo.

30 En algunas realizaciones, las diferentes secuencias de B y sus secuencias B' complementarias identifican diferentes cromosomas, y el método comprende además comparar el número de moléculas de ADN producto que comprenden una primera secuencia de o bien B o bien B' con el número de moléculas de ADN producto que comprenden una segunda secuencia de o bien B o bien B' para determinar si la muestra genómica tiene una aneuploidía.

En algunas realizaciones, el método puede comprender comparar los resultados de recuento de la etapa (c) con los resultados de recuento obtenidos a partir de una o más muestras de referencia.

35 En algunas realizaciones, la muestra genómica de prueba puede ser de una paciente que se sospecha que padece o que está en riesgo de padecer una enfermedad o estado, y los resultados de recuento de la etapa (c) proporcionan una indicación de si la paciente, o el feto de la misma, padece la enfermedad o estado.

En algunas realizaciones, la enfermedad o el estado puede ser un cáncer, una enfermedad infecciosa, una enfermedad inflamatoria, un rechazo de trasplante o una trisomía.

En algunas realizaciones, los fragmentos son fragmentos de restricción.

Breve descripción de las figuras

40 El experto en la técnica entenderá que los dibujos, descritos a continuación, son únicamente para fines de ilustración. No se pretende que los dibujos limiten el alcance de las presentes enseñanzas en modo alguno.

La figura 1 ilustra esquemáticamente algunas de las características del presente sistema de sonda.

La figura 2 ilustra esquemáticamente cómo la secuencia B sirve para identificar el locus de la secuencia A.

La figura 3 ilustra esquemáticamente algunas configuraciones del sistema de sonda a modo de ejemplo.

45 La figura 4 ilustra esquemáticamente algunas de las características de una realización de un método objeto.

La figura 5 ilustra esquemáticamente algunas de las características de una implementación de un método objeto.

La figura 6 ilustra esquemáticamente el diseño de sistemas de sonda.

La figura 7 muestra datos obtenidos usando dos sistemas de sonda diferentes.

La figura 8 muestra datos obtenidos a partir del análisis de muestras clínicas.

Definiciones

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

5 Los intervalos numéricos incluyen los números que definen el intervalo. A menos que se indique de otro modo, los ácidos nucleicos se escriben de izquierda a derecha en orientación de 5' a 3'; y, las secuencias de aminoácidos se escriben de izquierda a derecha en orientación de amino a carboxilo, respectivamente.

A menos que se defina de otro modo, todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el entendido comúnmente por 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 y Sons, Nueva York (1994), y Hale & Markham, THE HARPER COLLINS DICTIONARY OF BIOLOGY, Harper Perennial, N.Y. (1991) proporcionan el significado general de muchos de los términos usados en el presente documento. Aun así, a continuación se definen determinados términos por motivos de claridad y facilidad de referencia.

15 Debe indicarse que, tal como se usa en el presente documento y en las reivindicaciones adjuntas, las formas singulares “un”, “una” y “el/la” incluyen los referentes en plural, a menos que el contexto indique claramente otra cosa. Por ejemplo, el término “un cebador” se refiere a uno o más cebadores, es decir, un cebador único y cebadores múltiples. Se indica además que las reivindicaciones pueden redactarse para excluir cualquier elemento opcional. Como tal, se pretende que esta afirmación sirva como base antecedente para el uso de terminología tal exclusiva como “únicamente”, “solamente” y similares en relación con la mención de elementos de las reivindicaciones, o el uso de una limitación “negativa”.

20 El término “nucleótido” pretende incluir aquellos restos que contienen no sólo las bases conocidas de purina y pirimidina, sino también otras bases heterocíclicas que se han modificado. Dichas modificaciones incluyen purinas o 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 pueden contener no sólo azúcares de ribosa y desoxirribosa convencionales, sino también otros azúcares. Los nucleósidos o nucleótidos modificados también incluyen modificaciones en el resto de azúcar, por ejemplo, en los que uno o más de los grupos hidroxilo se reemplazan con átomos de halógeno o grupos alifáticos, se funcionalizan como éteres, aminas o similares.

30 Los términos “ácido nucleico” y “polinucleótido” se usan indistintamente en el presente documento para describir un polímero de cualquier longitud, por ejemplo, mayor de aproximadamente 2 bases, mayor de aproximadamente 10 bases, mayor de aproximadamente 100 bases, mayor de aproximadamente 500 bases, mayor de 1000 bases, hasta aproximadamente 10.000 o más bases, compuesto por nucleótidos, por ejemplo, desoxirribonucleótidos o ribonucleótidos, y pueden producirse enzimática o sintéticamente (por ejemplo, PNA tal como se describe en la patente estadounidense n.º 5.948.902 y las referencias citadas en ella) que pueden hibridarse con ácidos nucleicos que se producen de manera natural de manera específica de secuencia análoga a la de dos ácidos nucleicos que se producen de manera natural, por ejemplo, pueden participar en interacciones de emparejamiento de bases de Watson-Crick. Los nucleótidos que se producen de manera natural incluyen guanina, citosina, adenina, timina, uracilo (G, C, A, T y U, respectivamente). El ADN y el ARN tienen una estructura principal de azúcar de desoxirribosa y ribosa, respectivamente, mientras que la estructura principal del PNA está compuesta por unidades de repetición de N-(2-aminoetil)-glicina unidas mediante enlaces peptídicos. En el PNA, diversas bases de purina y pirimidina están unidas a la estructura principal mediante enlaces de metilencarbonilo. Un ácido nucleico bloqueado (LNA), a menudo denominado ARN inaccesible, es un nucleótido de ARN modificado. El resto de ribosa de un nucleótido de LNA está modificado con un puente adicional que conecta el oxígeno en 2' y el carbono en 4'. El puente “bloquea” la ribosa en la conformación 3'-endo (Norte), que a menudo se encuentra en los dúplex en forma de A. Los nucleótidos de LNA pueden mezclarse con residuos de ADN o ARN en el oligonucleótido cuando se desee. La expresión “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 de G' y un residuo de C', donde estos residuos corresponden a formas que se producen de manera no natural, es decir, análogos, de G y C, que forman pares de bases entre sí con estabilidad reducida, pero conservan la capacidad para formar pares de bases con residuos de C y G que se producen de manera natural, respectivamente. El ácido nucleico no estructurado se describe en el documento US20050233340, al que se hace referencia en el presente documento para la divulgación de UNA.

55 El término “oligonucleótido” tal como se usa en el presente documento indica un multímero monocatenario de nucleótidos de desde aproximadamente 2 hasta 200 nucleótidos, hasta 500 nucleótidos de longitud. Los oligonucleótidos pueden ser sintéticos o pueden obtenerse enzimáticamente y, en algunas realizaciones, tienen de 30 a 150 nucleótidos de longitud. Los oligonucleótidos pueden contener monómeros de ribonucleótidos (es decir, pueden ser oligorribonucleótidos) o monómeros de desoxirribonucleótidos. Un oligonucleótido puede tener de 10 a 20, de 21 a 30, 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”, tal como se usa en el presente documento, se refiere a un oligonucleótido que puede actuar como un punto de inicio de síntesis cuando se coloca en condiciones en las que se induce la 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 ADN polimerasa y a una temperatura y pH adecuados. El cebador puede ser monocatenario 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, incluyendo la temperatura, el origen del cebador y el uso del método. Por ejemplo, para aplicaciones de diagnóstico, dependiendo de la complejidad de la secuencia o fragmento diana, el cebador oligonucleotídico normalmente contiene 15-25 o más nucleótidos, aunque puede contener menos nucleótidos. Los cebadores en el presente 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 hibridarse con sus respectivas cadenas. Por tanto, no es necesario que la secuencia del cebador refleje la secuencia exacta del molde. Por ejemplo, un fragmento de nucleótido no complementario puede unirse al extremo 5' del cebador, siendo el resto de la secuencia del cebador complementaria a la cadena. Alternativamente, pueden intercalarse bases no complementarias o secuencias más largas en el cebador, siempre que la secuencia del cebador tenga suficiente complementariedad con la secuencia de la cadena para hibridarse con ella y, de este modo formar el molde para la síntesis del producto de extensión.

El término “hibridación” o “se hibrida” se refiere a un proceso en el que una cadena de ácido nucleico se aparea y forma un dúplex estable, ya sea 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 lleva a cabo mediante el apareamiento de dos cadenas de ácido nucleico complementarias en una reacción de hibridación. Puede hacerse que la reacción de hibridación sea altamente específica mediante el ajuste de las condiciones de hibridación (a menudo denominadas rigurosidad de hibridación) en las que tiene lugar la reacción de hibridación, de manera que la hibridación entre dos cadenas de ácido nucleico no formará un dúplex estable, por ejemplo, un dúplex que retiene una región bicatenaria en condiciones de rigurosidad normales, a menos que las dos cadenas de ácido nucleico contengan un determinado número de nucleótidos en secuencias específicas que sean sustancial o completamente complementarias. Las “condiciones de hibridación normal o rigurosidad normal” 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. Tal como se usa en el presente documento, el término “hibridar” o “hibridación” se refiere a cualquier proceso mediante el cual una cadena de ácido nucleico se une con una cadena complementaria a través de emparejamiento de bases.

Se considera que un ácido nucleico “puede hibridarse selectivamente” con una secuencia de ácido nucleico de referencia si las dos secuencias se hibridan específicamente entre sí en condiciones de hibridación y lavado de rigurosidad de moderada a alta. Se conocen condiciones de hibridación de rigurosidad de moderada y alta (véase, por ejemplo, Ausubel, *et al.*, Short Protocols in Molecular Biology, 3ª ed., Wiley & Sons 1995 y Sambrook *et al.*, Molecular Cloning: A Laboratory Manual, Tercera Edición, 2001 Cold Spring Harbor, N.Y.). Un ejemplo de condiciones de alta rigurosidad incluye la hibridación a aproximadamente 42°C en formamida al 50%, SSC 5X, disolución de Denhardt 5X, SDS al 0,5% y ADN portador desnaturalizado 100 ug/ml seguido de lavado dos veces en SSC 2X y SDS al 0,5% a temperatura ambiente y dos veces adicionales en SSC 0,1 X y SDS al 0,5% a 42°C.

La expresión “secuencia de código de barras” o “código de barras molecular”, tal como se usa en el presente documento, se refiere a una secuencia de nucleótidos única utilizada para a) identificar y/o rastrear el origen de un polinucleótido en una reacción y/o b) contar cuántas veces se secuencia una molécula inicial (por ejemplo, en los casos en que sustancialmente cada molécula de 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 código de barras pueden variar ampliamente en tamaño y composición; las siguientes referencias proporcionan una guía para seleccionar conjuntos de secuencias de códigos de barras apropiados para realizaciones particulares: Casbon (Nuc. Acids Res. 2011, 22 e81), Brenner, patente estadounidense n.º 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, patente estadounidense n.º 5.981.179; y similares. En realizaciones particulares, una secuencia de código de barras puede tener una longitud en el intervalo de desde 4 hasta 36 nucleótidos, o desde 6 hasta 30 nucleótidos, o desde 8 hasta 20 nucleótidos.

El término “secuenciación”, tal como se usa en el presente documento, se refiere a un método mediante el cual se obtiene 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.

La expresión “secuenciación de próxima generación” se refiere a las denominadas plataformas de secuenciación por síntesis o secuenciación por ligamiento en paralelo empleadas actualmente, por ejemplo, por Illumina, Life Technologies y Roche, etc. Los métodos de secuenciación de próxima generación también pueden incluir métodos de secuenciación por nanoporos o métodos basados en detección electrónica tales como, por ejemplo, la tecnología Ion Torrent comercializada por Life Technologies.

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

5 Los términos “determinar”, “medir”, “evaluar”, “valorar”, “someter a ensayo” y “analizar” se usan indistintamente en el presente documento para referirse a formas de medición e incluyen determinar si un elemento está presente o no. Estos términos incluyen determinaciones cuantitativas y/o cualitativas. La valoración puede ser relativa o absoluta.

10 La expresión “etiqueta de afinidad”, tal como se usa en el presente documento, se refiere a un resto que puede usarse para separar una molécula a la que se 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 específico, es decir, dos moléculas donde una de las moléculas, a través de medios químicos o físicos, está unida específicamente a la otra molécula. El miembro complementario del par de unión específico, denominado en la presente memoria “agente de captura” puede inmovilizarse (por ejemplo, a una superficie plana, una perla o un soporte de cromatografía) para producir un soporte de cromatografía de afinidad que se une específicamente a la etiqueta de afinidad. Dicho de otro modo, una “etiqueta de afinidad” puede unirse a un “agente de captura”, donde la etiqueta de afinidad se une específicamente al agente de captura, facilitando de ese modo la separación de la molécula a la que está unida la etiqueta de afinidad de otras moléculas que no contienen la etiqueta de afinidad.

15 Tal como se usa en el presente documento, la expresión “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 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.

20 La expresión “nucleótido terminal”, tal como se usa en el presente documento, se refiere al nucleótido o bien en el extremo 5' o bien en el 3' de una molécula de ácido nucleico. La molécula de ácido nucleico puede estar en forma bicatenaria (es decir, ser dúplex) o en forma monocatenaria.

25 El término “ligar”, tal como se usa en el presente documento, se refiere a la unión catalizada enzimáticamente 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.

Los términos “pluralidad”, “conjunto” y “población” se usan indistintamente para referirse a algo que contiene al menos 2 miembros. En determinados casos, una pluralidad puede tener al menos 10, al menos 100, al menos 1000, al menos 10.000, al menos 100.000, al menos 10^6 , al menos 10^7 , al menos 10^8 o al menos 10^9 o más miembros.

30 El término “digestión” pretende indicar un proceso mediante el cual un ácido nucleico se escinde por una enzima de restricción. Para digerir un ácido nucleico, una enzima de restricción y un ácido nucleico que contiene un sitio de reconocimiento para la enzima de restricción se ponen en contacto en condiciones adecuadas para que la enzima de restricción funcione. Se conocen las condiciones adecuadas para la actividad de las enzimas de restricción disponibles comercialmente y se suministran con esas enzimas en el momento de la adquisición.

35 Un “sitio de unión de oligonucleótido” se refiere a un sitio en el que un oligonucleótido se hibrida en un polinucleótido o fragmento diana. Si un oligonucleótido “proporciona” un sitio de unión para un cebador, entonces el cebador puede hibridarse con ese oligonucleótido o su complemento.

40 El término “separación”, tal 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 el otro intacto.

La expresión “región cromosómica de referencia”, tal 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.

45 El término “cadena”, tal como se usa en el presente documento, se refiere a un ácido nucleico constituido por nucleótidos unidos covalentemente entre sí mediante enlaces covalentes, por ejemplo, enlaces fosfodiéster.

50 En una célula, el ADN generalmente existe en una forma bicatenaria y, como tal, tiene dos cadenas complementarias de ácido nucleico denominadas en el presente documento las cadenas “superior” e “inferior”. En determinados casos, las cadenas complementarias de una región cromosómica pueden denominarse cadenas “más” y “menos”, las cadenas “primera” y “segunda”, las cadenas “codificantes” y “no codificantes”, las cadenas de “Watson” y “Crick” o las cadenas “sentido” y “antisentido”. La asignación de una cadena como cadena superior o inferior es arbitraria y no implica ninguna orientación, función o estructura particular. Se conocen las secuencias de nucleótidos de la primera cadena de varias regiones cromosómicas de mamífero a modo de ejemplo (por ejemplo, BAC, conjuntos, cromosomas, etc.) y pueden encontrarse en la base de datos Genbank de NCBI, por ejemplo.

55 La expresión “cadena superior”, tal como se usa en el presente documento, se refiere a cualquier cadena de un ácido nucleico pero no a ambas cadenas de un ácido nucleico. Cuando un oligonucleótido o un cebador se une o se

aparea “solo con una cadena superior”, se une a una sola cadena pero no a la otra. La expresión “cadena inferior”, tal 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 aparea “solo a una cadena”, se une a una sola cadena, por ejemplo, a la primera o a la segunda cadena, pero no a la otra cadena.

- 5 La expresión “unión covalente” se refiere a la producción de un enlace covalente entre dos moléculas independientes, por ejemplo, las cadenas superior e inferior de un ácido nucleico bicatenario. El ligamiento es un tipo de enlace covalente.

10 El término “desnaturalización”, tal como se usa en el presente documento, se refiere a la separación de al menos una parte de los pares de bases de un ácido nucleico dúplex colocando el dúplex en condiciones de desnaturalización adecuadas. Las condiciones de desnaturalización se conocen bien en la técnica. En una realización, para desnaturalizar un ácido nucleico dúplex, 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 determinadas realizaciones, un ácido nucleico puede desnaturalizarse exponiéndolo a una temperatura de al menos 90°C durante una cantidad de tiempo adecuada (por ejemplo, al menos 30 segundos, hasta 30 minutos). Los ácidos nucleicos también pueden desnaturalizarse químicamente (por ejemplo, usando urea o NaOH).

15 Tal como se usa en el presente documento, el término “marcador” se refiere a cualquier átomo o molécula que puede usarse para proporcionar un efecto detectable (preferiblemente cuantificable), y que puede unirse a un ácido nucleico o proteína. Los marcadores incluyen, entre otros, colorantes y radiomarcadores tales 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 los espectros de emisión por transferencia de energía de resonancia de fluorescencia (FRET). Los marcadores pueden proporcionar señales detectables por fluorescencia, radiactividad, 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 tener carga neutra. 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.

20 Los términos “oligonucleótido marcado” y “sonda marcada”, tal como se usan en el presente documento, se refieren a un oligonucleótido que tiene una etiqueta de afinidad (por ejemplo, un resto de biotina), un oligonucleótido modificado con átomos o grupos que permiten la separación o detección (por ejemplo, bromo-desoxiuridina, o partículas de oro coloidal que confieren diferente densidad), y un oligonucleótido modificado con o un marcador detectable ópticamente (por ejemplo, una fluorescencia u otro tipo de marcador emisor de luz). Los oligonucleótidos que contienen solo nucleótidos que se producen de manera natural no son oligonucleótidos marcados.

25 El término “extender”, tal 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 aparea a un ácido nucleico, el ácido nucleico actúa como molde para una reacción de extensión.

35 Tal como se usa en el presente documento, la expresión “extremos respectivos”, en la expresión “ligar un primer y un segundo oligonucleótidos a los extremos respectivos de un fragmento” pretende significar que se añade un oligonucleótido a un extremo del fragmento y se añade otro oligonucleótido al otro extremo del fragmento diana.

40 Tal como se usa en el presente documento, la expresión “adyacentes de manera que pueden ligarse” en el contexto de dos secuencias de oligonucleótidos que son adyacentes de manera que pueden ligarse entre sí, significa que no hay nucleótidos intermedios entre dos oligonucleótidos y que pueden ligarse entre sí.

Tal como se usa en el presente documento, la expresión “oligonucleótido de férula”, tal como se usa en el presente documento, se refiere a un oligonucleótido que, cuando se hibrida con otros dos o más polinucleótidos, actúa como una “férula” para colocar los polinucleótidos uno al lado del otro para que puedan ligarse entre sí, tal como se ilustra en la figura 1.

45 Tal como se usa en el presente documento, la expresión “una molécula de ácido nucleico circular” se refiere a una cadena que está en forma de un círculo cerrado que no tiene extremos libres 3' o 5'.

50 La expresión “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 la expresión. Por ejemplo, un RCA que corresponde a una secuencia en un genoma contiene la misma secuencia de nucleótidos que la secuencia en el genoma.

55 Determinados polinucleótidos descritos en el presente documento pueden denominarse mediante una fórmula (por ejemplo, “X'-A'-B'-Z'”). A menos que se indique de otro modo, los polinucleótidos definidos por una fórmula pueden orientarse en la dirección de 5' a 3' o en la dirección de 5' a 3'. Por ejemplo, los polinucleótidos definidos por la fórmula “X'-A'-B'-Z'” pueden ser “5'-X'-A'-B'-Z'-3'” o “3'-X'-A'-B'-Z'-5'”. Los componentes de la fórmula, por ejemplo, “A”, “X” y “B”, etc., se refieren a secuencias de nucleótidos que pueden definirse por separado dentro de un polinucleótido donde, a menos que esté implícito a partir del contexto (por ejemplo, en el caso de un complejo “que puede ligarse” de una fórmula particular), 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. Según convenio, el complemento de una secuencia mostrada en una fórmula se indicará con una prima (') de manera que el complemento de la secuencia "A" será "A'". Además, a menos que se indique de otro modo o esté implícito a partir del contexto, un polinucleótido definido por una fórmula puede tener una secuencia adicional, un sitio de unión a cebador, un código de barras molecular, un promotor o un espaciador, etc. en su extremo 3', en su extremo 5', o en ambos extremos 3' y 5'. Si un polinucleótido definido por una fórmula se describe como circular, entonces los extremos de esas moléculas están unidos entre sí, ya sea directa o indirectamente. Por ejemplo, en el caso de complejos circulares de fórmula X-A-B-Z-Y, entonces el extremo 5' de la molécula está unido, directa o indirectamente, al extremo 3' de la molécula para producir un círculo. Tal como resultará evidente, las diversas secuencias componentes de un polinucleótido (por ejemplo, A, B, C, X, Y, Z, etc.) pueden ser independientemente de cualquier longitud deseada, siempre que puedan realizar la función deseada (por ejemplo, hibridarse con otra secuencia). Por ejemplo, las diversas secuencias componentes de un polinucleótido pueden tener independientemente una longitud en el intervalo de 8-80 nucleótidos, por ejemplo, 10-50 nucleótidos o 12-30 nucleótidos.
- La expresión "complejo que puede ligarse", por ejemplo, de fórmula X-A-B-Z, se refiere a un complejo en el que los diversos oligonucleótidos son adyacentes de manera que pueden ligarse entre sí (en forma circular o lineal), unidos entre sí por un oligonucleótido de férula, tal como se muestra en la figura 1.
- La expresión "complejo circular que puede ligarse", por ejemplo, de fórmula X-A-B-Z-Y, se refiere a un complejo circular en el que los diversos oligonucleótidos son adyacentes de manera que pueden ligarse entre sí en un círculo, unidos por un oligonucleótido de férula.
- Los términos "locus", "locus genómico", tal como se usan en el presente documento, se refieren a una región definida 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 locus puede ser una región de un cromosoma que es de tan solo 100 kb, y puede ser tan larga como un brazo cromosómico o un cromosoma completo.
- Los términos "primer locus" y "segundo locus" se refieren a diferentes loci, es decir, diferentes regiones en un genoma, por ejemplo, diferentes brazos cromosómicos o diferentes cromosomas.
- Los términos "fragmentos de un locus" se refieren a una población de fragmentos definidos (que pueden obtenerse usando una enzima de restricción o reprogramando una endonucleasa guiada por ARN tal como CAS9) de un locus particular. No es necesario analizar todos los fragmentos de un locus. Debido a que se han publicado las secuencias de diversos genomas, el diseño de oligonucleótidos que se hibridan con un fragmento de un locus es rutinario.
- La expresión "complementario a un fragmento" se refiere a una secuencia que es complementaria a una cadena (ya sea la cadena superior o la inferior) de un fragmento.
- La expresión "secuencia genómica", tal como se usa en el presente documento, se refiere a una secuencia que se produce en un genoma.
- El término "variable", en el contexto de dos o más secuencias de ácido nucleico que son variables, se refiere a dos o más ácidos nucleicos que tienen diferentes secuencias de nucleótidos en uno en relación con el otro. Dicho de otro modo, si los polinucleótidos de una población tienen una secuencia variable o una secuencia particular "varía", entonces la secuencia de nucleótidos de las moléculas de polinucleótido de la población varía de una molécula a otra. No debe considerarse que el término "variable" requiere que cada molécula en una población tenga una secuencia diferente a las otras moléculas en una población.
- Si dos ácidos nucleicos (por ejemplo, las secuencias A y A') son "complementarios", se hibridan entre sí en condiciones de alta rigurosidad. En muchos casos, dos secuencias que son complementarias tienen al menos 10, por ejemplo, al menos 12, al menos 15, al menos 20 o al menos 25 nucleótidos de complementariedad y en determinados casos pueden tener una, dos o tres bases no complementarias.
- El término "identifica", en el contexto de una secuencia que identifica un locus, se refiere a un código de barras molecular que es único para el locus. Una secuencia de este tipo no procede del propio locus, sino que más bien es un código de barras molecular, que habitualmente tiene una secuencia que no está presente en la muestra que está analizándose, que se añade a los fragmentos de un locus que está analizándose y que identifica esos fragmentos como procedentes del locus. Por ejemplo, si los fragmentos de un primer locus se ligan a una primera secuencia de identificador y los fragmentos de un segundo locus se ligan a una segunda secuencia de identificador, entonces puede determinarse el origen de esos fragmentos (el locus al que corresponden) detectando qué secuencia de identificador se ha ligado a esos fragmentos.
- La expresión "orientación invertida" en el contexto de dos secuencias que se hibridan con otras secuencias en una orientación invertida, se refiere a una estructura en la que los extremos 5' y 3' de una de las secuencias se hibridan con la otra de una manera en que los extremos están orientados uno hacia el otro, tal como se ilustra en la parte superior de la figura 3B.

Tal como se usa en el presente documento, la expresión “amplificación de círculo rodante” o “RCA” para abreviar se refiere a una amplificación isotérmica que genera copias concatemerizadas lineales de un molde de ácido nucleico circular usando una polimerasa de desplazamiento de cadena. La RCA se conoce bien en las técnicas de biología molecular y se describe en una variedad de publicaciones incluyendo, pero sin limitarse a Lizardi *et al* (Nat. Genet. 1998 19: 225-232), Schweitzer *et al* (Proc. Natl. Acad. Sci. 2000 97: 10113-10119), Wiltshire *et al* (Clin. Chem. 2000 46: 1990-1993) y Schweitzer *et al* (Curr. Opin. Biotech 2001 12: 21-27).

Tal como se usa en el presente documento, la expresión “productos de amplificación de círculo rodante” se refiere a los productos concatemerizados de una reacción de amplificación de círculo rodante. Tal como se usa en el presente documento, la expresión “productos de amplificación de círculo rodante marcados con fluorescencia” se refiere a productos de amplificación de círculo rodante que se han marcado con fluorescencia, por ejemplo, hibridando un oligonucleótido marcado con fluorescencia con los productos de amplificación de círculo rodante u otros medios (por ejemplo, incorporando un nucleótido fluorescente en el producto durante la amplificación).

Tal como se usa en el presente documento, el término “zona”, en el contexto de una zona de un soporte o una zona de una imagen, se refiere a una zona contigua o no contigua. Por ejemplo, si un método implica contar el número de productos de RCA marcados en una zona, la zona en la que se cuentan los productos de RCA puede ser un solo espacio contiguo o múltiples espacios no contiguos.

Tal como se usa en el presente documento, la expresión “formación de imágenes” se refiere a un proceso mediante el cual señales ópticas procedentes de la superficie de un objeto se detectan y almacenan como datos en asociación con una ubicación (es decir, un “píxel”). Puede reconstruirse una imagen digital del objeto a partir de estos datos. Puede crearse una imagen de una zona de un soporte usando una sola imagen o una o más imágenes.

Tal como se usa en el presente documento, la expresión “productos de RCA marcados individualmente” se refiere a moléculas de RCA individuales que están marcadas.

Tal como se usa en el presente documento, el término “contar” se refiere a determinar el número de objetos individuales en una colección mayor. “Contar” requiere detectar señales independientes de objetos individuales en una pluralidad (no una señal colectiva de la pluralidad de objetos) y luego determinar cuántos objetos hay en la pluralidad contando las señales individuales. En el contexto del presente método, el “recuento” se realiza determinando el número de señales individuales en una matriz de señales.

Tal como se usa en el presente documento, el término “matriz” con referencia a una matriz de productos de RCA se refiere a un conjunto de productos de RCA individuales sobre una superficie plana, donde los productos de RCA están separados espacialmente entre sí en el plano de la superficie (en la medida permitida por la distribución de Poisson si la matriz es verdaderamente aleatoria). Una matriz “aleatoria” es una matriz en la que los elementos, por ejemplo, productos de RCA, se distribuyen sobre la superficie de un sustrato en posiciones que no están predeterminadas. En algunos casos, la distribución de productos de RCA sobre una matriz aleatoria puede describirse mediante estadísticas de Poisson, de modo que, por ejemplo, la distribución de distancias entre productos de RCA de una matriz aleatoria se aproxima mediante una distribución de Poisson.

Pueden aparecer otras definiciones de términos a lo largo de toda la memoria descriptiva

Descripción de realizaciones a modo de ejemplo

Antes de que se describan las diversas realizaciones, debe entenderse que el alcance de la presente invención está limitado solo por las reivindicaciones adjuntas.

Los encabezados de sección usados en el presente documento son solo para fines organizativos y no deben interpretarse como limitativos del contenido descrito en modo alguno.

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

La mención de cualquier publicación es por su divulgación antes de la fecha de presentación y no debe interpretarse como una admisión de que las presentes reivindicaciones no tienen derecho a anticiparse a tal publicación en virtud de una invención anterior. Además, las fechas de publicación proporcionadas pueden ser diferentes de las fechas de publicación reales que puede ser necesario confirmar independientemente.

Tal como resultará evidente para los expertos en la materia al leer esta descripción, cada una de las realizaciones individuales descritas e ilustradas en el presente documento tiene características y componentes diferenciados que pueden separarse fácilmente o combinarse con las características de cualquiera de las otras realizaciones. Cualquier método mencionado puede llevarse a cabo en el orden de acontecimientos citados o en cualquier otro orden que sea lógicamente posible.

Composiciones de sonda

Algunas realizaciones del sistema de sonda pueden comprender: (a) un conjunto de oligonucleótidos identificadores de secuencia B; (b) un conjunto de oligonucleótidos de férula de fórmula X'-A'-B'-Z', en el que: dentro del conjunto: (i) las secuencias A' y B' varían, y (ii) las secuencias X' y Z' son diferentes entre sí y no son variables; y, dentro de cada oligonucleótido de férula: (i) la secuencia A' es complementaria a un fragmento genómico de la muestra de ácido nucleico y (ii) la secuencia B' es complementaria a al menos un miembro del conjunto de oligonucleótidos identificadores; y (c) una o más secuencias de sonda que comprenden X y Z, donde las secuencias X y Z no son variables y se hibridan con las secuencias X' y Z'; en el que cada oligonucleótido de férula puede hibridarse con: (i) las secuencias de sonda, (ii) un miembro del conjunto de oligonucleótidos identificadores y, (iii) el fragmento genómico, produciendo de ese modo un complejo que puede ligarse de fórmula X-A-B-Z. Tal como se describirá en mayor detalle a continuación, en algunas realizaciones los diferentes oligonucleótidos identificadores y sus secuencias B' complementarias identifican diferentes cromosomas, por ejemplo, los cromosomas 21, 18 y 13.

La figura 1 muestra el complejo que puede ligarse de fórmula X-A-B-Z, cuya estructura caracteriza el presente sistema de sonda. Tal como se muestra en la figura 1, en las secuencias complejas X, A, B y Z son adyacentes de manera que pueden ligarse entre sí, manteniéndose en su posición mediante un oligonucleótido de férula. Tal como se indica en la figura 1, la secuencia A es un fragmento diana de un genoma (por ejemplo, una cadena de un fragmento de restricción), y la secuencia B identifica el locus (por ejemplo, una región particular en un cromosoma, un brazo de cromosoma particular o un cromosoma particular, etc.) de la que se deriva la secuencia A adyacente. La relación entre las secuencias A y B se ilustra en la figura 2, que ilustra un conjunto de sonda sencillo, hibridado con diversos fragmentos genómicos (A₁ a A₆). Tal como se muestra en la figura 2, los fragmentos genómicos en los tres complejos superiores (de secuencia A₁, A₂ y A₃) proceden de un primer locus (por ejemplo, el cromosoma 21) y los fragmentos genómicos en los tres complejos inferiores (de secuencia A₄, A₅ y A₆) proceden de un segundo locus (por ejemplo, el cromosoma 18). El locus del que se derivan los fragmentos genómicos en los tres complejos superiores se identifica mediante una sola secuencia (B₁), y el locus del que se derivan los fragmentos genómicos en los tres complejos inferiores se identifica mediante una secuencia diferente (B₂). Las secuencias X y Z son iguales en todos los complejos ilustrados.

Tal como resultará evidente, el conjunto de oligonucleótidos de férula puede ser tan complejo como se desee y, en algunas realizaciones, la secuencia A' puede tener una complejidad de al menos 100, al menos 1.000, al menos 5.000, al menos 10.000 o al menos 50.000 o más, lo que significa que los oligonucleótidos de férula pueden hibridarse colectivamente con al menos 100, al menos 1.000, al menos 5.000, al menos 10.000 o al menos 50.000 o más fragmentos de ADN genómico. La secuencia B' en el conjunto de oligonucleótidos de férula puede ser mucho menos diversa porque simplemente sirve como identificador de locus. Como tal, en el conjunto de oligonucleótidos de férula, la secuencia B' puede tener una complejidad de al menos 2, por ejemplo, 3 ó 4, aunque la secuencia B' puede tener una complejidad de al menos 10, al menos 100 o al menos 1000 en algunas implementaciones. Tal como resultará evidente, puesto que la secuencia B' es complementaria a la secuencia B, la complejidad del conjunto de oligonucleótidos específicos de locus puede ser igual que la complejidad de la secuencia B'. Por ejemplo, si hay tres oligonucleótidos identificadores, puede haber tres secuencias B' diferentes. El número de oligonucleótidos de férula en un conjunto puede variar enormemente, dependiendo de la longitud del locus y del número de fragmentos diana. En algunas realizaciones, cada conjunto de oligonucleótidos de férula puede contener al menos 10, al menos 50, al menos 100, al menos 500, al menos 1.000, al menos 5.000, al menos 10.000 o al menos 50.000 oligonucleótidos de férula diferentes.

Por ejemplo, en algunas realizaciones, un conjunto de oligonucleótidos de férula puede contener: (i) una primera subpoblación de oligonucleótidos de férula que contienen menos de 100 secuencias A', por ejemplo, el conjunto de A_{1,x}, x=1-100+, que son complementarios a diferentes fragmentos de un primer locus (por ejemplo, fragmentos del cromosoma 21 o, por ejemplo, el conjunto de A_{1,x}, x=1-100+), donde cada una de estas subpoblaciones de oligonucleótidos de férula tienen la misma secuencia B', por ejemplo, B₁'; (ii) una segunda subpoblación de oligonucleótidos de férula que contienen menos de 100 secuencias A', por ejemplo, el conjunto de A_{2,x}, x=1-100+, que son complementarios a diferentes fragmentos de un segundo locus (por ejemplo, fragmentos del cromosoma 18 o por ejemplo, el conjunto de A_{2,x}, x=1-100+), donde cada una de estas subpoblaciones de oligonucleótidos de férula tienen la misma secuencia B', por ejemplo, B₂', que es diferente de la secuencia B' de la primera (o de cualquier otra) subpoblación; (iii) una tercera subpoblación de oligonucleótidos de férula que contienen menos de 100 secuencias A', por ejemplo, el conjunto de A_{3,x}, x=1-100+, que son complementarios a diferentes fragmentos de un tercer locus (por ejemplo, fragmentos del cromosoma 18 o, por ejemplo, el conjunto de A_{3,x}, x=1-100+), donde cada una de estas subpoblaciones de oligonucleótidos de férula tienen la misma secuencia B', por ejemplo, B₃', que es diferente de la secuencia B' de cualquier otra subpoblación; (iv) una cuarta subpoblación opcional de oligonucleótidos de férula que contienen menos de 100 secuencias A', por ejemplo, el conjunto de A_{4,x}, x=1-100+, que son complementarios a diferentes fragmentos de un cuarto locus (por ejemplo, fragmentos de otro cromosoma o, por ejemplo, el conjunto de A_{4,x}, x=1-100+) donde cada una de estas subpoblaciones de oligonucleótidos de férula tienen una secuencia B', por ejemplo, B₄', que es diferente de la secuencia B' de cualquier otra subpoblación.

Tal como se ilustra en la figura 3, el sistema de sonda puede configurarse en una variedad de modos diferentes dependiendo de cómo vaya a usarse. Por ejemplo, tal como se ilustra en las figuras 3A, C y D, las secuencias X y Z pueden estar en moléculas diferentes y, como resultado el complejo que puede ligarse es lineal. En estas

realizaciones, la una o más sondas que contienen las secuencias X e Y pueden comprender un primer oligonucleótido que comprende la secuencia X y un segundo oligonucleótido que comprende la secuencia Y. En estas realizaciones, no es necesario que los oligonucleótidos primero y segundo tengan cola, tal como se muestra en la figura 1A. En estas realizaciones, tras el ligamiento, los productos de ligamiento pueden amplificarse usando, por ejemplo, cebadores de PCR comentados que se hibridan con las secuencias X y Z. En algunas realizaciones (tal como se muestra en las figuras 3C y D), los oligonucleótidos primero y/o segundo pueden tener por sí mismos una cola para proporcionar un sitio de unión a cebador para facilitar la amplificación y el recuento. En algunas realizaciones, una cola puede contener un indexador molecular (por ejemplo, una secuencia aleatoria) que permita contar el número de productos de ligamiento originales una vez que esas moléculas se han amplificado y secuenciado. En realizaciones alternativas, y tal como se muestra en la figura 3B, la una o más sondas que contienen secuencias X e Y pueden ser una sola sonda de estructura principal de fórmula X-Y-Z. En estas realizaciones y tal como se muestra, el complejo que puede ligarse es un complejo que puede ligarse circular de fórmula X-A-B-Z-Y, donde la secuencia Y se une a las secuencias X y Z. En otra realización ilustrada en la figura 3E, la una o más sondas que contienen las secuencias X y Z pueden formar parte del propio oligonucleótido de férula. En estas realizaciones, el producto de ligamiento puede tener forma de "mancuerna", tal como se muestra en la figura 3E.

En estas realizaciones, el sistema de sonda puede comprender además un par de cebadores de PCR que se hibridan con la una o más sondas que comprenden las secuencias X y Z, permitiendo de ese modo que se amplifique la parte central del producto de ligamiento (es decir, la parte que contiene las secuencias A y B). En algunas realizaciones, por ejemplo, la realización mostrada en la figura 3B, el sistema de sonda puede comprender además un cebador de amplificación de círculo rodante que se hibrida con una secuencia en la sonda de estructura principal, facilitando de ese modo la amplificación de esos productos mediante la amplificación de círculo rodante. En algunas realizaciones, el sistema de sonda puede comprender un cebador de amplificación de círculo rodante que hibrida una secuencia con la sonda de estructura principal, y hasta cuatro oligonucleótidos marcados de manera distinguible, en los que cada uno de oligonucleótidos marcados de manera distinguible se hibrida con el complemento de una secuencia B'. Esto se explicará en mayor detalle a continuación.

Como tales, algunas realizaciones del sistema de sonda pueden comprender oligonucleótidos de férula, una sonda de estructura principal, y uno o más oligonucleótidos específicos de locus. El sistema de sonda también puede comprender uno o más cebadores de amplificación, tales como un cebador de amplificación de círculo rodante que se hibrida con una secuencia en la sonda de estructura principal o un par de cebadores de PCR que se hibridan con sitios en la sonda de estructura principal, y, opcionalmente, una o más sondas marcadas que se hibridan con el complemento del oligonucleótidos específicos de locus.

Tal como se indicó anteriormente, la secuencia A' varía entre los diferentes miembros del conjunto, y las secuencias de A' se diseñan cada una para ser complementarias con un fragmento diana diferente de un genoma. Las secuencias de A' pueden variar independientemente de longitud y secuencia y, en algún caso, pueden estar en el intervalo de 8 a 80 nucleótidos, por ejemplo, de 10 a 60 nucleótidos, de longitud, dependiendo de la longitud y secuencia de los fragmentos diana. La secuencia B' identifica el locus del que se deriva el fragmento adyacente (por ejemplo, un cromosoma particular tal como el cromosoma 18 o el 21, etc.). La secuencia B' puede ser de cualquier longitud adecuada, pero en algunas realizaciones está en el intervalo de 8 a 30 nucleótidos de longitud. Dentro de cualquier ensayo individual, las secuencias X' y Z' son diferentes entre sí, y no son variables. La secuencia X' y Z' puede ser de cualquier longitud adecuada, pero en algunas realizaciones están independientemente en el intervalo de 8 a 30 nucleótidos de longitud, aunque pueden usarse secuencias más largas o más cortas. La longitud total de los oligonucleótidos de férula puede estar en el intervalo de 50 a 200 nucleótidos. En algunas realizaciones, los oligonucleótidos de férula pueden biotinilarse, permitiendo de ese modo que los productos de ligamiento (comentados a continuación) se aislen de otros productos no ligados antes de la amplificación. Tal como resultará evidente, las secuencias X y Z (que pueden ser de cualquier longitud adecuada, pero que en algunas realizaciones están independientemente en el intervalo de 8 a 30 nucleótidos de longitud) no son variables y se hibridan con las secuencias X' y Z'. El oligonucleótido específico de locus es de secuencia B que, de nuevo, puede ser de cualquier longitud adecuada, por ejemplo, en el intervalo de 8 a 30 nucleótidos de longitud.

Tal como se indicó anteriormente, los complejos producidos usando el sistema de sonda descrito anteriormente pueden ser lineales o circulares (tal como se muestra en la figura 3). La figura 4 ilustra algunas de las características de la realización circular ilustrada en la figura 3B.

Tal como se muestra en la figura 4, en algunas realizaciones, el sistema de sonda puede comprender el conjunto de oligonucleótidos de férula 2 (de fórmula X'-A'-B'-Z', que puede estar en la orientación de 5' a 3' o de 3' a 5'), una sonda de estructura principal 6 de fórmula X-Y-Z, donde las secuencias X y Z no son variables y se hibridan con las secuencias X' y Z' en una orientación invertida (es decir, de modo que los extremos de la estructura principal apunten uno hacia el otro, tal como se muestra), y un conjunto de oligonucleótidos específicos de locus 8 es de secuencia B. La secuencia Y en la sonda de estructura principal puede ser de cualquier longitud conveniente, por ejemplo, de 20 a 100 nucleótidos. La longitud total de la sonda de estructura principal 6 puede estar en el intervalo de 50 a 300 nucleótidos de longitud, o más en determinados casos.

Tal como se muestra en la figura 4, el conjunto de sonda se caracteriza porque los diversos oligonucleótidos pueden

5 hibridarse con fragmentos genómicos para producir un primer conjunto de complejos circulares que pueden ligarse
 10 (es decir, un complejo en el que los extremos de la sonda de estructura principal 6, un oligonucleótido específico
 de locus 8 y un fragmento genómico 4 son adyacentes de manera que pueden ligarse entre sí y se mantienen
 adyacentes de manera que pueden ligarse entre sí mediante un oligonucleótido de férula 2). Tal como se muestra en
 el ejemplo ilustrado, la sonda de estructura principal 6, el oligonucleótido específico de locus 8 y el fragmento 4 se
 hibridan con el primer oligonucleótidos de férula 2 para producir un conjunto de complejos circulares que pueden
 ligarse 10 de fórmula X-A-B-Z-Y, donde la secuencia Y se une a las secuencias X y Z. Los fragmentos 4 que están
 presentes en este conjunto de complejos circulares que pueden ligarse 10 pueden proceder de al menos 2, al menos
 5, al menos 10 o al menos 50 o más loci diferentes (por ejemplo, cromosomas diferentes), y la identidad del locus
 del que se deriva un fragmento adyacente (por ejemplo, los cromosomas particulares) se proporciona por el
 oligonucleótido específico de locus 8, que tiene la misma secuencia para cada locus. En este ejemplo, las
 secuencias A y A' (que corresponden a las secuencias de diferentes fragmentos genómicos) varían, B y B' (el
 identificador de locus) varían, y las secuencias X, Y y Z no varían.

15 Tal como se describirá en mayor detalle a continuación, en esta realización, el sistema de sonda (que comprende un
 primer conjunto de oligonucleótidos de férula 2, una sonda de estructura principal 6 y un oligonucleótido específico
 de locus 8) pueden hibridarse con una muestra que comprende fragmentos de un genoma 4 para producir un primer
 conjunto de complejos circulares que pueden ligarse de fórmula X-A-B-Z-Y 10, tal como se muestra. Tras el
 ligamiento de los complejos circulares que pueden ligarse para producir un primer conjunto de moléculas de ADN
 circulares 12 de fórmula X-A-B-Z-Y, el primer conjunto de moléculas de ADN circulares puede amplificarse mediante
 20 amplificación de círculo rodante (RCA) para producir un primer conjunto de productos de RCA 16. La RCA puede
 realizarse usando un cebador de amplificación de círculo rodante 14 que se hibrida con una secuencia en la sonda
 de estructura principal 6, tal como se ilustra en la figura 4, o cebadores de PCR que se hibridan con sitios que
 flanquean el fragmento ligado. Como tal, en determinadas realizaciones, el sistema de sonda puede comprender
 adicionalmente un cebador de amplificación de círculo rodante 14, cebador que se hibrida con una secuencia en la
 25 sonda de estructura principal 6, o un par de cebadores de PCR que se hibridan con sitios que flanquean el
 fragmento ligado. Tras la RCA, puede determinarse entonces el "origen" del fragmento clonado en un producto de
 RCA 16 particular (es decir, el locus, por ejemplo, el cromosoma particular, del que se deriva el fragmento genómico
 clonado) hibridando un primer oligonucleótido 18 marcado con el complemento de secuencia B (es decir, B'), o
 mediante secuenciación. Tal como resultará evidente, el oligonucleótido marcado 18 puede comprender al menos
 algo de la secuencia B. Como tal, en determinadas realizaciones, el sistema de sonda puede comprender
 30 adicionalmente un oligonucleótido marcado que se hibrida con el complemento de primer oligonucleótido específico
 de locus 8.

35 Tal como resultará evidente, si van a detectarse secuencias de dos o más loci diferentes en la misma reacción, el
 sistema de sonda puede comprender oligonucleótidos marcados de manera distinguible adicionales, uno por cada
 identificador de locus B, de modo que ambos conjuntos de productos de RCA pueden identificarse al mismo tiempo.
 En estas realizaciones, el sistema de sonda puede comprender además hasta cuatro oligonucleótidos marcados de
 manera distinguible (por ejemplo, B₁, B₂, B₃, B₄), donde cada uno de los oligonucleótidos marcados de manera
 distinguible se hibrida con el complemento de una secuencia B' (por ejemplo, B₁', B₂', B₃', B₄').

40 Tal como resultará evidente, los fragmentos con los que se hibridan los oligonucleótidos de férula son fragmentos de
 restricción del genoma que está analizándose. Además, cualquiera de las sondas, oligonucleótidos o cebadores
 descritos anteriormente (por ejemplo, la sonda de estructura principal) puede contener un código de barras
 molecular (por ejemplo, una secuencia de indexación tal como una secuencia aleatoria o semialeatoria) de manera
 que cada molécula de ADN circular puede distinguirse mediante la combinación del fragmento clonado y el código
 de barras, permitiendo de ese modo contar cuántas moléculas iniciales se secuenciaron, incluso después de que las
 45 moléculas se hayan amplificado (véase, por ejemplo, Casbon *et al*).

Métodos

50 También se proporciona en el presente documento un método que comprende: (a) hibridar un sistema de sonda tal
 como se describió anteriormente, con una muestra genómica de prueba que comprende fragmentos de un genoma
 para producir complejos que pueden ligarse de fórmula X-A-B-Z; (b) ligar los complejos que pueden ligarse para
 producir moléculas de ADN producto de fórmula X-A-B-Z; y (c) contar las moléculas de ADN producto
 correspondientes a cada identificador de locus de secuencia B. En algunas realizaciones, el recuento puede
 realizarse secuenciando las moléculas de ADN producto, o productos de amplificación de los mismos, para producir
 lecturas de secuencia, y contar el número de lecturas de secuencia que comprende cada secuencia de B.

55 En realizaciones en las que las moléculas de ADN producto son circulares, el recuento puede comprender amplificar
 las moléculas de ADN producto mediante amplificación de círculo rodante, y contar el número de productos de
 amplificación que comprenden cada secuencia de B. En estas realizaciones, el método puede comprender marcar
 los productos de RCA usando sondas marcadas de manera distinguible que se hibridan con la secuencia B, y el
 recuento se realiza contando el número de productos de RCA para cada marcador distinguible. Los principios
 generales de una implementación de este método se muestran en la figura 4. Tal como resultará evidente, los
 60 fragmentos con los que se hibridan los oligonucleótidos de férula pueden ser (independientemente) fragmentos de
 restricción de cadenas superior o inferior del genoma que está analizándose. Estos fragmentos pueden generarse

digiriendo el genoma con una o más enzimas de restricción (por ejemplo, una combinación de enzimas que tienen cuatro secuencias de reconocimiento de bases), y luego desnaturalizando la muestra digerida. Como tal, los fragmentos que se clonan tienen extremos definidos, permitiendo de ese modo el diseño de oligonucleótidos de férula para clonar esos fragmentos. Hay otros modos de generar fragmentos que tienen extremos definidos (por ejemplo, métodos que usan endonucleasa Flap, exonucleasa, relleno de huecos, etc.).

Tal como se indicó anteriormente, este método puede multiplexarse para proporcionar un modo de analizar dos o más loci diferentes, tal como se muestra en la figura 5. Con referencia a la figura 5, una muestra que contiene fragmentos de ADN genómico 40 puede: a) hibridarse con un sistema de sonda 42 que comprende: (i) un primer conjunto de sondas de férula, tal como se describió anteriormente; (ii) un primer oligonucleótido específico de locus, tal como se describió anteriormente; (iii) un segundo conjunto de sondas de férula, tal como se describió anteriormente; (iv) un segundo oligonucleótido específico de locus, tal como se describió anteriormente; y, (v) una sonda de estructura principal, tal como se describió anteriormente, para producir una mezcla 44 que comprende un primer conjunto de complejos circulares que pueden ligarse de fórmula X-A-B-Z-Y (que contienen fragmentos del primer locus, por ejemplo, un primer cromosoma, así como fragmentos de un segundo locus, por ejemplo, un segundo cromosoma). A continuación, el método comprende (b) ligar los complejos circulares que pueden ligarse para producir una mezcla de moléculas de ADN circulares 46 (que contienen los conjuntos de moléculas de ADN circulares primero y segundo), y, tras tratar la muestra con una exonucleasa para eliminar las moléculas de ácido nucleico moléculas lineales, (c) amplificar las moléculas de ADN circulares 46 mediante amplificación de círculo rodante usando un solo cebador que se hibrida con la sonda de estructura principal, para producir productos de RCA 48. El locus de cada uno de los fragmentos contenidos dentro de cada producto de RCA puede identificarse mediante la hibridación de los productos de RCA con sondas de oligonucleótido primero y segundo marcados de manera distinguible, que se hibridan con el complemento del oligonucleótido específico de locus que está presente en cada uno de los productos, para producir una muestra marcada 50. En estas realizaciones, el método puede comprender: (d) detectar por separado: (i) productos de RCA que contienen fragmentos de un primer locus usando una sonda marcada que se hibrida con una primera secuencia de identificador de locus y (ii) productos de RCA que contienen fragmentos de un segundo locus usando una sonda marcada que se hibrida con una segunda secuencia de identificador de locus, en el que las sondas marcadas se marcan de manera distinguible. Tal como se indicó anteriormente, tras el ligamiento, si los oligonucleótidos de férula están biotinilados, los productos circulares pueden aislarse de productos no ligados usando, por ejemplo, perlas de estreptavidina. En cualquier caso, la muestra ligada puede tratarse con una exonucleasa, eliminando de ese modo las moléculas de ADN lineales de la reacción. Este principio puede ampliarse para contar el número de productos de ligamiento producidos por cualquier número de loci (por ejemplo, 3, 4, hasta 10 o hasta 100 o más loci).

En algunas realizaciones, la etapa de detección puede (d) comprender: (i) depositar los productos de RCA sobre un soporte; y, (ii) contar por separado el número de los productos de RCA marcados individuales que se marcan con un marcador y el número de productos de RCA marcados individuales marcados con otro marcador en una zona del soporte. Tal como se entenderá, la hibridación de los oligonucleótidos marcados puede realizarse antes de que los productos de RCA se distribuyan sobre el soporte, o después de que los productos de RCA se distribuyan sobre el soporte.

Dicho de otro modo, el número de productos de amplificación de círculo rodante correspondientes a cada locus puede estimarse, por ejemplo, distribuyendo los productos de RCA sobre la superficie de un soporte (un portaobjetos o membrana porosa), hibridando los productos de RCA usando oligonucleótidos marcados (por ejemplo, oligonucleótidos marcados con fluorescencia) y luego contando el número de señales diferenciadas en una zona del soporte, por ejemplo, usando un lector de fluorescencia. El marcaje puede realizarse antes o después de que los productos se hayan distribuido sobre el soporte y, puesto que cada producto de RCA contiene miles de copias de las mismas secuencias, debe haber miles de sitios de unión para los oligonucleótidos marcados, aumentando de ese modo la señal. En realizaciones de multiplexación (por ejemplo, en las que van a contarse productos de RCA correspondientes a dos loci diferentes), los productos de RCA correspondientes a un locus pueden marcarse con un fluoróforo y los productos de RCA correspondientes a otro locus pueden marcarse con un fluoróforo diferente, permitiendo de ese modo que los diferentes productos de RCA se cuenten por separado.

En determinadas realizaciones, el método puede comprender (a) filtrar una muestra líquida que contiene los productos de amplificación de círculo rodante (RCA) a través de una membrana capilar transparente porosa, concentrando de ese modo los productos de RCA y produciendo una matriz de los productos de RCA sobre la membrana; (b) marcar con fluorescencia los productos de RCA antes o después de la etapa (a); y, (c) contar el número de los productos de RCA marcados individuales en una zona de la membrana, proporcionando de ese modo una estimación del número de los productos de RCA marcados en la muestra. En algunas realizaciones, la membrana capilar transparente porosa puede ser una membrana de óxido de aluminio anódica porosa. En estas realizaciones, la etapa de marcaje (b) puede realizarse hibridando oligonucleótidos marcados con fluorescencia con los productos de RCA, antes o después de la etapa (a). En determinadas realizaciones, el método puede comprender obtener imágenes de una zona de la membrana para producir una o más imágenes y contar el número de los productos de RCA marcados individuales en la una o más imágenes. Ejemplos de tales métodos se describen en el documento PCT/IB2016/052495, presentado el 2 de mayo de 2016.

La cuantificación de señales de productos de RCA individuales es importante porque, en muchas aplicaciones (por

ejemplo, en el diagnóstico prenatal no invasivo mediante análisis de ADN), es necesario determinar el número de fragmentos correspondientes a cromosomas particulares (por ejemplo, el cromosoma 21) de manera precisa y sin sesgo. Los métodos de análisis típicos usan PCR que, tal como se conoce bien, es un procedimiento muy sesgado en que algunas secuencias se amplifican con eficacias mucho mayores que otras. Esto hace que las estrategias basadas en PCR no resulten prácticas para muchos esfuerzos de diagnóstico.

En realizaciones particulares, la muestra puede contener múltiples poblaciones de productos de RCA (por ejemplo, dos, tres o cuatro o más poblaciones de productos de RCA, tal como una primera población de productos de RCA marcados y una segunda población de productos de RCA), donde las diferentes poblaciones de productos de RCA se marcan de manera distinguible, lo que significa que los miembros individuales de cada una de las poblaciones de marcadores de productos de RCA pueden detectarse y contarse independientemente, incluso cuando las poblaciones son mixtas. Los pares de marcadores fluorescentes que pueden distinguirse adecuados útiles en los métodos objeto incluyen, por ejemplo, Cy-3 y Cy-5 (Amersham Inc., Piscataway, NJ), Quasar 570 y Quasar 670 (Biosearch Technology, Novato CA), Alexafluor555 y Alexafluor647 (Molecular Probes, Eugene, OR), BODIPY V-1002 y BODIPY V1005 (Molecular Probes, Eugene, OR), POPO-3 y TOTO-3 (Molecular Probes, Eugene, OR), y POPRO3 TOPRO3 (Molecular Probes, Eugene, OR). Marcadores detectables de manera distinguible adecuados adicionales pueden encontrarse en, por ejemplo, Kricka *et al.* (Ann Clin Biochem. 39:114-29, 2002). Por ejemplo, los productos de RCA pueden marcarse con cualquier combinación de ATTO, ALEXA, CY, o colorantes de cianina diméricos tales como YOYO, TOTO etc. También pueden usarse otros marcadores.

En algunos casos, una población de productos de RCA puede marcarse de manera distinguible marcándola con múltiples marcadores, aumentando de ese modo las posibilidades de multiplexación. Por ejemplo, en algunos casos una población puede marcarse con dos colorantes que pueden distinguirse (por ejemplo, Cy3 y Cy5) que, cuando se leen, podrán distinguirse de poblaciones que se marcan con los colorantes individuales (por ejemplo, Cy3 o Cy5). En algunas realizaciones, una primera población de productos de RCA representa una población de productos de RCA marcados "de prueba" y una segunda población de productos de RCA representa una población de productos de RCA "de referencia" con el que puede compararse el número de los primeros productos de RCA. Por ejemplo, en algunas realizaciones, una primera población de productos de RCA puede corresponder a una primera región cromosómica (por ejemplo, un primer cromosoma tal como el cromosoma 21) y una segunda población de productos de RCA puede corresponder a una segunda región cromosómica (por ejemplo, un segundo cromosoma tal como el cromosoma 13 o el 18 o una región diferente del primer cromosoma) y el número de la primera población de productos de RCA y la segunda población de productos de RCA puede contarse y compararse para determinar si existe una diferencia en el número de copias de las regiones (lo que indica que existe duplicación o delección de la región de prueba). En algunas realizaciones, la muestra contiene al menos una primera población de productos de RCA y una segunda población de productos de RCA, en la que las poblaciones primera y segunda de productos de RCA se marcan de manera distinguible en la etapa de marcaje (b)). En estas realizaciones, el método comprende contar el número de primeros productos de RCA marcados en una zona de la membrana y contar el número de segundos productos de RCA marcados en una zona (la misma zona o una zona diferente) de la membrana, proporcionando de ese modo una estimación del número de las poblaciones primera y segunda de productos de RCA en la muestra. Esta realización puede implicar además comparar el número de primeros productos de RCA en la muestra con el número de segundos productos de RCA en la muestra.

En algunas de estas realizaciones del método, el método puede comprender obtener imágenes de las poblaciones primera y segunda de productos de RCA marcados para producir una o más imágenes (por ejemplo, una primera imagen y una segunda imagen, respectivamente) y, opcionalmente, (i) contar el número de productos de RCA marcados en la una o más imágenes, proporcionando de ese modo una estimación del número de poblaciones primera y segunda de productos de RCA marcados en la muestra. Las poblaciones primera y segunda de productos de RCA marcados pueden detectarse por separado usando métodos conocidos (por ejemplo, usando filtros apropiados, etc.). Estas realizaciones del método pueden comprender además comparar el número de primeros productos de RCA marcados en la muestra con el número de segundos productos de RCA marcados en la muestra. Esta etapa del método puede implicar contar al menos 1.000 (por ejemplo, al menos 5.000, al menos 10.000, al menos 20.000, al menos 50.000, al menos 100.000, al menos 500.000 hasta 1 M o más) productos de RCA marcados en la primera población, al menos 1.000 (por ejemplo, al menos 5.000, al menos 10.000, al menos 20.000 o al menos 50.000, al menos 100.000, al menos 500.000 hasta 1 M o más) productos de RCA marcados en una zona de la membrana y contar, garantizando de ese modo que pueda calcularse una diferencia en el número de copias con rigor estadístico.

En realizaciones alternativas, fragmentos clonados en las moléculas de ADN (y, opcionalmente, cualquier secuencia de indexación en las moléculas de ADN circulares) pueden amplificarse mediante PCR usando cebadores de PCR que se hibridan con o son iguales que los sitios que flanquean esas secuencias. En esta realización, un producto de PCR puede amplificarse usando los cebadores. En esta realización, la cantidad del producto puede cuantificarse mediante cualquier ensayo de qPCR adecuado, por ejemplo, un ensayo de TaqMan o similar. En otra realización, el producto puede secuenciarse (con o sin amplificación). En estas realizaciones, la cantidad de moléculas circulares correspondientes a cada locus puede estimarse contando el número de lecturas de secuencia correspondientes al locus (por ejemplo, contando cuántas lecturas de secuencia tienen una secuencia de código de barras específica de locus particular). En algunas realizaciones, si se usa una secuencia de indexación, el número de moléculas circulares correspondientes a cada locus puede contarse determinando cuántas secuencias de código de barras

moleculares diferentes están asociadas con cada secuencia de código de barras específica de locus.

Tal como resultará evidente, en esta realización, los cebadores usados pueden contener secuencias que son compatibles con el uso en, por ejemplo, el método del terminador reversible de Illumina, el método de pirosecuenciación de Roche (454), la secuenciación por ligamiento de Life Technologies (la plataforma SOLiD) o la plataforma Ion Torrent de Life Technologies. Ejemplos de tales métodos se describen en las referencias siguientes: 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* (Brief 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), que se refieren a descripciones generales de los métodos y a las etapas particulares de los métodos, incluyendo todos los productos de partida, reactivos y productos finales para cada una de las etapas.

La muestra genómica de prueba puede ser de una paciente que se sospecha que padece o que está en riesgo de padecer una enfermedad o estado, y el resultado de la etapa (c) una indicación de si la paciente, o el feto de la misma, padece la enfermedad o el estado. En algunas realizaciones, la enfermedad o el estado puede ser un cáncer, una enfermedad infecciosa, una enfermedad inflamatoria, un rechazo de trasplante o un defecto cromosómico tal como una trisomía.

Tal como se indicó anteriormente, en algunos casos la muestra que está analizándose usando este método puede ser una muestra de ADN obtenida de sangre, por ejemplo, de la sangre de una mujer embarazada. En estas realizaciones, el método puede usarse para detectar anomalías cromosómicas en el feto en desarrollo (tal como se describió anteriormente) o para calcular la fracción de ADN fetal en la muestra, por ejemplo.

Las anomalías ilustrativas en el número de copias que pueden detectarse usando el método incluye, pero no se limitan a, trisomía 21, trisomía 13, trisomía 18, trisomía 16, XXY, XYY, XXX, monosomía X, monosomía 21, monosomía 22, monosomía 16 y monosomía 15. En la siguiente tabla se enumeran anomalías adicionales en el número de copias que pueden detectarse usando el presente método.

Cromosoma	Anomalía y asociación con enfermedad
X:	XO (síndrome de Turner)
Y:	XXY (síndrome de Klinefelter)
Y:	XYY (síndrome de doble Y)
Y:	XXX (síndrome de trisomía X)
Y:	XXXX (síndrome de cuatro X)
Y:	Deleción de Xp21 (síndrome de Duchenne/Becker, hipoplasia suprarrenal congénita, enfermedad granulomatosa crónica)
Y:	Deleción de Xp22 (deficiencia de esteroide sulfatasa)
Y:	Deleción de Xq26 (enfermedad linfoproliferativa ligada al cromosoma X)
1:	1p, somática (neuroblastoma)
1:	monosomía (neuroblastoma)
1:	trisomía (neuroblastoma)
2:	monosomía (retraso del crecimiento, retraso mental y del desarrollo, y anomalías físicas secundarias)
2:	trisomía 2q (retraso del crecimiento, retraso mental y del desarrollo y anomalías físicas secundarias)
3:	monosomía (linfoma no Hodgkin)
3:	trisomía, somática (linfoma no Hodgkin)
4:	monosomía (leucemia no linfocítica aguda (ANLL))
4:	trisomía, somática (leucemia no linfocítica aguda (ANLL))
5:	5p (Cri du chat; síndrome de Lejeune)
5:	5q, somática (síndrome mielodisplásico)
5:	monosomía (síndrome mielodisplásico)
5:	trisomía (síndrome mielodisplásico)
6:	monosomía (sarcoma de células claras)
6:	trisomía, somática (sarcoma de células claras)
7:	deleción de 7q11.23 (síndrome de William)
7:	monosomía (síndrome de monosomía 7 de la infancia; somática: adenomas corticales renales; síndrome mielodisplásico)
7:	trisomía (síndrome de monosomía 7 de la infancia; somática: adenomas corticales renales; síndrome mielodisplásico)
8:	deleción de 8q24.1 (síndrome de Langer-Giedon)
8:	monosomía (síndrome mielodisplásico; síndrome de Warkany; somática: leucemia mielógena crónica)
8:	trisomía (síndrome mielodisplásico; síndrome de Warkany; somática: leucemia mielógena crónica)
9:	monosomía 9p (síndrome de Alfi)
9:	monosomía 9p (síndrome de Rethore)

9:	trisomía parcial (síndrome de Rethore)
9:	trisomía (síndrome de trisomía 9 completa; síndrome de trisomía 9 en mosaico)
10:	monosomía (ALL o ANLL)
10:	trisomía, somática (ALL o ANLL)
11:	11p- (aniridia; tumor de Wilms)
11:	11q- (síndrome de Jacobsen)
11:	monosomía (linajes mieloides afectados (ANLL, MDS))
11:	trisomía, somática (linajes mieloides afectados (ANLL, MDS))
12:	monosomía (CLL, tumor de células de la granulosa juvenil (JGCT))
12:	trisomía, somática (CLL, tumor de células de la granulosa juvenil (JGCT))
13:	13q- (síndrome 13q; síndrome de Orbeli)
13:	deleción de 13q14 (retinoblastoma)
13:	monosomía (síndrome de Patau)
13:	trisomía (síndrome de Patau)
14:	monosomía (trastornos mieloides (MDS, ANLL, CML atípica)
14:	trisomía, somática (trastornos mieloides (MDS, ANLL, CML atípica)
15:	deleción de 15q11-q13 (Prader-Willi, síndrome de Angelman)
15:	monosomía (Prader-Willi, síndrome de Angelman)
15:	trisomía, somática (linajes mieloides y linfoides afectados, por ejemplo, MDS, ANLL, ALL, CLL)
16:	deleción de 16q13.3 (Rubenstein-Taybi)
16:	monosomía (carcinomas de células renales papilares (malignos))
16:	trisomía, somática (carcinomas de células renales papilares (malignos))
17:	17p-somática (síndrome 17p en tumores malignos mieloides)
17:	deleción de 17q11.2 (Smith-Magenis)
17:	17q13.3 (Miller-Dieker)
17:	monosomía (adenomas corticales renales)
17:	trisomía, somática (adenomas corticales renales)
17:	17p11.2-12 (síndrome de Charcot-Marie-Tooth tipo 1; HNPP)
17:	trisomía (síndrome de Charcot-Marie-Tooth tipo 1; HNPP)
18:	18p- (síndrome de monosomía parcial 18p o síndrome de Grouchy-Lamy-Thieffry)
18:	18q- (síndrome de Grouchy-Lamy-Salmon-Landry)
18:	monosomía (síndrome de Edwards)
18:	trisomía (síndrome de Edwards)
19:	monosomía (síndrome de Edwards)
19:	trisomía (síndrome de Edwards)
20:	20p- (síndrome de trisomía 20p)
20:	deleción de 20p11.2-12 (Alagille)
20:	20q- (somática: MDS, ANLL, policitemia vera, leucemia neutrofílica crónica)
20:	monosomía (carcinomas de células renales papilares (malignos))
20:	trisomía, somática (carcinomas de células renales papilares (malignos))
21:	monosomía (síndrome de Down)
21:	trisomía (síndrome de Down)
22:	deleción de 22q11.2 (síndrome de DiGeorge, síndrome velocardiofacial, síndrome de anomalía conotruncal facial, síndrome de Opitz G/BBB autosómico dominante, síndrome cardiofacial de Caylor)
22:	monosomía (síndrome de trisomía 22 completa)
22:	trisomía (síndrome de trisomía 22 completa)

El método descrito en el presente documento puede emplearse para analizar ADN genómico prácticamente de cualquier organismo incluyendo, pero sin limitarse a, plantas, animales (por ejemplo, reptiles, mamíferos, insectos, gusanos, peces, etc.), muestras de tejido, bacterias, hongos (por ejemplo, levaduras), fagos, virus, tejido cadavérico, muestras arqueológicas/antiguas, etc. En determinadas realizaciones, el ADN genómico usado en el método puede derivarse de un mamífero, donde en determinadas realizaciones el mamífero es un ser humano. En realizaciones a modo 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 estar compuesta por células en cultivo 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 de la escena de un crimen). En realizaciones particulares, la muestra de ácido nucleico puede obtenerse de una muestra biológica tal como células, tejidos, fluidos corporales y heces. Los fluidos corporales de interés incluyen, pero no se limitan a, sangre, suero, plasma, saliva, moco, flema, líquido cefalorraquídeo, líquido pleural, lágrimas, líquido del conducto quilífero, linfa, esputo, líquido cefalorraquídeo, líquido sinovial, orina, líquido amniótico y semen. En realizaciones particulares, puede obtenerse una muestra de un sujeto, por ejemplo, un ser humano. En algunas realizaciones, la muestra analizada puede ser una muestra de ADNI obtenida de sangre, por ejemplo, de la sangre de una mujer embarazada.

Por ejemplo, en algunas realizaciones, puede obtenerse una muestra de ADN y la muestra puede digerirse con una o más enzimas de restricción (o una endonucleasa guiada por ARN tal como cas9) para producir fragmentos

predecibles (cuya mediana de tamaño puede estar en el intervalo de 20-100 bases). El método descrito anteriormente puede realizarse en el ADN digerido, y el número de fragmentos correspondientes a un locus (por ejemplo, un cromosoma) puede compararse con el número de fragmentos correspondientes a otro locus (por ejemplo, otro cromosoma) usando el método descrito en el presente documento. Tal como se indica, el método puede usarse para identificar diferencias en el número de copias, por ejemplo, aneuploidías cromosómicas, que están asociadas con una enfermedad o estado.

Tal como se indicó anteriormente, en algunos casos la muestra analizada puede ser una muestra de ADN obtenida de sangre, por ejemplo, de la sangre de una mujer embarazada. En estas realizaciones, el método puede usarse para detectar anomalías cromosómicas en el feto en desarrollo o para calcular la fracción de ADN fetal en la muestra, por ejemplo.

Kits

Mediante esta descripción también se proporcionan kits para la puesta en práctica de los métodos objeto, tal como se describió anteriormente. En determinadas realizaciones, el kit puede comprender: (a) un conjunto de oligonucleótidos de fórmula de $X'-A'-B'-Z'$, en el que: dentro del conjunto: (i) la secuencia de A' y B' varía, y (ii) las secuencias de X' y Z' son diferentes entre sí y no son variables; y dentro de cada molécula: (i) la secuencia A' es complementaria a un fragmento de un genoma y (ii) la secuencia B' identifica el locus del que se deriva el fragmento genómico que se hibrida con la secuencia de A' adyacente; (b) una o más sondas que comprenden las secuencias X y Z, en las que: i. las secuencias X y Z no son variables y se hibridan con las secuencias X' y Z'; y (c) un conjunto de oligonucleótidos específicos de locus de secuencia B; y en el que: cada oligonucleótido de fórmula de (a) puede hibridarse con (i) las secuencias de sonda de (b); (ii) un oligonucleótido específico de locus de (c); y, (iii) un fragmento genómico de (a), para producir un complejo que puede ligarse de fórmula X-A-B-Z, en el que la secuencia B identifica el locus de la secuencia A adyacente. En algunas realizaciones, la una o más sondas de (b) comprenden un primer oligonucleótido que comprende la secuencia X y un segundo oligonucleótido que comprende la secuencia Y. En algunas realizaciones, el kit puede comprender además un par de cebadores de PCR que se hibridan con la una o más sondas que comprenden las secuencias X e Y. En determinadas realizaciones, la una o más sondas de (b) son una sonda de estructura principal de fórmula X-Y-Z, y el complejo que puede ligarse es un complejo que puede ligarse circular de fórmula X-A-B-Z-Y, donde la secuencia Y se une a las secuencias X y Z, y la secuencia B identifica el locus de la secuencia A adyacente. En estas realizaciones, el kit puede comprender además un cebador de amplificación de círculo rodante que se hibrida con una secuencia en la sonda de estructura principal. En estas realizaciones, el kit puede comprender una pluralidad de oligonucleótidos marcados de manera distinguible, en el que cada uno de los oligonucleótidos marcados de manera distinguible se hibrida con el complemento de una secuencia B'. El kit puede contener adicionalmente una ligasa y/o una polimerasa de desplazamiento de cadena para realizar amplificación de círculo rodante.

Los diversos componentes del kit pueden estar presentes en recipientes independientes o determinados componentes compatibles (por ejemplo, los conjuntos primero y segundo de sondas de fórmula y las sondas específicas de locus primera y segunda) pueden combinarse previamente en un solo recipiente, según se desee.

Además de los componentes mencionados anteriormente, el kit objeto puede incluir además instrucciones para el uso de los componentes del kit para la puesta en práctica del método objeto.

Ejemplos

Los siguientes ejemplos se presentan para proporcionar a los expertos habituales en la técnica una divulgación y una descripción adicionales de cómo realizar y usar la presente invención, y no pretenden limitar el alcance de lo que los inventores consideran su invención, que es la de las reivindicaciones adjuntas, ni pretenden indicar que los experimentos a continuación son todos o los únicos experimentos realizados.

EJEMPLO I

Método inicial de validación de datos

El fin de este experimento es comparar los métodos que usan oligonucleótidos de estructura principal que son específicos de cromosoma (por ejemplo, el oligonucleótido de estructura principal usado para capturar fragmentos de un primer cromosoma, por ejemplo, el cromosoma 21, es diferente del oligonucleótido de estructura principal usado para capturar fragmentos de un segundo cromosoma, por ejemplo, el cromosoma 18, tal como se describe en los documentos WO2015083001 y WO2015083002), con métodos en los que se usa el mismo oligonucleótido de estructura principal para todos los cromosomas examinados. Esto se ilustra en la figura 6. Tal como se muestra, en el diseño "nuevo", el origen del fragmento clonado se determina usando una secuencia específica de cromosoma (por ejemplo, A o B) que se clona en el mismo producto circular que el fragmento diana. En el método nuevo, se usa un solo oligonucleótido de estructura principal (en comparación con múltiples oligonucleótidos de estructura principal en el método anterior), y los fragmentos clonados de todos los cromosomas pueden amplificarse usando el mismo cebador de RCA o un solo par de cebadores de PCR.

Se digirió ADN de línea celular (10 ng) desnaturalizado y se hibridó con los diseños de sonda "antiguo" y "nuevo".

5 Tras la hibridación y el ligamiento, las regiones de ligamiento se sometieron a tratamiento con exonucleasa para eliminar el ADN no circularizado en la disolución. Los productos circulares restantes sirvieron como moldes en una reacción de RCA, que produjo copias concatémicas de los productos circulares. Estos productos de RCA se marcaron con oligonucleótidos marcados con fluorescencia complementarios a la secuencia "de férula" y se depositaron en un soporte sólido para su detección.

Trece muestras de ADNI de mujeres embarazadas se sometieron a la misma reacción tal como se describió anteriormente.

10 Para todas las reacciones, se contó el número de objetos individuales (productos de RCA) en cada color. Se calculó la razón del número de objetos en color A/B para cada muestra y se calculó el coeficiente de variación como una medida de precisión del ensayo. El bajo coeficiente de variación permite mediciones precisas de muestras con baja fracción fetal. Esto se ilustró añadiendo muestras que contenían una baja cantidad de adición conocida de muestra de línea celular con trisomía 21.

Según los datos mostrados en la figura 7, el diseño nuevo genera un CV menor tanto para el ADN de línea celular como para el ADNI, lo que permite una medición más precisa del ADN fetal con anomalías cromosómicas.

15 Sin querer restringirse a ninguna teoría particular, se cree que este método puede ser menos sensible para impurezas en la muestra.

EJEMPLO II

Análisis de muestras clínicas

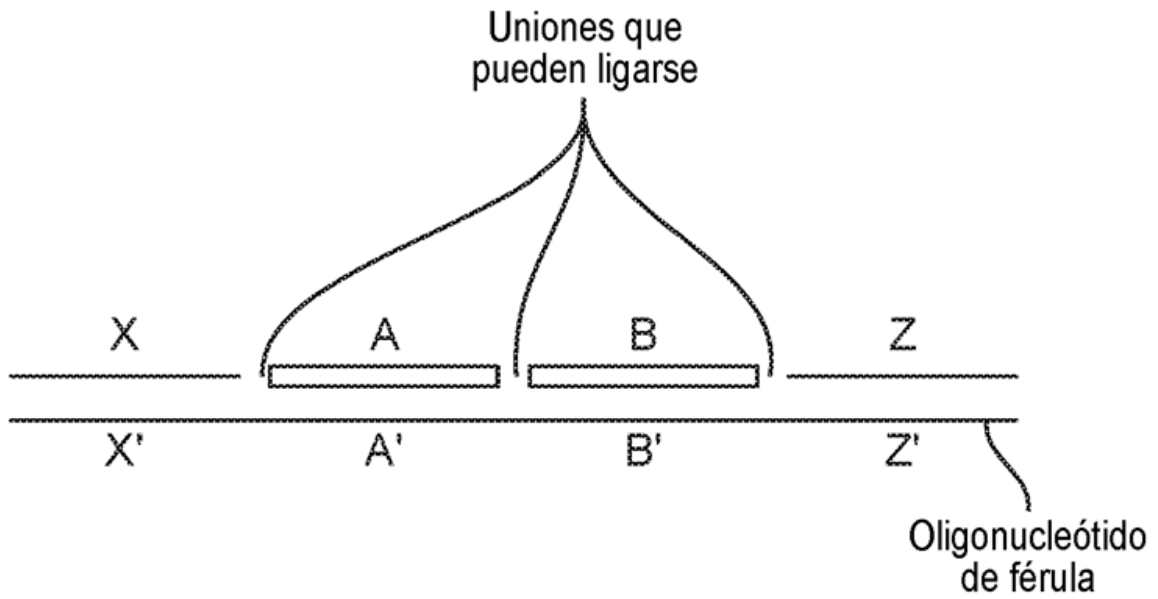
20 Se prepararon muestras de ADNI de 26 mujeres embarazadas normales y 4 mujeres que llevan un feto con trisomía 21. Se centrifugó sangre (10 ml) de cada paciente para separar el plasma de los glóbulos rojos y la capa leucocitaria. El plasma correspondiente (~3-5 ml/paciente) se sometió a un protocolo de extracción de ADN basado en perlas, lo que dio como resultado ADNI extraído diluido en 50ul de tampón.

25 Entonces se sometió el ADNI al método en el presente documento descrito anteriormente y se analizó mediante recuento digital de los productos de círculo rodante usando un microscopio de fluorescencia. Los 4 casos positivos se detectaron por encima de una puntuación z mayor de 3. Se calculó que el CV de las muestras normales era del 0,49%, lo que demuestra la alta precisión del ensayo.

REIVINDICACIONES

1. Sistema de sonda para analizar una muestra de ácido nucleico, que comprende:
 - (a) un conjunto de oligonucleótidos identificadores de secuencia B;
 - (b) un conjunto de oligonucleótidos de férula de fórmula $X'-A'-B'-Z'$, en el que:
 - 5 dentro del conjunto: (i) las secuencias A' y B' varían, y (ii) las secuencias X' y Z' son diferentes entre sí y no son variables; y,
 - dentro de cada oligonucleótido de férula: (i) la secuencia A' es complementaria a un fragmento genómico de la muestra de ácido nucleico y (ii) la secuencia B' es complementaria a al menos un miembro del conjunto de oligonucleótidos identificadores;
 - 10 en el que cada oligonucleótido identificador o su secuencia B' complementaria en un oligonucleótido de férula identifica (i) un locus en un genoma del que se deriva el fragmento genómico, o (ii) el cromosoma del que se deriva el fragmento genómico; y
 - (c) una o más secuencias de sonda que comprenden X y Z, donde las secuencias X y Z no son variables y se hibridan con las secuencias X' y Z';
 - 15 en el que cada oligonucleótido de férula puede hibridarse con: (i) las secuencias de sonda, (ii) un miembro del conjunto de oligonucleótidos identificadores y, (iii) el fragmento genómico, produciendo de ese modo un complejo que puede ligarse de fórmula X-A-B-Z.
2. Sistema de sonda según cualquier reivindicación anterior, en el que el conjunto de oligonucleótidos identificadores comprende al menos dos oligonucleótidos identificadores de secuencia B diferente y, dentro del conjunto de oligonucleótidos de férula hay: al menos 100 secuencias A' diferentes; y, al menos dos secuencias B' diferentes que son complementarias a los al menos dos oligonucleótidos identificadores diferentes.
3. Sistema de sonda según cualquier reivindicación anterior, en el que cada oligonucleótido identificador o su secuencia B' complementaria en un oligonucleótido de férula corresponde al fragmento genómico.
- 25 4. Sistema de sonda según cualquier reivindicación anterior, en el que cada oligonucleótido identificador o su secuencia B' complementaria en un oligonucleótido de férula identifica uno o más del cromosoma 21, el cromosoma 18 y el cromosoma 13.
5. Sistema de sonda según cualquier reivindicación anterior, en el que el fragmento genómico es un fragmento de restricción.
- 30 6. Sistema de sonda según cualquier reivindicación anterior, en el que la una o más secuencias de sonda de (c) comprenden además un oligonucleótido que comprende la secuencia Y, y en el que el complejo que puede ligarse es lineal.
7. Sistema de sonda según cualquier reivindicación anterior, que comprende además un par de cebadores de PCR que se hibridan con la una o más sondas de (c).
- 35 8. Sistema de sonda según cualquiera de las reivindicaciones 1-5, en el que la una o más secuencias de sonda de (c) comprenden una sonda de estructura principal de fórmula X-Y-Z, donde Y comprende una secuencia de oligonucleótido, de manera que el complejo que puede ligarse es un complejo que puede ligarse circular de fórmula X-A-B-Z-Y, donde la secuencia Y se une a las secuencias X y Z.
9. Sistema de sonda según la reivindicación 8, que comprende además un cebador de amplificación de círculo rodante que se hibrida con una secuencia en la sonda de estructura principal.
- 40 10. Sistema de sonda según la reivindicación 8, que comprende además:
 - (A) un cebador de amplificación de círculo rodante que hibrida una secuencia con la sonda de estructura principal; y
 - (B) hasta cuatro oligonucleótidos de detección marcados de manera distinguible, en el que cada uno de los oligonucleótidos de detección marcados distinguibles se hibrida con una secuencia B'.
 - 45
11. Método que comprende:
 - (a) hibridar un sistema de sonda según cualquier reivindicación anterior con una muestra genómica de prueba que comprende fragmentos genómicos, para producir complejos que pueden ligarse de fórmula X-A-B-Z;

- (b) ligar los complejos que pueden ligarse para producir moléculas de ADN producto de fórmula X-A-B-Z; y
- (c) contar las moléculas de ADN producto correspondientes a cada secuencia B o B'.
- 5 12. Método según la reivindicación 11, en el que las moléculas de ADN producto son circulares, y el recuento comprende amplificar las moléculas de ADN producto mediante amplificación de círculo rodante, y contar el número de productos de amplificación que comprenden cada secuencia de B o B'.
13. Método según la reivindicación 12, en el que el método comprende marcar los productos de RCA usando sondas marcadas de manera distinguible que se hibridan con la secuencia B', y el recuento se realiza contando el número de productos de RCA para cada marcador distinguible.
- 10 14. Método según la reivindicación 13, en el que el método comprende además: i. depositar los productos de RCA sobre un soporte plano; y ii. contar el número de los productos de RCA marcados individuales en una zona del soporte.
15. Método según la reivindicación 14, en el que el soporte es una membrana capilar transparente porosa.
- 15 16. Método según cualquiera de las reivindicaciones 11-15, en el que las diferentes secuencias de B y sus secuencias B' complementarias identifican diferentes cromosomas, y el método comprende además comparar el número de moléculas de ADN producto que comprenden una primera secuencia de o bien B o bien B' con el número de moléculas de ADN producto que comprenden una segunda secuencia de o bien B o bien B' para determinar si la muestra genómica tiene una aneuploidía.
- 20 17. Método según cualquiera de las reivindicaciones 11-16, en el que el método comprende comparar los resultados de recuento de la etapa (c) con los resultados de recuento obtenidos a partir de una o más muestras de referencia.
- 25 18. Método según cualquiera de las reivindicaciones 11-17, en el que la muestra genómica de prueba es de una paciente que se sospecha que padece o que está en riesgo de padecer una enfermedad o estado, y los resultados de recuento de la etapa (c) proporcionan una indicación de si la paciente, o el feto de la misma, padece la enfermedad o estado.



En la cadena superior:

- la secuencia A es un fragmento de ADN genómico;
- la secuencia B identifica el locus, por ejemplo, el cromosoma del que se deriva el fragmento adyacente del ADN genómico;
- las secuencias X y Z son secuencias de sonda.

FIG. 1

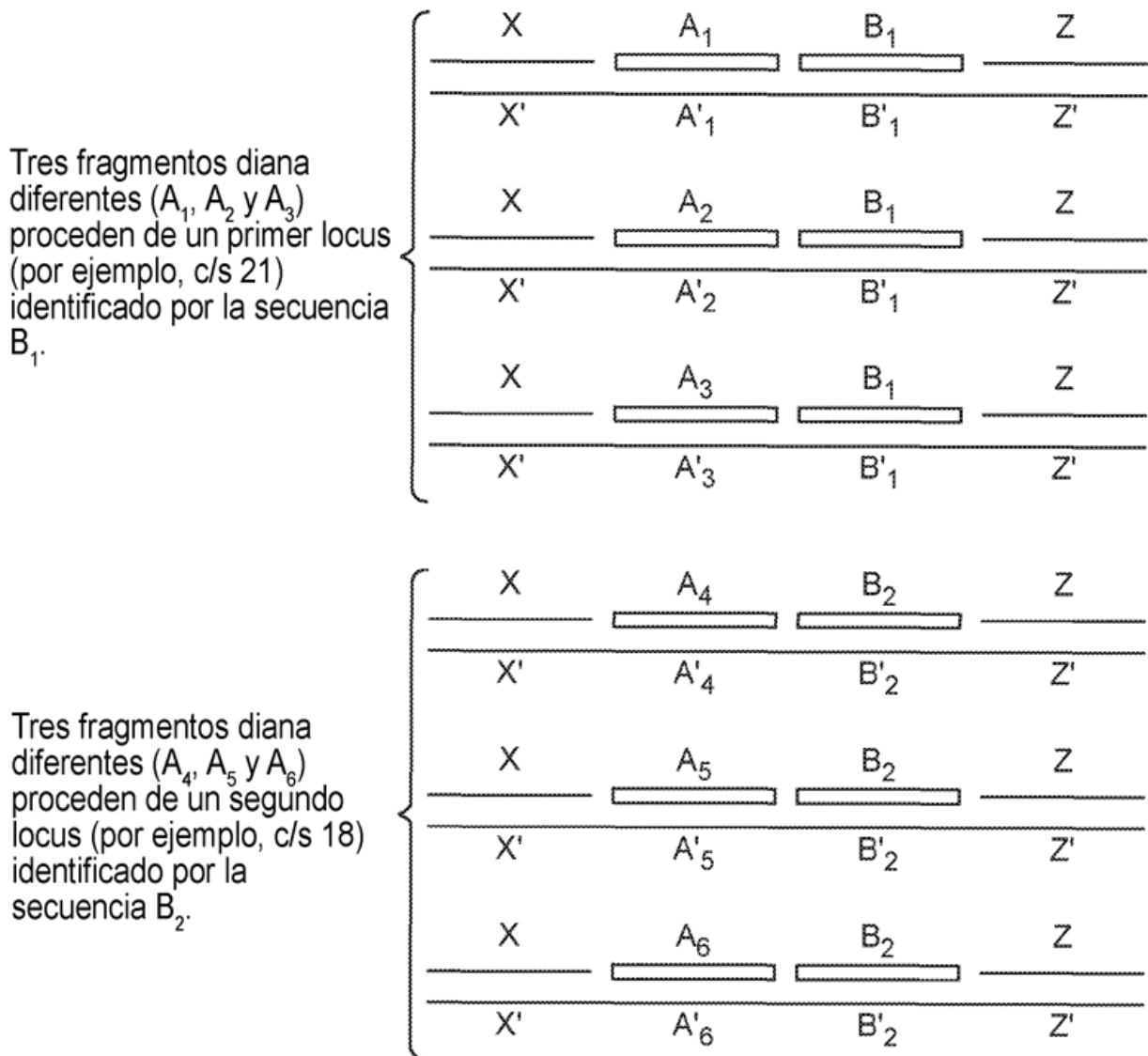


FIG. 2

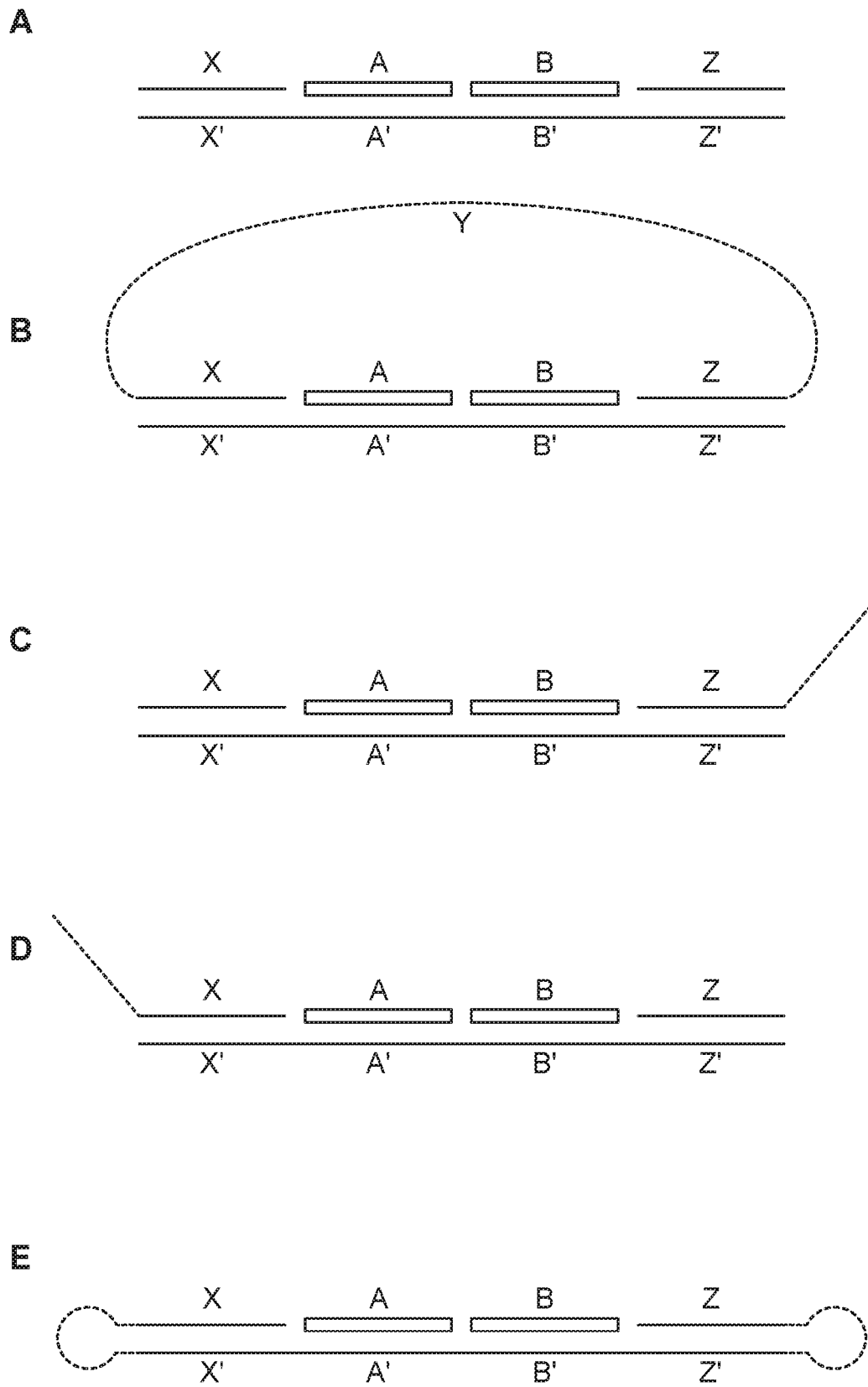


FIG. 3

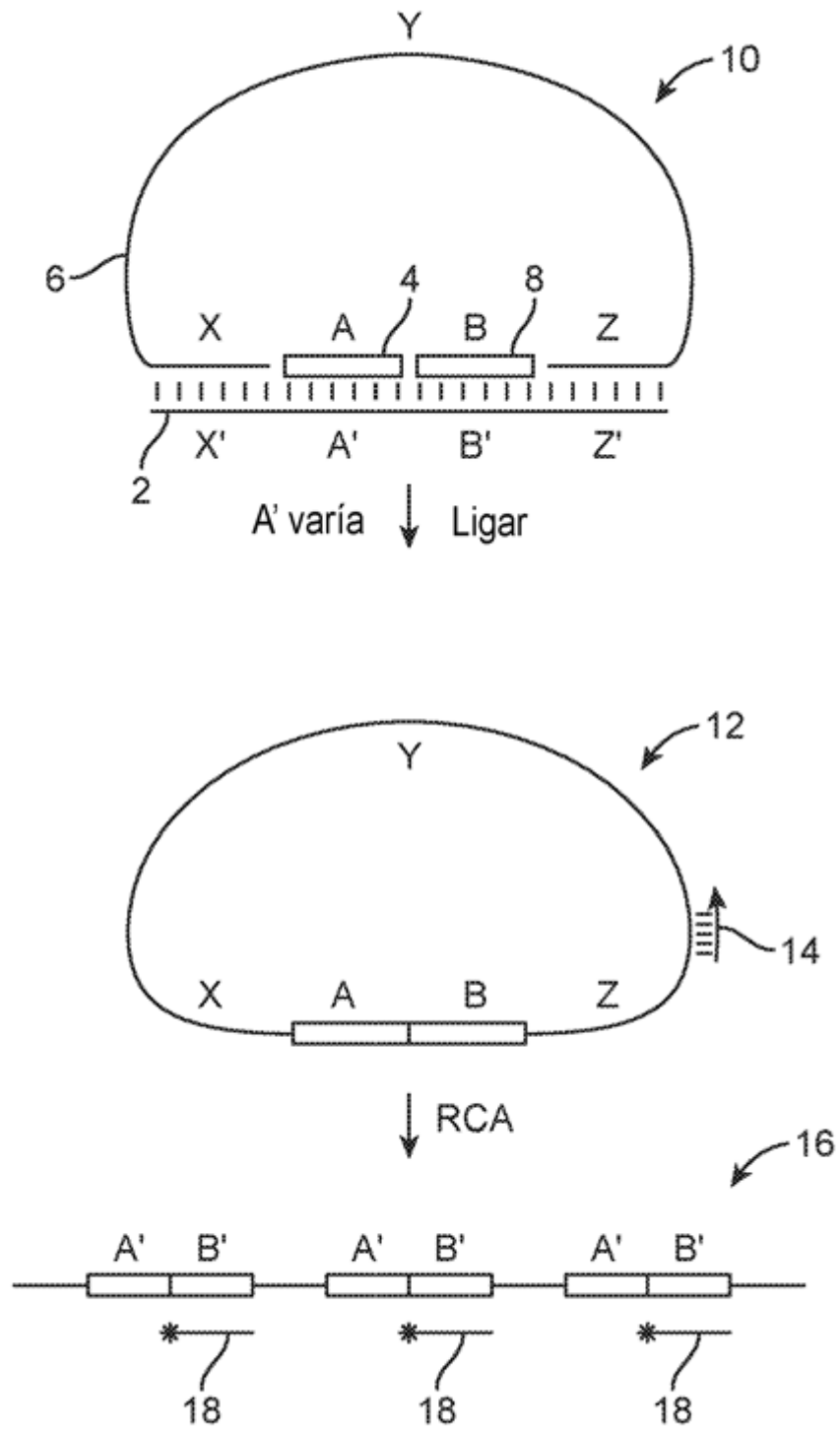


FIG. 4

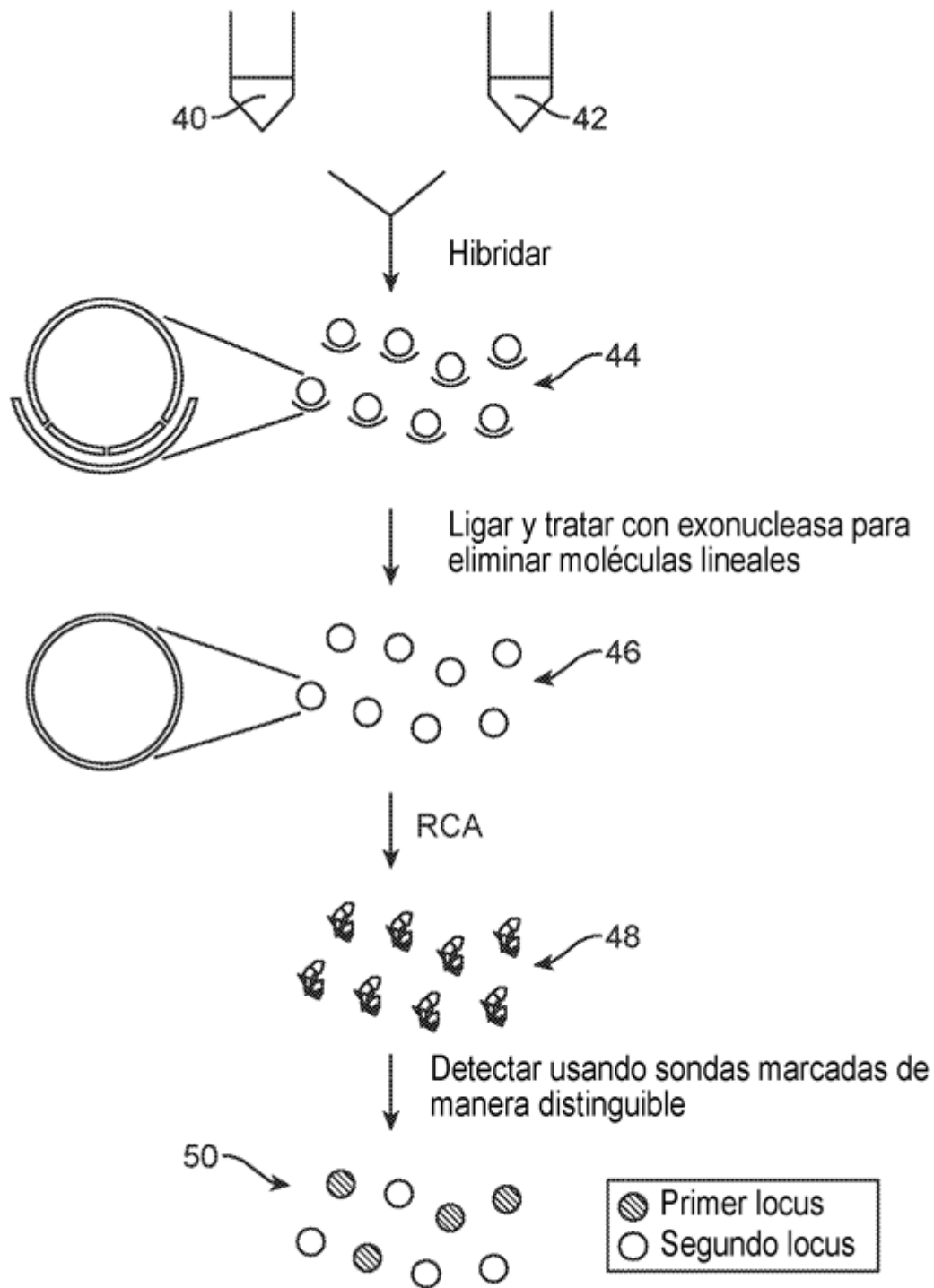


FIG. 5

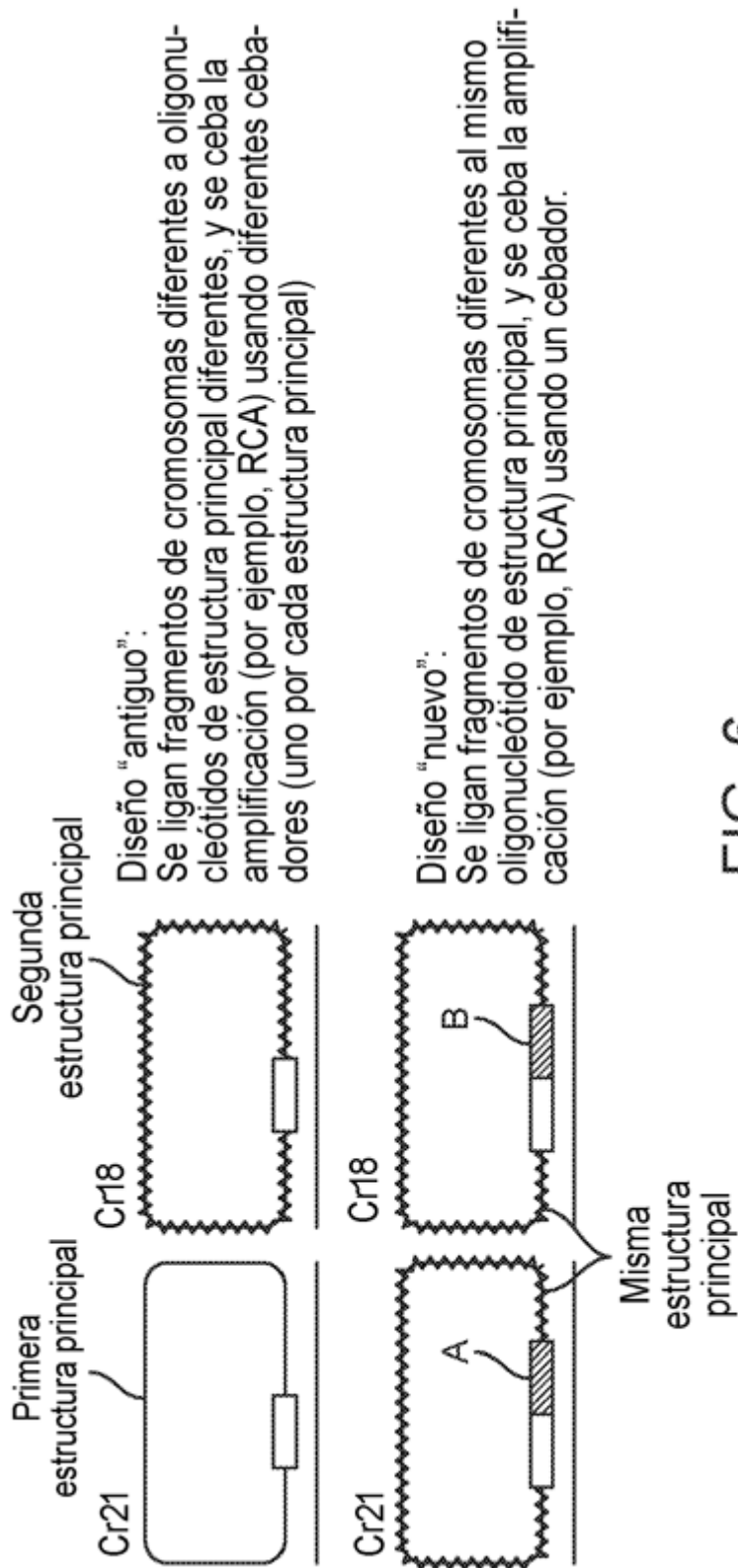
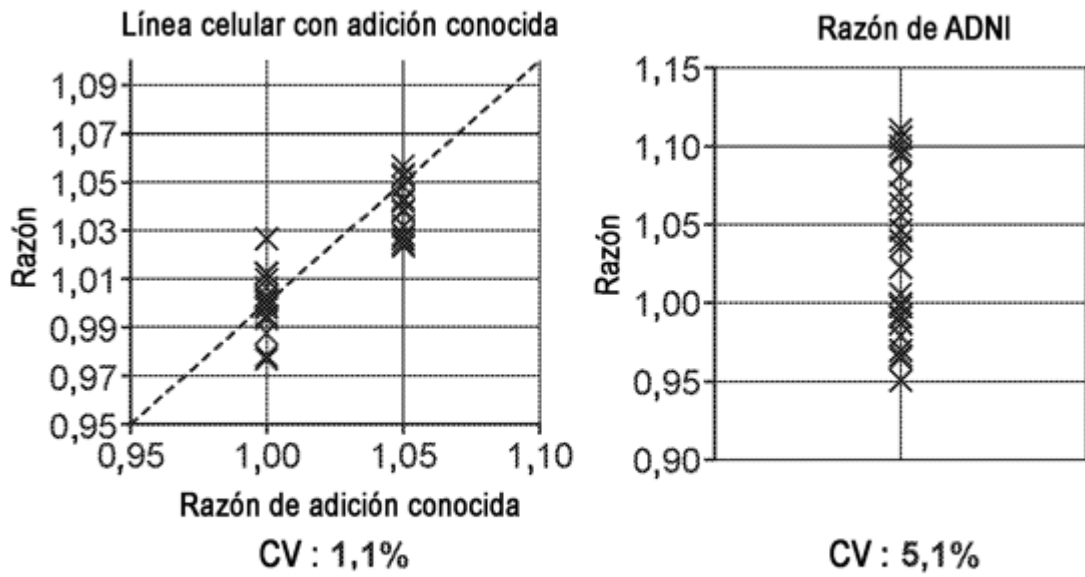


FIG. 6

Diseño "antiguo"



Diseño "nuevo"

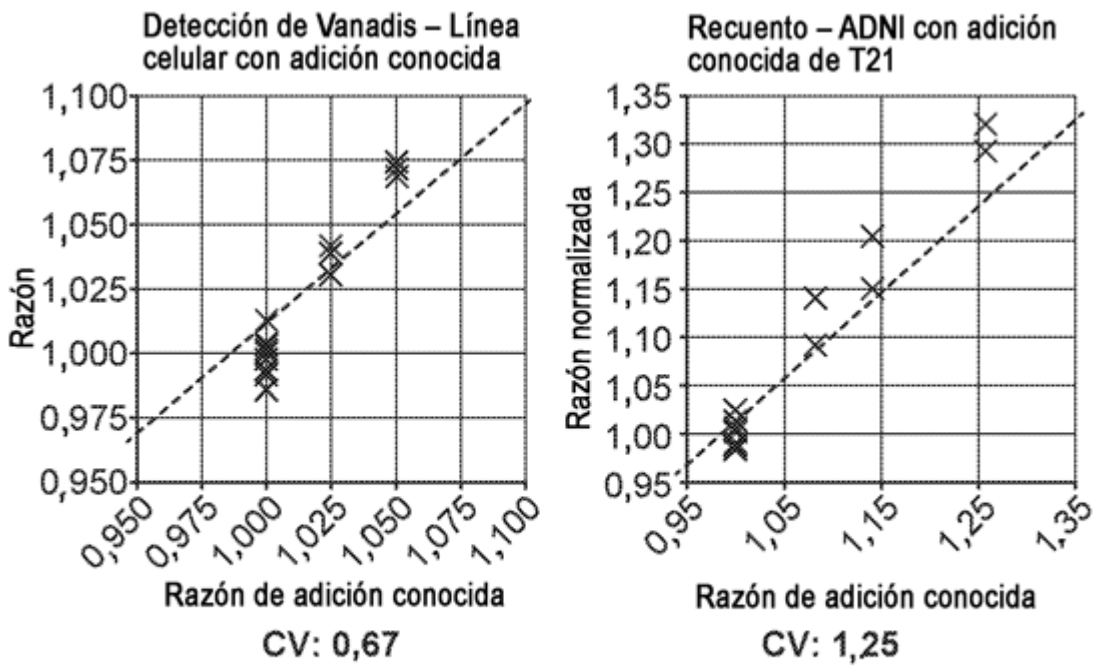


FIG. 7

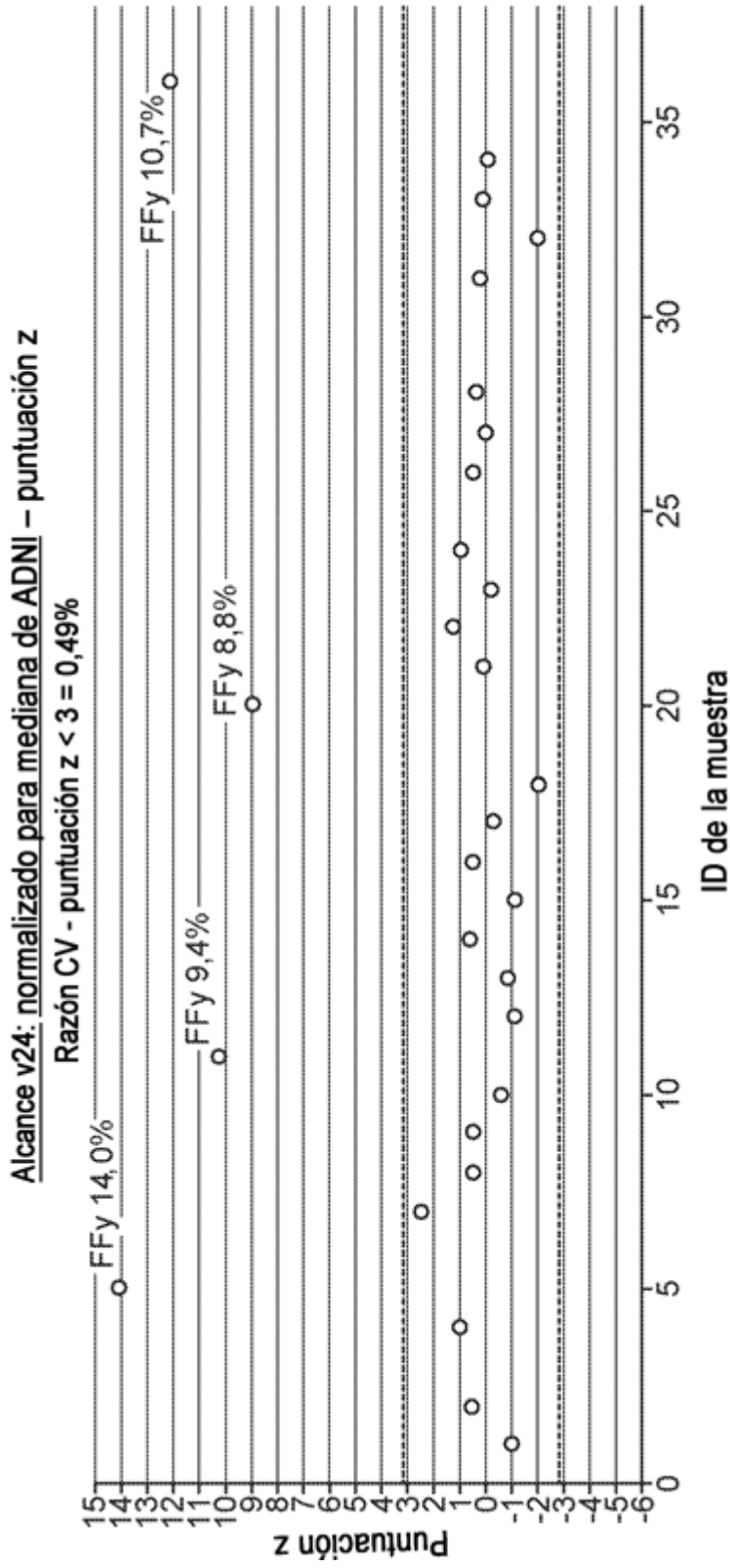


FIG. 8