

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **3 013 064**

51 Int. Cl.:

C12Q 1/6869 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **25.08.2023** E 23193551 (1)

97 Fecha y número de publicación de la concesión europea: **29.01.2025** EP 4332238

54 Título: **Métodos para la detección y cuantificación precisas y paralelas de ácidos nucleicos**

30 Prioridad:

31.08.2022 US 202217899851

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

10.04.2025

73 Titular/es:

**GENOMILL HEALTH OY (100.00%)
Itäinen Pitkätatu 4 B
20520 Turku, FI**

72 Inventor/es:

**PURSIHEIMO, JUHA-PEKKA;
HIRVONEN, TATU;
KORKIAKOSKI, ANTONI y
TAMMINEN, MANU**

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

Observaciones:

Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes

ES 3 013 064 T3

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

DESCRIPCIÓN

Métodos para la detección y cuantificación precisas y paralelas de ácidos nucleicos

5 **Campo técnico**

La presente descripción de la invención se refiere a métodos de secuenciación de ADN de última generación mejorados para la cuantificación precisa y masivamente paralela de una o más dianas de ácido nucleico. Más particularmente, la descripción se refiere a métodos y kits que comprenden sondas para detectar y cuantificar dianas genéticas en conjuntos de ADN complejos usadas principalmente para la detección de dianas y variantes genéticas. La invención usa una o más sondas de ácido nucleico específicas de diana por diana genética, un oligo de bucle de código de barras y uno o más oligos de puente.

15 **Antecedentes**

Con el avance en la tecnología del estudio de la variación genética, la detección de la misma en plantas y animales no es engorrosa. Sin embargo, la detección y cuantificación precisa de variaciones genéticas tales como mutaciones, en particular en muestras que tienen señales débiles es actualmente aún engorrosa, laboriosa y costosa, a pesar de los costes de secuenciación disminuidos. Diversos problemas pueden expresarse de forma más precisa tales como la especificidad para detectar señales genéticas contra un fondo consenso, sensibilidad para detectar señales genéticas débiles, precisión para la cuantificación precisa de las señales detectadas, número de rendimiento de dianas genéticas dirigidas por ensayo, coste por ensayo, escalado para determinar la escala del coste del ensayo cuando se ensayan múltiples muestras en paralelo y rotación para determinar cuánto es el tiempo desde el muestreo a los resultados.

25 Actualmente, los métodos de cuantificación habituales para biopsias líquidas y ensayos conceptualmente similares (tales como detección de genes de resistencia a los antibióticos), incluyen PCR cuantitativa (qPCR), qPCR de matriz, PCR digital, amplificación de sonda dependiente de ligamiento múltiple (MLPA) o cuantificación a partir de datos de secuenciación de ADN de nueva generación. Aunque los métodos de cuantificación son métodos robustos y bien establecidos, cada uno de los métodos está asociado a problemas específicos analizados con mayor detalle a continuación:

30 PCR cuantitativa: La PCR cuantitativa (qPCR), es una técnica que incluye la amplificación de una molécula de ADN marcada como diana durante la PCR, es decir, en tiempo real. La PCR en tiempo real puede usarse cuantitativamente (PCR en tiempo real cuantitativa) y semi-cuantitativamente, es decir, por encima/por debajo de una determinada cantidad de moléculas de ADN (PCR en tiempo real semi cuantitativa). La PCR cuantitativa (qPCR) es un método de referencia de la cuantificación de dianas genéticas. Actualmente, el coste de laboratorio de una reacción de qPCR es aproximadamente 2 \$. Sin embargo, teniendo en cuenta la mano de obra considerable en tiempo (coste laboral) para configurar la reacción, la necesidad de curvas patrón, junto con réplicas para cada diana cuantificada, el coste real es de hecho mucho más alto. La cantidad de mano de obra en el tiempo escala abruptamente con un número en aumento de muestras ya que se requiere un experimento de cuantificación distinto para cada diana genética.

40 PCR de matriz: Las matrices de PCR son las herramientas más fiables para analizar la expresión de un panel enfocado de genes relevantes, de rutas o de enfermedades. Cada placa de 96 pocillos, placa de 384 pocillos o disco de 100 pocillos de Matriz de PCR incluye ensayos de cebador optimizados SYBR Green para un panel exhaustivamente investigado de paneles enfocados de genes. Una nueva iteración de la tecnología qPCR es una qPCR de matriz que miniaturiza las reacciones de qPCR individuales. La PCR de matriz disminuye el coste de una reacción de qPCR individual y mejora la escalabilidad del método a múltiples dianas y muestras. Sin embargo, el método está actualmente limitado a perfilar 384 dianas de 12 muestras (o a la inversa 12 dianas de 384 muestras) a un coste de miles de dólares por chip más un gran coste de capital de la infraestructura de lectura. Perfilar miles de muestras usando el ajuste anteriormente mencionado, por lo tanto, se mantiene prohibitivamente caro.

50 PCR digital: La reacción en cadena de la polimerasa digital (PCR digital, PCR Digital, dPCR o dePCR) es un método para proporcionar la cuantificación absoluta de dianas a través de microfluidos de gotitas y detección de fluorescencia. La metodología es relativamente rentable (una diana por muestra cuesta alrededor de 3 \$) pero la mano de obra para preparar, ajustar y ejecutar experimentos individuales para cada diana en cada muestra escala mal a miles de muestras.

55 La amplificación de sonda dependiente de ligamiento múltiple (MLPA) proporciona un enfoque para simplificar la detección de múltiples dianas genéticas en muestras individuales. Sin embargo, la MLPA proporciona solo la cuantificación relativa de dianas y requiere un experimento de detección distinto para cada muestra. Más recientemente, una variante de MLPA introduce conceptos de códigos de barras de ADN. El concepto permite una mejor resolución cuantitativa y multiplexación de muestras que el flujo de trabajo de MLPA tradicional.

60 Enfoques basados en secuenciación de última generación: La secuenciación de última generación (NGS), también conocida como secuenciación de alto rendimiento que hace al análisis de expresión génica basada en secuencia una alternativa “digital” a técnicas analógicas. La diana a tener en cuenta de los datos de secuenciación de ADN de última generación se está volviendo cada vez más atractiva ya que los costes de secuenciación de ADN se mantienen disminuyendo y se usa actualmente por ejemplo en pruebas prenatales no invasivas. Sin embargo, el enfoque actual

5 padece altos costes de preparación de bibliotecas de secuenciación y esfuerzos de secuenciación que se malgastan en dianas genéticas no relevantes. Por ejemplo, en biopsias líquidas relacionadas con cáncer, los enfoques no dirigidos dan como resultado el malgasto de esfuerzo de secuenciación en locus oncológicamente no relevantes. En diagnósticos fetales, el muestreo no dirigido de locus limita considerablemente las opciones estadísticas para interpretar los datos. Guardant Health Inc proporciona enfoques de secuenciación más dirigidos, donde una matriz de sondas de captura de ARN enriquece dianas para la secuenciación de ADN de última generación.

10 Akhras y col. (2007) PLoS ONE 2(2):e223 describen un ensayo de detección de patógenos en múltiplex que implica sondas específicas de diana de código de barras, circularización y secuenciación de dianas. El uso de un oligonucleótido de puente para ligar las sondas específicas de diana también se describe.

El documento WO2018109206 describe métodos para la detección de analitos en una muestra usando sondas de candado y amplificación por círculo rodante. El uso de un oligo de puente no se describe.

15 El documento WO2019038372 describe un enfoque de secuenciación de nueva generación en donde las secuencias diana de interés se amplifican selectivamente mediante la transcripción in vitro de complejos de ligamiento que contienen un promotor para la polimerasa T7, seguido de síntesis de ADNc y secuenciación. Aunque este método permite la detección y cuantificación precisas y paralelas de muchas secuencias diana en una muestra, las muestras más complejas, de volumen grande, diluidas y/o impuras siguen siendo un reto.

20 Por lo tanto, a la luz del análisis anterior, existe una necesidad de superar las desventajas anteriormente mencionadas tales como, aunque no de forma limitativa, especificidad, sensibilidad, precisión, rendimiento, coste, escalado y tiempo a través de una cuantificación precisa y masivamente paralela de dianas de ácido nucleico.

25 **Resumen de la invención**

La presente invención proporciona un método para usar secuenciación de nueva generación para cuantificación de diana altamente sensible, escalable y precisa a partir de muestras de gran volumen (hasta decenas de mililitros) y/o material de muestra diluido y/o no purificado. Se evita una etapa de amplificación de ARN tal como se describe en el documento WO2019038372, haciendo que el método sea más sencillo. Es más, el método de la presente invención incluye el uso de un oligonucleótido de bucle de código de barras distinto para la identificación de las muestras. El uso de un oligonucleótido de bucle de código de barras permite la consulta simultánea de dianas y la indexación de muestras, lo que permite una agrupación eficiente de las muestras y, en consecuencia, un ensayo de cuantificación más rentable y flexible y proporciona una forma alternativa de introducir secuencias de identificadores moleculares únicos (UMI). El protocolo resultante omite el requisito de indexación de muestras por separado, lo que permite longitudes de lectura más cortas y aporta beneficios tanto a las aplicaciones de lectura corta como a las de secuenciación por nanoporos.

40 En un primer aspecto principal, la invención se refiere a un método para la detección de alto rendimiento de una o más secuencia de nucleótidos diana en una pluralidad de muestras, comprendiendo el método las etapas de:

(i) proporcionar para cada secuencia de nucleótidos diana en cada una de las muestras: una primera sonda, una segunda sonda, un oligo de bucle de código de barras y un oligo de puente o una pluralidad de oligonucleótidos de puente capaces de hibridarse con el oligo de bucle de código de barras para formar un complejo de oligo de puente,

45 en donde la primera sonda comprende una primera secuencia específica de oligo de puente en el extremo 5' de la primera sonda y una primera porción específica de diana en el extremo 3' de la primera sonda;

50 en donde la segunda sonda comprende una segunda porción específica de diana en el extremo 5' de la segunda sonda y una segunda secuencia específica de oligo de puente en el extremo 3' de la segunda sonda;

55 en donde el oligo de bucle de código de barras comprende, a partir del extremo 5' de la molécula, una tercera secuencia específica de oligo de puente, una secuencia de bucle de código de barras y una cuarta secuencia específica de oligo de puente,

en donde el oligo de puente o la pluralidad de oligonucleótidos de puente contiene secuencias complementarias a la primera secuencia específica de oligo de puente y a la segunda secuencia específica de oligo de puente en la primera sonda y la segunda sonda, respectivamente, y secuencias complementarias a la tercera secuencia específica de oligo de puente y a la cuarta secuencia específica de oligo de puente en el oligo de bucle de código de barras;

60 y en donde, opcionalmente, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende una secuencia de reconocimiento para una endonucleasa;

65 (ii) poner en contacto, para cada una de las una o más secuencias de nucleótidos diana, la primera sonda y la segunda sonda con, preferiblemente para cada una de las muestras en un tubo distinto, el oligo de bucle de código de barras y el oligo de puente o la pluralidad de oligonucleótidos de puente y permitir autohibridarse en complejos de ligamiento;

- (iii) poner en contacto los ácidos nucleicos presentes en cada una de las muestras que va a someterse a prueba para las secuencias de nucleótidos diana, con los complejos de ligamiento;
- 5 (iv) permitir que la primera porción específica de diana y la segunda porción específica de diana de la respectiva primera sonda y segunda sonda se hibriden con secciones esencialmente adyacentes en la secuencia diana, formando de este modo complejos de hibridación;
- 10 (v) agrupar opcionalmente los complejos de hibridación de la pluralidad de muestras;
- (vi) ligar las sondas en los complejos de hibridación para proporcionar complejos de ligamiento ligados;
- 15 (vii) amplificar ácidos nucleicos de los uno o más complejos de ligamiento ligados usando amplificación por círculo rodante con una polimerasa de desplazamiento de hebra, obteniendo de este modo secuencias concatémicas monocatenarias;
- (viii) opcionalmente, siempre que esté presente una secuencia de reconocimiento como la especificada en la etapa (i), realizar una etapa para obtener fragmentos de ácido nucleico mediante:
- 20 (a) escindir las secuencias concatémicas monocatenarias obtenidas en la etapa (vii), o
- (b) someter las secuencias concatémicas monocatenarias obtenidas en la etapa (vii) para hibridarse con un oligonucleótido específico que contiene una secuencia de reconocimiento para una endonucleasa en donde el oligonucleótido se hibrida con la secuencia de reconocimiento especificada en la etapa (i) de modo que se obtiene un sitio de reconocimiento para la endonucleasa y escindir los complejos hibridados con dicha endonucleasa;
- 25 (ix) someter la secuencia concatémica obtenida en la etapa (vii) o los fragmentos de ácido nucleico obtenidos en la etapa (viii) a una tecnología de secuenciación de alto rendimiento para determinar la secuencia o secuencias de códigos de barras; y
- 30 (x) identificar la presencia y/o el número de la secuencia de nucleótidos diana en la pluralidad de muestras mediante la determinación de al menos parte de la primera porción específica de diana y/o la segunda porción específica de diana, y/o al menos parte de la secuencia de código de barras correspondiente al código de barras en el oligo de bucle de código de barras, en donde las etapas (v) y (vi) pueden realizarse en cualquier orden.
- 35

Descripción de las figuras

La figura 1 ilustra un diagrama de flujo del ensayo de ligamiento multiplexado (MLA, *Multiplexed Ligation Assay*) según una realización de la invención.

Las figuras 2A, 2B, 2C y 2D ilustran las principales combinaciones de sondas según las realizaciones de la invención.

La figura 3 ilustra los productos de RCA del flujo de trabajo antes (carril 2) y después (carril 1) de la digestión mediante una endonucleasa de restricción.

La figura 4 ilustra la respuesta lineal del flujo de trabajo experimental para disminuir logarítmicamente el número de dianas genéticas en cuatro reacciones replicadas, inferida a partir de datos de secuenciación de ADN de nueva generación enumerando los códigos de barras moleculares. Cada fila representa tres concentraciones de la secuencia diana. La respuesta es lineal en tres órdenes de magnitud.

Descripción detallada de la invención

Definiciones

- 55 Secuencia de nucleótidos diana: La expresión secuencia de nucleótidos diana puede ser cualquier secuencia de nucleótidos de interés de la cual se requiere su detección. Se entenderá que la expresión dada se refiere a una secuencia de nucleótidos contiguos así como a moléculas de ácido nucleico con la secuencia complementaria. La secuencia diana en algunas realizaciones es una secuencia de nucleótidos que representa o está asociada a un polimorfismo.
- 60 Polimorfismo: El término polimorfismo se refiere a una aparición de dos o más secuencias o alelos alternativos determinados genéticamente en una población. Un marcador o sitio polimórfico es el locus en el cual se produce la divergencia de secuencia. Un locus polimórfico puede ser tan pequeño como un par de bases.
- 65 Muestras: El término muestras se usa en la presente memoria para dos o más muestras que contienen dos o más secuencias diana. Las muestras como se proporcionan en un método según la invención pueden haberse preparado para extraer al menos los ácidos nucleicos diana y hacer estos accesibles a las sondas como se usan en la invención.

- 5 En particular, en algunas realizaciones, las muestras comprenden cada una al menos dos secuencias diana diferentes, preferiblemente al menos 100, más preferiblemente al menos 250, más preferiblemente al menos 500, con máxima preferencia al menos 2000, o más. El término muestras puede referirse a pero no se limita a dos o más muestras obtenidas de un cuerpo humano/animal, incluidas orina, biopsias, saliva y otras secreciones, extractos de humedad exhalada, tejido, plasma sanguíneo (biopsias líquidas), o dos o más muestras obtenidas del ambiente, incluidos agua, agua residual, suelo, plantas, o dos o más muestras que contienen virus o bacterias o similares. En una realización, la pluralidad de muestras incluye una muestra de sangre, una muestra de saliva, una muestra de orina o una muestra de heces, una muestra de otro fluido corporal o un extracto de material corporal, por ejemplo escamas de pelo o de piel.
- 10 Sonda: El término sonda es un fragmento de ADN o ARN de longitud variable (habitualmente 50-1000 bases de longitud, preferiblemente 50 - 200 bases de longitud) que puede usarse en muestras de ADN o ARN para detectar la presencia de secuencias de nucleótidos (la diana de ADN o ARN) que son complementarias a la secuencia en la sonda. Las secciones de las sondas de oligonucleótidos que son complementarias a la secuencia diana se diseñan de tal manera que para cada secuencia diana en una muestra, se proporciona un par de una sonda primera y segunda, por lo que cada una de las sondas contiene una sección en su extremo final que es complementaria a una parte de la secuencia diana. Adicionalmente, la presente descripción describe un oligo de puente o complejo de oligo de puente que se usa para unir la primera sonda y la segunda sonda. Por otra parte, la presente descripción describe un oligo de bucle de código de barras que comprende una sección de bucle flanqueada por dos secciones que pueden hibridarse con los uno o más oligos de puente. La sección de bucle no se hibrida con los uno o más oligos de puente y comprende un código de barras.
- 15 20 Universal: El término universal cuando se usa para describir un procedimiento de amplificación se refiere a una secuencia que permite el uso de un único cebador o un conjunto de cebadores para una pluralidad de reacciones de amplificación. El uso de tales cebadores simplifica enormemente la multiplexación en que solo son necesarios dos cebadores para amplificar una pluralidad de secuencias de ácido nucleico seleccionadas. El término universal cuando se usa para describir un sitio de cebado es un sitio al cual hibridará un cebador universal. También debe indicarse que pueden usarse “conjuntos” de secuencias cebadoras/cebadores universales.
- 25 Hibridación: El término hibridación describe el proceso de que las moléculas de ADN o ARN hibriden a ADN o ARN complementario. La replicación de ADN o ARN y la transcripción de ADN en ARN dependen ambas de la hibridación de nucleótidos.
- 30 Ligamiento: El término ligamiento es la unión de dos fragmentos de ácido nucleico a través de la acción de una enzima. Las ADN ligasas son enzimas capaces de catalizar la formación de un enlace fosfodiéster entre (los extremos de) dos hebras polinucleotídicas unidas en sitios adyacentes en una hebra complementaria. En una realización, el ligamiento también puede realizarse químicamente, en particular, si ambos extremos adyacentes de los polinucleótidos se modifican para permitir el ligamiento químico.
- 35 Amplificación: El término amplificación como se usa en la presente memoria denota el uso de una ADN polimerasa para aumentar la concentración de una secuencia de nucleótido particular en una mezcla de secuencias de nucleótidos. La “PCR” o “Reacción en Cadena de la Polimerasa” es un procedimiento rápido para la amplificación enzimática *in vitro* de un segmento de ADN/ARN específico. El ADN/ARN a amplificar puede desnaturalizarse calentando la muestra. El término cebador es una hebra de ARN o ADN (generalmente de aproximadamente 18-22 bases) que sirve como un punto de partida para la síntesis de ADN. Se requiere para la replicación de ADN porque las enzimas que catalizan este proceso, ADN polimerasas, solo pueden añadir nuevos nucleótidos a una hebra existente de ADN.
- 40 45 Polimerasa: Una polimerasa es una enzima que sintetiza cadenas largas o polímeros de ácidos nucleicos. La ADN polimerasa y la ARN polimerasa se usan para ensamblar moléculas de ADN y ARN, respectivamente, copiando una hebra molde de ADN o ARN usando interacciones de apareamiento de bases.
- 50 Alto rendimiento: La expresión alto rendimiento denota la capacidad de procesar y cribar simultáneamente un gran número de muestras de ADN; así como cribar simultáneamente grandes números de diferentes locus genéticos dentro de una única muestra de ADN. La secuenciación o cribado de alto rendimiento (High-Throughput Sequencing or Screening), a menudo abreviado HTS es un método para la experimentación científica específicamente relevante para cribar eficazmente grandes cantidades de muestras simultáneamente.
- 55 Endonucleasa: Una endonucleasa es una enzima que escinde o corta la hebra doble o simple del ADN en un lugar específico o aleatorio.
- 60 Código de barras: Las sondas y los oligos usados en la presente invención pueden comprender uno o más códigos de barras que consisten en secuencias de nucleótidos. Las secuencias de códigos de barras pueden comprender secuencias identificadoras de secuencias de nucleótidos diana, secuencias identificadoras de muestras y/o códigos de barras moleculares (también denominados identificadores moleculares únicos) para la enumeración de dianas. Las secuencias de códigos de barras pueden comprender secuencias aleatorias.
- 65 Tal como se describió anteriormente, la descripción se refiere más particularmente a un método para la detección de alto rendimiento de detección de secuencias de nucleótidos diana en un número muy grande de muestras aprovechando los

ensayos dependientes de ligamiento. La descripción proporciona un método para determinar las secuencias de dianas genéticas en conjuntos de ácido nucleico complejos usando técnicas permitidas por la secuenciación de última generación. La descripción también proporciona un método para perfilar múltiples dianas genéticas en varias muestras, preferiblemente un número muy grande de muestras, aprovechando los ensayos dependientes de ligamiento. La descripción proporciona un método para la amplificación de sonda dependiente de ligamiento múltiple que permite consultar diferentes ácidos nucleicos diana en una pluralidad de muestras. Los métodos de la presente invención permiten la secuenciación de las una o más secuencias de nucleótidos diana en una pluralidad de muestras proporcionando una pluralidad de diferentes conjuntos de sondas para diferentes ácidos nucleicos diana. Se usan identificadores de secuencia única para la identificación de las dianas genéticas y la cuantificación absoluta de muestras individuales a partir del conjunto de muestras cuando se procesan los datos de secuenciación.

En un primer aspecto principal, la invención se refiere a un método para la detección de alto rendimiento de una o más secuencia de nucleótidos diana en una pluralidad de muestras, comprendiendo el método las etapas de:

- (i) proporcionar para cada secuencia de nucleótidos diana en cada una de las muestras: una primera sonda, una segunda sonda, un oligo de bucle de código de barras y un oligo de puente o una pluralidad de oligonucleótidos de puente capaces de hibridarse con el oligo de bucle de código de barras para formar un complejo de oligo de puente,
 - en donde la primera sonda comprende una primera secuencia específica de oligo de puente en el extremo 5' de la primera sonda y una primera porción específica de diana en el extremo 3' de la primera sonda;
 - en donde la segunda sonda comprende una segunda porción específica de diana en el extremo 5' de la segunda sonda y una segunda secuencia específica de oligo de puente en el extremo 3' de la segunda sonda;
 - en donde el oligo de bucle de código de barras comprende, a partir del extremo 5' de la molécula, una tercera secuencia específica de oligo de puente, una secuencia de bucle de código de barras y una cuarta secuencia específica de oligo de puente,
 - en donde el oligo de puente o la pluralidad de oligonucleótidos de puente contiene secuencias complementarias a la primera secuencia específica de oligo de puente y a la segunda secuencia específica de oligo de puente en la primera sonda y la segunda sonda, respectivamente, y secuencias complementarias a la tercera secuencia específica de oligo de puente y a la cuarta secuencia específica de oligo de puente en el oligo de bucle de código de barras;
 - y en donde, opcionalmente, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende una secuencia de reconocimiento para una endonucleasa;
- (ii) poner en contacto, para cada una de las una o más secuencias de nucleótidos diana, la primera sonda y la segunda sonda con, preferiblemente para cada una de las muestras en un tubo distinto, el oligo de bucle de código de barras y el oligo de puente o la pluralidad de oligonucleótidos de puente y permitir autohibridarse en complejos de ligamiento;
- (iii) poner en contacto los ácidos nucleicos presentes en cada una de las muestras que va a someterse a prueba para las secuencias de nucleótidos diana, con los complejos de ligamiento;
- (iv) permitir que la primera porción específica de diana y la segunda porción específica de diana de la respectiva primera sonda y la segunda sonda se hibriden con secciones esencialmente adyacentes en la secuencia diana, formando de este modo complejos de hibridación (u opcionalmente un complejo de hibridación);
- (v) agrupar opcionalmente los complejos de hibridación de la pluralidad de muestras;
- (vi) ligar las sondas en los complejos de hibridación para proporcionar complejos de ligamiento ligados;
- (vii) amplificar ácidos nucleicos de los uno o más complejos de ligamiento ligados usando amplificación por círculo rodante con una polimerasa de desplazamiento de hebra, obteniendo de este modo secuencias concatémicas monocatenarias;
- (viii) opcionalmente, siempre que esté presente una secuencia de reconocimiento como la especificada en la etapa (i), realizar una etapa para obtener fragmentos de ácido nucleico mediante:
 - (a) escindir las secuencias concatémicas monocatenarias obtenidas en la etapa (vii), o
 - (b) someter las secuencias concatémicas monocatenarias obtenidas en la etapa (vii) para hibridarse con un oligonucleótido específico que contiene una secuencia de reconocimiento para una endonucleasa en donde el oligonucleótido se hibrida con la secuencia de reconocimiento especificada en la etapa (i) de modo que se obtiene un sitio de reconocimiento para la endonucleasa y escindir los complejos hibridados con dicha endonucleasa;

(ix) someter la secuencia concatémica obtenida en la etapa (vii) o los fragmentos de ácido nucleico obtenidos en la etapa (viii) a una tecnología de secuenciación de alto rendimiento para determinar la secuencia o secuencias de códigos de barras; y (x) identificar la presencia y/o el número de la secuencia de nucleótidos diana en la pluralidad de muestras mediante la determinación de al menos parte de la primera porción específica de diana y/o la segunda porción específica de diana, y/o al menos parte de la secuencia de códigos de barras correspondiente al código de barras en el oligo de bucle de código de barras,

en donde las etapas (v) y (vi) pueden realizarse en cualquier orden.

La figura 1 proporciona una ilustración no limitativa de una realización del método de la invención.

Los métodos de la presente invención utilizan cuatro o más moléculas de ácido nucleico, de las cuales dos sondas de ácido nucleico específicas de diana (primera sonda y segunda sonda) son específicas para una diana genética y otras dos o más sondas de ácido nucleico típicamente son universales (oligo de puente o complejo de oligo de puente y oligo de bucle de código de barras). La primera sonda, la segunda sonda y el oligo de bucle de código de barras se hibridan con las una o más sondas de puente formando un complejo de ligamiento. Los complejos de ligamiento (que contienen una o más secuencias de códigos de barra) que tienen sitios de identificación diana en la muestra de ADN o ARN se permiten hibridar contra secuencias diana complementarias de la muestra de consulta. Después de la hibridación, la sonda primera y segunda se ligan química o enzimáticamente mediante una ADN ligasa para formar el complejo de ligamiento ligado. En la presente invención, se formará una pluralidad de dichos complejos de ligamiento ligados durante el análisis de muestras en la pluralidad de muestras a analizar.

Una "pluralidad de muestras" puede referirse a, pero no se limita a, dos o más muestras obtenidas de un cuerpo humano o animal, incluidas biopsias, saliva y otras secreciones, extractos de humedad exhalada, tejido, plasma sanguíneo (biopsias líquidas), dos o más muestras obtenidas del ambiente, incluidos agua, agua residual, suelo, plantas, o dos o más muestras que contienen virus o bacterias o similares. En una realización, la muestra se usa sin ninguna purificación o concentración previa de ácido nucleico. En otra realización, la muestra puede tratarse previamente, por ejemplo, lisando células para exponer ácido nucleico.

La secuencia de nucleótidos diana puede incluir cualquier secuencia de nucleótidos de interés de la cual se requiere la detección. La secuencia de nucleótidos diana de la descripción puede obtenerse a partir de, pero no se limita a, una fracción de ADN en la sangre del paciente o una fracción de ADN en sangre materna. Una fracción del ADN en la sangre del paciente puede obtenerse por ejemplo a partir de células de cáncer apoptóticas/necróticas o una fracción de ADN en sangre materna de origen fetal y/o materno. Además, los resultados del análisis se usan para, por ejemplo, evaluar el riesgo de un individuo a un tipo dado de cáncer, para determinar la eficacia de un tratamiento dado contra un cáncer dado, el desarrollo de una mutación relacionada con resistencia a fármacos en un tumor o el riesgo de que un feto lleve trastornos genéticos tales como trisomías habituales de síndromes de Down, Patau y Edwards. En determinadas realizaciones, el método comprende proporcionar, para cada secuencia de nucleótidos diana, una pluralidad de diferentes conjuntos de sonda.

Como se usa en la presente memoria, la expresión conjuntos de sonda incluye una primera sonda, una segunda sonda, un oligo de bucle de código de barras y uno o más oligos de puente.

En determinadas realizaciones, la primera sonda incluye, empezando desde el extremo 5' de la molécula, opcionalmente un fosfato 5', una primera secuencia específica de oligo de puente, opcionalmente una primera secuencia universal, opcionalmente una primera secuencia de código de barras y una primera porción específica de diana en su extremo 3'.

En determinadas realizaciones, la segunda sonda incluye, empezando desde el extremo 5' de la molécula, opcionalmente un fosfato 5', una segunda porción específica de diana, opcionalmente una segunda secuencia de código de barras, opcionalmente una segunda secuencia universal y una segunda secuencia específica de oligo de puente en su extremo 3'.

El oligo de puente o la pluralidad de oligos de puente contiene secuencias complementarias a las secuencias específicas de oligo de puente primera y segunda en la sonda primera y segunda, respectivamente, opcionalmente una secuencia universal y secuencias complementarias a la tercera secuencia específica de oligo de puente y la cuarta secuencia específica de oligo de puente en el oligo de bucle de código de barras.

El oligo de bucle de código de barras comprende, a partir del extremo 5' de la molécula, una tercera secuencia específica de oligo de puente, una secuencia de bucle de código de barras y una cuarta secuencia específica de oligo de puente. El código de barras puede usarse para permitir definir de forma única el complejo dentro de todos los complejos de ligamiento en todas las muestras analizadas. El oligo de bucle de código de barras puede tener cualquier longitud adecuada y, por ejemplo, puede tener una longitud de entre 30 y 100 pb, tal como entre 40 y 60 pb de longitud.

Opcionalmente, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras o los uno o más oligos de puente comprende una secuencia de reconocimiento para una endonucleasa. Una secuencia de reconocimiento de endonucleasas permite la escisión de la secuencia concatémica. En una realización, la secuencia de reconocimiento es una secuencia de reconocimiento para una endonucleasa de restricción tal como EcoRI. En otra realización, la secuencia de reconocimiento es una secuencia de reconocimiento para una

endonucleasa de migración tal como I-CeuI. En otra realización, la secuencia de reconocimiento es una secuencia de reconocimiento para un sistema de corte guiado tipo DNAaseI o CRISPR-Cas. En otra realización, la secuencia de reconocimiento es una secuencia de reconocimiento para una endonucleasa de corte.

5 En una realización, el oligo de bucle de código de barras contiene una o más secuencias de reconocimiento para una endonucleasa, tal como una endonucleasa de corte. En otra realización, dicho oligo de bucle de código de barras contiene dos secuencias de reconocimiento que son capaces de hibridarse entre sí de tal modo que se obtiene un sitio de reconocimiento de endonucleasas bicatenarias. En otra realización del presente documento, el método no comprende la etapa (viii) como se ha especificado anteriormente, sino que (en su lugar) comprende, entre las etapas (vii) y (ix), una etapa para permitir que dos secuencias de reconocimiento en el oligo de bucle de código de barras se hibriden y escindir el sitio de reconocimiento de la endonucleasa bicatenaria resultante con una endonucleasa, tal como una endonucleasa de corte, que tiene especificidad para dicho sitio de reconocimiento.

15 Opcionalmente, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras o los uno o más oligos de puente comprende un primer resto de captura. Un primer resto de captura, cuando se usa en la presente memoria, se refiere a un resto, tal como un grupo químico, que permite capturar la sonda, el complejo de ligamiento o el complejo de hibridación por, es decir, unido a, un segundo resto de captura que está ligado a un soporte sólido. Se puede usar cualquier resto de captura adecuado conocido en la técnica para este propósito. Un ejemplo adecuado bien conocido es la captura de moléculas biotiniladas usando perlas magnéticas recubiertas de estreptavidina. Por lo tanto, en una realización, el primer resto de captura es un resto de biotina, que puede interactuar con un resto de estreptavidina o avidina (el segundo resto de captura) ligado a un soporte sólido, tal como una perla magnética. Otras opciones incluyen derivados de biotina tales como doble biotina, destiobiotina o biotina fotoescindible que puede usarse para la conjugación con estreptavidina/avidina. Otras opciones incluyen el uso de grupos tiol y acridita para conjugación de acridita/acrilamida, grupos alquino y azida para química de click y digoxigenina para la conjugación de anticuerpos anti-digoxigenina. Las parejas de conjugación pueden proporcionarse en cualquier superficie sólida, tales como perlas (magnéticas o de otro modo) o soportes sólidos. En consecuencia, en una realización del método de la invención, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras o los uno o más oligos de puente comprende un primer resto de captura, y entre las etapas (iv) y (v) se realiza una etapa intermedia (iv)(a) que comprende poner el complejo o complejos de hibridación en contacto con un soporte sólido que comprende un segundo resto de captura, lo que permite que el primer resto de captura y el segundo resto de captura interactúen de tal modo que los complejos de hibridación se ligan al soporte sólido y separar los complejos de hibridación ligados a un soporte sólido a partir de componentes de las muestras que no están ligados al soporte sólido.

35 La primera porción específica de diana, la segunda porción específica de diana, las primeras secuencias específicas de oligo de puente, las segundas secuencias específicas de oligo de puente, las terceras secuencias específicas de oligo de puente y/o las cuartas secuencias específicas de oligo de puente, opcionalmente contienen independientemente entre sí al menos un nucleótido químicamente modificado para aumentar la unión a la sonda. Las modificaciones químicas que aumentan la unión de sondas incluyen, aunque no de forma limitativa, ácidos ribonucleicos, ácidos péptido nucleicos y ácidos nucleicos bloqueados (p. ej., como se ilustra en la figura 3 del documento WO2019038372).

45 En una realización, la porción de puente de la primera sonda, la segunda sonda, el oligo de puente de código de barras o estos tres, comprende o comprenden bases químicamente modificadas para permitir la unión mejorada a los uno o más oligos de puente. En otra realización, la primera porción específica de diana, la segunda porción específica de diana, las primeras secuencias específicas de oligo de puente, las segundas secuencias específicas de oligo de puente, las terceras secuencias específicas de oligo de puente y/o las cuartas secuencias específicas de oligo de puente, contienen independientemente unas de otras, uno o más nucleótidos modificados químicamente. En determinadas realizaciones, las modificaciones químicas permiten el ligamiento químico de sondas adyacentes.

50 En algunas realizaciones, las sondas mencionadas anteriormente se unen a loci genéticos completamente adyacentes, es decir, secciones adyacentes de la secuencia de nucleótidos diana o hasta 500 pares de bases de separación, por ejemplo, hasta 200 pares de bases de separación, tal como hasta 50 pares de bases de separación, hasta 40 pares de bases de separación, hasta 30 pares de bases de separación, hasta 20 pares de bases de separación, hasta 10 pares de bases de separación o hasta 5 pares de bases de separación.

55 En algunas realizaciones, la primera sonda, la segunda sonda, el oligo de bucle de código de barras o los uno o más oligos de puente pueden incluir secuencias adaptadoras para una plataforma de secuenciación de ADN tal como (pero sin limitarse a) Illumina. Estas secuencias adaptadoras permiten que las bibliotecas de secuenciación resultantes se unan a las partes de detección de los dispositivos de secuenciación tales como las células de flujo Illumina.

60 Es más, en algunas realizaciones, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende:

65 (i) de una a cinco bases protuyentes en 3' (es decir, bases adicionales que no forman doble hélice con la segunda sonda), y/o

(ii) 3' fosfato, y/o

(iii) una o más modificaciones de fosforotioato dentro de tres posiciones desde el extremo 3'.

5 Antes de poner en contacto las sondas con la muestra que comprende las secuencias diana, la primera sonda, la segunda sonda y el oligo de bucle de código de barras se ponen en contacto con el oligo de puente (o una pluralidad de oligonucleótidos capaces de formar un complejo de oligo de puente), preferiblemente para cada una de las muestras en un tubo distinto y se permite la hibridación en complejos de ligamiento (etapa (ii)). El oligo de puente y el bucle de código de barras pueden prehibridarse antes de la hibridación con la primera y segunda sonda o todas las etapas de hibridación pueden realizarse a la vez. En una realización en donde el puente no es un oligo, sino una pluralidad de oligonucleótidos, tal como dos oligonucleótidos, capaces de hibridarse con el oligo de bucle de código de barras para formar un complejo de oligo de puente (ilustrado en la presente memoria en 2B), la pluralidad de oligonucleótidos puede prehibridarse con el oligo de bucle de código de barras antes de hibridarse con la primera y segunda sonda o todas las etapas de hibridación pueden realizarse a la vez o el oligo de bucle de código de barras puede hibridarse con el complejo de sonda prehibridado durante la captura de la diana.

Preferiblemente cada complejo de ligamiento es único para la combinación de la primera secuencia específica de diana, la segunda secuencia específica de diana y una o más secuencias de código de barras. Esto permite la enumeración de las secuencias diana después de la amplificación y el análisis de los resultados.

20 En lo sucesivo, una o más secuencias de nucleótidos diana en la pluralidad de muestras se pone en contacto con la pluralidad de complejos de ligamiento (etapa (iii)). La primera porción específica de diana y la segunda porción específica de diana de la respectiva primera sonda y segunda sonda hibridan con secciones esencialmente adyacentes en la secuencia diana, formando de este modo un complejo de hibridación (etapa (iv)). Como se ha mencionado anteriormente, las secciones esencialmente adyacentes de la secuencia diana pueden estar inmediatamente adyacentes o puede haber un hueco de hasta 500 pares de bases de separación.

30 En algunas realizaciones, la muestra tiene un volumen de más de 100 microlitros, por ejemplo, más de 1 ml. En una realización adicional, la muestra tiene una concentración de ácido nucleico inferior a 5 pmol, tal como inferior a 1 pmol, por ejemplo, inferior a 200 fmol. En una realización, la pluralidad de muestras incluye una o más muestras de sangre, una o más muestras de saliva, una o más muestras de orina o una o más muestras de heces.

35 Posteriormente, en algunas realizaciones, si al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras o los uno o más oligos de puente comprende un primer resto de captura, el/los complejo(s) de hibridación se pone(n) en contacto con un soporte sólido que comprende un segundo resto de captura y el primer resto de captura y el segundo resto de captura se permiten interactuar de tal modo que el complejo o complejos de hibridación se ligan al soporte sólido (etapa opcional (iv)(a)). Después de eso, los complejos de hibridación unidos al soporte sólido se separan de los componentes de las muestras que no están ligados al soporte sólido. Si los soportes sólidos son perlas magnéticas, las perlas pueden inmovilizarse usando un imán y la muestra líquida restante puede retirarse. Opcionalmente, se realiza una etapa de lavado antes de proceder.

45 La etapa (iv)(a) da como resultado una purificación y enriquecimiento para ácido nucleico, lo que permite obtener resultados mejorados en particular para muestras altamente impuras. En una realización, el método de la invención no comprende una etapa de enriquecimiento de ácidos nucleicos antes de la etapa (iv)(a). Por lo tanto, en una realización, el método no contiene antes de la etapa (vi) una etapa en la que los ácidos nucleicos de la muestra original se concentran más de 2 veces, más de 10 veces, o más de 100 veces. En otra realización, el método de la invención no incluye una etapa de purificación posterior al ligamiento en la etapa (vi).

50 Posteriormente, el ligamiento de las sondas en los complejos hibridados formados se lleva a cabo bien enzimática o químicamente para proporcionar complejos de ligamiento ligados (etapa (vi)). Opcionalmente, como parte de la etapa (vi), un hueco entre la primera sonda y la segunda sonda, entre la primera sonda y el oligo de bucle de código de barras y/o entre la segunda sonda y el oligo de bucle de código de barras, si está presente, puede llenarse introduciendo una polimerasa y uno o más nucleótidos. La polimerasa añade nucleótidos (a) complementarios a la secuencia o secuencias de oligo de puente y de este modo rellena los huecos entre la primera sonda, la segunda sonda y el oligo de bucle de código de barras, lo que da como resultado el ligamiento de las sondas y la inclusión del oligo de bucle de código de barras en la hebra complementaria de puente. El oligo de puente o el complejo de oligo de puente se extiende desde el sitio 5' o el sitio 3' complementario a las sondas ligadas de tal manera que la secuencia identificadora de secuencia diana presente en la primera sonda o la segunda sonda se integra en el oligo de puente o el complejo de oligo de puente. Preferiblemente, se usa una polimerasa que no rompe el ADN bicatenario, tal como por ejemplo una polimerasa Taq, para no interferir con el ligamiento de la primera a la segunda sonda cuando están ambas hibridadas a la secuencia diana. En una realización, el oligo de puente, o uno o más oligonucleótidos de la pluralidad de oligonucleótidos de puente, comprende, en una región no complementaria a la primera sonda, la segunda sonda o el oligo de bucle de código de barras, una pluralidad de análogos de bases universales para permitir la incorporación de secuencias aleatorias adecuadas para usar como código de barras molecular para la enumeración de dianas. En tal realización, como parte de la etapa (vi), se realiza una etapa de

llenado de huecos usando polimerasa y nucleótidos con el fin de generar tales secuencias aleatorias. En una realización, la pluralidad de análogos de bases universales es una pluralidad de 5-nitroindoles.

5 Los complejos de ligamiento ligados se agrupan después opcionalmente a partir de una o más muestras diana (etapa (vi)). Las etapas (v) y (vi) pueden realizarse en el orden especificado o alternativamente en orden inverso.

10 A continuación, los ácidos nucleicos se amplifican a partir de uno o más complejos de ligamiento ligados, obteniendo de este modo secuencias concatémicas monocatenarias (etapa (vii)). La amplificación se realiza usando amplificación por círculo rodante con una polimerasa de desplazamiento de hebra, tal como la polimerasa phi29 (UniProtKB-P03680; DPO_BPPH2) o una polimerasa Bst (P52026; DPO1_GEOSE).

15 En una realización, después de la etapa (vi), pero antes de la etapa (vii), se realizan una etapa (a) y una etapa (b), en donde la etapa (a) comprende permitir que los complejos de ligamiento ligados se disocien de la secuencia de nucleótidos diana y la etapa (b) comprende añadir una sonda específica de diana que comprende una secuencia correspondiente a la secuencia de nucleótidos diana, en donde dicha sonda específica de diana es capaz de hibridarse con los complejos de ligamiento ligados y permitir que la sonda específica de diana se hibride con los complejos de ligamiento ligados, formando de este modo moldes de amplificación. En tal realización, dichos moldes de amplificación se amplifican mediante amplificación por círculo rodante con una polimerasa de desplazamiento de hebra en la etapa (vii).

20 Opcionalmente, siempre que esté presente una secuencia de reconocimiento como la especificada en la etapa (i), se realiza una etapa (viii) para obtener fragmentos de ácido nucleico mediante:

(a) escindir las secuencias concatémicas monocatenarias obtenidas en la etapa (vii), o

25 (b) someter las secuencias concatémicas monocatenarias obtenidas en la etapa (vii) para hibridarse con un oligonucleótido específico que contiene una secuencia de reconocimiento para una endonucleasa en donde el oligonucleótido se hibrida con la secuencia de reconocimiento especificada en la etapa (i) de modo que se obtiene un sitio de reconocimiento para la endonucleasa y escindir los complejos hibridados con dicha endonucleasa.

30 Posteriormente, en la etapa (ix), las secuencias concatémicas obtenidas en la etapa (vii) o, si se realizó la etapa (viii), los fragmentos de ácido nucleico obtenidos en la etapa (viii), se someten a una tecnología de secuenciación de alto rendimiento para determinar la secuencia o secuencias de códigos de barras.

35 Opcionalmente, después de la amplificación, los soportes sólidos, si están presentes, se eliminan y el sobrenadante se usa para su posterior procesamiento. Por ejemplo, si los soportes sólidos son partículas magnéticas, estos pueden retirarse usando un imán. En algunas otras realizaciones del método de la invención, la interacción entre el primer resto de captura y el segundo resto de captura se interrumpe inmediatamente después de la etapa (v), después de la etapa (vi) o después de la etapa (vii). Por ejemplo, si el primer resto de captura es biotina y el segundo resto de captura es estreptavidina, la interacción puede interrumpirse añadiendo biotina soluble en exceso. Si la estreptavidina se une a partículas magnéticas, se puede retirar posteriormente usando un imán.

40 Además, opcionalmente, inmediatamente antes de la etapa (ix) se realiza una amplificación por PCR usando cebadores que se unen a partes universales de la primera y segunda sonda, en donde dichos cebadores incluyen opcionalmente secuencias adaptadoras para la posterior secuenciación en la etapa (ix).

45 En otra realización, la secuenciación en la etapa (ix) se realiza usando secuenciación por nanoporos, en donde opcionalmente la secuencia concatémica obtenida en la etapa (vii) se fragmenta usando complejos de transposición. Las técnicas adecuadas para la secuenciación por nanoporos se han revisado en Wang y col. (2021 Nat Biotechnol 39(11): 1348).

50 La identificación de la presencia y/o el número de la secuencia de nucleótidos diana en la pluralidad de muestras puede realizarse mediante la determinación de al menos parte de la primera y/o segunda porción específica de diana, y/o al menos parte de un código de barras mediante tecnología de secuenciación de alto rendimiento (etapas (ix) y (x)), por ejemplo, usando una plataforma de secuenciación de última generación incluyendo, sin limitación, Illumina iSeq, MiSeq, HiSeq, NextSeq o NovaSeq. Preferiblemente, la enumeración de diana genética se permite contando el número de códigos de barra moleculares por diana y por muestra. Las muestras se separan (desconvolucionan) de los datos de secuencia y las dianas de secuencia se cuantifican in silico después de la secuenciación de ADN.

60 Las ventajas de la presente invención incluyen, aunque no de forma limitativa ensayos de cuantificación con bajo coste, alta simplicidad, alta especificidad, alta sensibilidad, alta precisión, alto rendimiento, alta escalabilidad y alta rotación en comparación con tecnologías de secuenciación de ácido nucleico tradicional. Otro aspecto de la presente invención es que los métodos de la presente invención permiten agrupar múltiples muestras debido a la indexación de las muestras en una fase inicial del flujo de trabajo, lo que proporciona mejoras en los costes y velocidad del ensayo. Otro aspecto de la presente invención es que los métodos de la presente invención permitan la cuantificación precisa y masivamente paralela de una pluralidad de dianas de ácido nucleico en múltiples muestras incluidas poblaciones humanas y animales, e incluyendo grandes volúmenes de material de muestra sin purificar. Tal como se menciona, en una realización preferida,

la muestra, tal como una muestra de orina, se usa sin ninguna purificación o concentración previa de ácido nucleico. En otra realización, la muestra puede tratarse previamente, por ejemplo, lisando células para exponer ácido nucleico. Una ventaja particular de la invención es permitir la detección y la amplificación de la secuencia diana de interés usando diseños de sonda únicos, es decir, el triplete de sonda. Las sondas se diseñan con nucleótidos modificados especialmente situados que mejoran la hibridación y la eficiencia de unión. La mejora en las propiedades de unión da lugar a una especificidad, sensibilidad y precisión del ensayo más altas. Los métodos de la presente invención son igualmente aplicables para estudiar variantes genéticas y encuentran aplicación en diagnóstico y pronóstico, incluidas pero no limitadas a genotipar la muestra o muestras para una o más secuencias y/o polimorfismos, tales como SNP y/o indeles, diagnósticos de cáncer o trastornos cromosómicos fetales a partir de la sangre materna. En una realización preferida, para dos o más muestras o para dos o más combinaciones de locus/alelo, se usan secuencias de código de barras para genotipar las muestras para una o más secuencias y/o polimorfismos, tales como SNP y/o indeles.

En otro aspecto, la invención proporciona un kit de partes que comprende una pluralidad de recipientes, en donde al menos un recipiente comprende uno o más conjuntos de la primera sonda y la segunda sonda, al menos un recipiente comprende un oligo de bucle de código de barras y al menos un recipiente comprende uno o más oligos de puente o una pluralidad de oligonucleótidos de puente capaces de formar un complejo de oligo de puente con el oligo de bucle de código de barras, en donde la primera sonda comprende una primera secuencia específica de oligo de puente en el extremo 5' de la primera sonda y una primera porción específica de diana en el extremo 3' de la primera sonda;

en donde la segunda sonda comprende una segunda porción específica de diana en el extremo 5' de la segunda sonda y una segunda secuencia específica de oligo de puente en el extremo 3' de la segunda sonda;

en donde el oligo de bucle de código de barras comprende, a partir del extremo 5' de la molécula, una tercera secuencia específica de oligo de puente, una secuencia de bucle de código de barras y una cuarta secuencia específica de oligo de puente,

en donde el oligo de puente o la pluralidad de oligonucleótidos de puente contiene secuencias complementarias a la primera secuencia específica de oligo de puente y a la segunda secuencia específica de oligo de puente en la primera sonda y la segunda sonda, respectivamente, y secuencias complementarias a la tercera secuencia específica de oligo de puente y a la cuarta secuencia específica de oligo de puente en el oligo de bucle de código de barras;

en donde, opcionalmente, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende una secuencia de reconocimiento para una endonucleasa;

y en donde opcionalmente el kit de partes comprende además un oligonucleótido capaz de hibridarse con dicha secuencia de reconocimiento de tal modo que se obtiene un sitio de reconocimiento para dicha endonucleasa. Preferiblemente, el extremo 3' de las primeras sondas o el extremo 5' de las segundas sondas, o ambos, se modifican para permitir el ligamiento químico de las primeras sondas a las segundas sondas.

Preferiblemente, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende uno o más nucleótidos químicamente modificados en la secuencia complementaria a una secuencia de la primera sonda o en la secuencia complementaria a una secuencia de la segunda sonda o ambas.

Preferiblemente, el extremo 3' de la primera sonda o el extremo 5' de la segunda sonda, o ambos, se modifican para permitir el ligamiento química de la primera sonda a la segunda sonda.

Preferiblemente, la porción de puente de la primera sonda o la segunda sonda, o ambas, o el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprenden bases modificadas químicamente para permitir una mejor unión al oligo de puente o complejo de oligo de puente.

En una realización particular, al menos un recipiente que comprende el conjunto de la sonda primera y segunda, al menos un recipiente que comprende un oligo de bucle de código de barras y al menos un recipiente que comprende el oligo de puente o una pluralidad de oligonucleótidos capaces de hibridarse entre sí para formar un complejo de oligo de puente, son el mismo recipiente. En tal caso, las cuatro o más sondas pueden prehibridarse y haber formado un complejo ligado.

Una ventaja particular de la invención es permitir la detección y la amplificación de la secuencia diana de interés usando diseños de sonda únicos. Las sondas se diseñan con propiedades de unión mejoradas que dan lugar a una especificidad, sensibilidad y precisión de ensayo más altas. La presente invención encuentra aplicación en el área de la biología molecular, biología evolutiva, metagenómica, genotipado y más específicamente, pero sin limitarse a diagnósticos de cáncer o trastornos cromosómicos fetales, incluyendo pero sin limitarse a genotipado de muestra(s) para una o más secuencias y/o polimorfismos, tales como SNP y/o indeles.

En una realización preferida particular, el oligo de bucle de código de barras comprende información para identificar la muestra e incluye un código de barras único. En tal caso, la primera y la segunda sonda es universalmente aplicable a todas las muestras (y solo comprende información para identificar la diana). En una realización preferida, por lo tanto, se

proporciona un método o un kit según la invención, en donde el oligo de bucle de código de barras comprende un código de barras que comprende una secuencia única que permite la enumeración de las secuencias diana de cada muestra.

Ejemplos

5 Método

1. Formación de complejos de sonda

10 Los complejos de sonda contienen secuencias requeridas para el direccionamiento genómico, la indexación de muestras y la construcción de bibliotecas de secuenciación Illumina.

Se permiten que los complejos sonda de cuatro partes se formen (tal como se ilustra en la figura 2), comprendiendo:

- 15 (a) una primera sonda que tiene, a partir del extremo 5' de la molécula, una primera secuencia específica de oligo de puente y una primera porción específica de diana en el extremo 3' de la primera sonda;
- (b) una segunda sonda que tiene, a partir del extremo 5' de la molécula, una segunda porción específica de diana, una segunda secuencia de código de barras y una segunda secuencia específica de oligo de puente en el extremo 3' de la segunda sonda;
- 20 (c) un oligo de puente con secuencias complementarias a la primera secuencia específica de oligo de puente y a la segunda secuencia específica de oligo de puente en la primera sonda y en la segunda sonda, respectivamente;
- 25 y (d) un oligo de bucle que tiene secuencias complementarias al oligo de puente.

Los complejos de sonda se construyen combinando las tres partes (puente, brazo derecho y brazo izquierdo) en cantidades equimolares en reacción de hibridación. La reacción se lleva a cabo en un termociclador (programa de hibridación en la tabla 1).

Tabla 1

Etapa	Temperatura	Tiempo
1	+95 °C	5 min
2	+95 °C	1 min
-4 °C/ciclo, pasar a 2 10x		
3	+55 °C	10 min
4	+55 °C	1 min
-5 °C/ciclo, pasar a 5 5x		
5	+4 °C	Mantener

2. Captura de diana

Se seleccionan como diana regiones genómicas específicas que contienen la(s) mutación/mutaciones de interés. El ADN purificado (por ejemplo, de tejido, plasma, orina o saliva) puede usarse como muestra o las muestras pueden no purificarse, pero solo procesarse previamente, por ejemplo, por ebullición y/o centrifugación.

Los complejos de sonda se hibridan con regiones diana mediante interacciones complementarias de secuencia base. Para iniciar la captura de diana, las sondas de reacción y el ADN diana se mezclan y se incuban en un termociclador (captura de destino y programa GapFill en la tabla 2). Los bucles de códigos de barras prehibridados (con una secuencia de índice específica para cada muestra) se añaden a la reacción de localización específica.

Tabla 2

Etapa	Temperatura	Tiempo	Proceso
1	+85 °C	4 min	Desnaturalización
2	+75 °C	2,5 min	
3	+65 °C	2,5 min	
4	+55 °C	120 min	Captura de diana
5	+50 °C	10 min	GapFill
6	+45 °C	45 min	

7	+4 °C	Mantener	
---	-------	----------	--

3. Reacción GapFill

5 Después de la captura de la diana, se agrupan los complejos de sonda de las distintas reacciones de localización específica y posteriormente se extienden y ligan añadiendo una combinación de Phusion ADN polimerasa, nucleótidos y Ampligase ADN ligasa e incubando 45 minutos a +45 °C.

10 4. Tratamientos con exonucleasas

Después de GapFill, las moléculas lineales se eliminan añadiendo 1 µl de exonucleasa termolábil 1 (NEB, n.º M0568L) y 1 µl de exonucleasa RecJF (NEB, n.º M0264L) e incubando 30 minutos a +37 °C. Las exonucleasas se inactivan incubando 12 minutos a +92 °C.

15 5. Amplificación de círculo rodante

20 Después de la extensión y ligamiento, las moléculas de sonda circulares se someten a amplificación de círculo rodante (RCA). Para la reacción de RCA, una reacción de captura de diana se mezcla con la mezcla de reacción de RCA que contiene polimerasa EquipPhi29 (Thermo Scientific). La reacción se incuba a +42 °C durante 30 min - 2 horas. Después de la reacción de RCA, la eficiencia de la reacción se analiza midiendo la concentración de ADN monocatenario (ADNmc) con fluorómetro Qubit.

25 6. Digestión enzimática

La reacción de RCA produce una molécula de ADNmc concatémica larga que tiene múltiples copias de la biblioteca diana. Cada biblioteca diana completa se separa mediante la secuencia de reconocimiento de la enzima de restricción EcoRI. Esta secuencia permite el corte específico de la secuencia del concatémico largo mediante hibridación con un oligonucleótido específico que contiene la secuencia de reconocimiento de la enzima de restricción EcoRI y la liberación de bibliotecas diana listas. Los productos de RCA se digieren con EcoRI durante 1 hora a +37 °C.

30 7. PCR de biblioteca

Los productos de RCA digeridos se extienden a bibliotecas de secuenciación en una reacción de PCR donde los adaptadores de secuenciación truncados presentes en la sonda derecha se extienden a adaptadores de secuenciación de longitud completa compatibles con celdas de flujo.

35 8. Purificación de la biblioteca

Después de la PCR de biblioteca, las moléculas de la biblioteca se purifican extrayéndolas de geles de agarosa después de la electroforesis o con perlas de selección de tamaño (tales como Macherey Nagel NucleoMag).

40 9. Secuenciación

45 Las bibliotecas compatibles con MiSeq o iSeq100 purificadas se someten a análisis de secuencia usando los instrumentos de secuenciación del estado de la técnica. Es importante destacar que las bibliotecas se pueden convertir para ajustarse en cualquier plataforma de secuenciación disponible mediante modificaciones de oligonucleótidos simples. Los datos de secuenciación se procesan usando una combinación de herramientas de línea de comandos Unix y lenguajes de programación Python y R. En resumen, la razón para el procesamiento de secuencias es identificar las secuencias de sonda dentro de cada secuencia leída, la zona genómica entre ellos y contar el número de códigos de barras moleculares asociados con cada diana genética.

50 Experimento 1

55 En el primer experimento, la mezcla de sondas consistía en una colección de cuatro complejos de sondas indexados en bucle en cinco muestras distintas, donde las cuatro primeras muestras tenían un índice de bucle único y la quinta tenía muestras 1-4 agrupadas. Se marcaron con diana cuatro fusiones de genes. Los oligonucleótidos diana tenían secuencias de reconocimiento únicas que permiten la identificación de cada diana.

60 Como muestra, se mezclaron tres oligonucleótidos diana sintéticos en concentraciones crecientes logarítmicamente. Las reacciones de captura de diana, extensión y ligamiento, amplificación por círculo rodante, posterior digestión con EcoRI y PCR de preparación de biblioteca se llevaron a cabo tal como se describió anteriormente. Un ejemplo de la biblioteca de secuenciación resultante se muestra en la figura 3.

65 Las bibliotecas listas se secuenciaron con los instrumentos MiSeq y iSeq100 y las regiones diana se detectaron dentro de los datos de secuencia haciendo coincidir las secuencias de sonda dentro de cada lectura, identificando el área de

secuencia genómica entre las secuencias de sonda y contando los códigos de barras moleculares. Los datos de recuento reflejaron con precisión las proporciones de las moléculas de oligo de bucle agrupadas (figura 4).

Descripción detallada de las figuras 1 y 2

5 La Fig. 1 ilustra el flujo de trabajo de una realización de la invención descrita. En la etapa 1, los ácidos nucleicos (ADN o ARN) extraídos de una muestra (102) se ponen en contacto con un conjunto de complejos (104) de ligamiento. Los complejos de ligamiento hibridan en los ácidos (106) nucleicos diana. En la etapa 2, los complejos de ligamiento unidos a diana se capturan opcionalmente desde el material de muestra, dejando detrás de las impurezas de la muestra (103). En la etapa 3, los complejos de ligamiento ligados de múltiples muestras (110) se ponen en conjunto juntos (112). En la etapa 4, los complejos de ligamiento hibridados agrupados se ligan, dando como resultado complejos de ligamiento ligados. En la etapa 5, las secuencias de sonda se amplifican mediante amplificación por círculo rodante usando la polimerasa phi29 u otra polimerasa de desplazamiento de hebra, dando como resultado copias concatémicas largas de las sondas (116). En la etapa 6, las copias concatémicas de sonda se escinden opcionalmente en unidades monoméricas usando una endonucleasa de restricción tal como EcoRI o una nucleasa de migración tal como I-CeuI y, opcionalmente, se amplifican adicionalmente usando PCR o PCR en emulsión (117). En la etapa 7, el ADN amplificado se secuencia usando secuenciación de ADN de nueva generación. En la etapa 8, los resultados de secuenciación de ADN se convierten en recuentos diana usando un proceso bioinformático.

20 La figura 2A ilustra una estructura principal de cuadruplete de sondas según una realización en la presente memoria. La pluralidad de entidades de sonda incluye una primera sonda (202), una segunda sonda (201), un oligo (200) de puente y un oligo (217) de bucle de código de barras. En este caso, el complejo de sondas contiene huecos o cortes entre la primera sonda y el oligo (210 y 213) de bucle de código de barras, entre la segunda sonda y el oligo (207 y 215) de bucle de código de barras y entre la primera y segunda sonda (203 y 204). Estos espacios se llenan introduciendo una polimerasa y uno o más nucleótidos. Para este proceso se puede usar una mezcla de fragmento Stoffel, polimerasa Taq o polimerasa Phusion y ADN ligasa tal como Ampligase. La polimerasa llena estos huecos y la acción posterior de la ADN ligasa da como resultado el ligamiento de la sonda, los oligos de puente y de bucle de código de barras en un complejo circular.

30 15-25 bases de la primera sonda incluye una secuencia (210) de unión de puente, que incluye opcionalmente bases químicamente modificadas para la unión eficiente del oligo de puente. La primera sonda incluye además 15-30 bases desde el extremo 5', uniéndose a la diana genética (203). Algunos o todos de los nucleótidos de 210 pueden incluir modificaciones químicas que aumenten la afinidad de las sondas a la diana o al puente (209). La última base de la primera sonda incluye opcionalmente un resto fosfato para ligamiento enzimático o una modificación que permita el ligamiento químico al extremo 5' de la sonda (205) adyacente.

35 La primera base de la segunda sonda incluye opcionalmente un resto fosfato para ligamiento enzimático o una modificación que permita el ligamiento químico al extremo 5' de la sonda (206) adyacente. Las 15-30 bases del extremo 5' de la segunda sonda incluyen una parte de la segunda sonda que se une a la diana (204) genética. Las últimas 15-25 bases de la segunda sonda (207) son complementarias inversas al oligo (208) de puente. Algunos o todos los nucleótidos de 203, 204, 207 o 210 pueden incluir modificaciones químicas que aumenten la afinidad de las sondas a la diana o al oligo de puente.

40 Las primeras 15-25 bases desde el extremo 5' del oligo (209) de puente, son complementarias inversas a la secuencia específica de oligo de puente de la primera sonda (210) y opcionalmente incluyen nucleótidos químicamente modificados para una unión aumentada. Las últimas 15-25 bases del oligo (208) de puente, son complementarias inversas a la secuencia de la segunda sonda (207) y opcionalmente incluyen nucleótidos químicamente modificados para una unión aumentada. El extremo 5' del oligo de puente incluye opcionalmente un resto (211) de captura usado para capturar los complejos de ligamiento. Es más, el oligo de puente comprende las secuencias 214 y 216, complementarias a las secuencias 213 y 215 del oligo de bucle de código de barras. El extremo 3' del oligo de puente incluye opcionalmente un resto (212) fosfato (u otro elemento escindible) para evitar la extensión durante el llenado de huecos.

50 Las primeras 15-25 bases del extremo 5' del oligo (215) de bucle de código de barras son complementarias inversas a la secuencia 216 de oligo de puente. El oligo de bucle de código de barras comprende una región de bucle que comprende un código de barras (218). Las últimas 15-25 bases del oligo (213) de bucle de código de barras son complementarias inversas a la secuencia 214 de oligo de puente.

55 La figura 2B ilustra una estructura principal de un conjunto de sondas que tiene una pluralidad de entidades de sonda según una realización de la invención. Las sondas corresponden a las de la figura 2A, excepto que esta realización comprende dos oligos (200 y 220) de puente, en donde el oligo 200 de puente comprende 209 y 214 y el oligo 220 de puente comprende 208 y 216.

60 La figura 2C ilustra una estructura principal de un conjunto de sondas que tiene una pluralidad de entidades de sonda según una realización de la invención. Las sondas corresponden a las de la figura 2A, excepto que el oligo 200 de puente contiene secuencias (219) que comprenden una pluralidad de análogos de bases universales para permitir la incorporación de secuencias aleatorias adecuadas para usar como código de barras molecular para la enumeración de dianas. Pueden estar presentes una o ambas secuencias 219 indicadas.

65

La figura 2D ilustra una estructura principal de cuadruplete de sondas según una realización en la presente memoria. La pluralidad de entidades de sonda incluye una primera sonda (202), una segunda sonda (201), un oligo (200) de puente y un oligo (217) de bucle de código de barras. En este caso, el complejo de sondas contiene huecos entre la primera sonda y el oligo (210 y 213) de bucle de código de barras, entre la segunda sonda y el oligo (207 y 215) de bucle de código de barras y entre la primera y segunda sonda (203 y 204). Estos espacios se llenan introduciendo una polimerasa y uno o más nucleótidos. Para este proceso se puede usar una mezcla de fragmento Stoffel, polimerasa Taq o polimerasa Phusion, y ADN ligasa tal como Ampligase. La polimerasa llena estos huecos y la acción posterior de la ADN ligasa da como resultado el ligamiento de la sonda, los oligos de puente y de bucle de código de barras en un complejo circular.

15-25 bases de la primera sonda incluye una secuencia (210) de unión de puente, que incluye opcionalmente bases químicamente modificadas para la unión eficiente del oligo de puente. La primera sonda incluye además un sitio de unión para un cebador (221) de amplificación y una secuencia (222) de código de barras y 15-30 bases desde el extremo 5', una secuencia que se une a la diana (203) genética. Algunos o todos de los nucleótidos de 210 pueden incluir modificaciones químicas que aumenten la afinidad de las sondas a la diana o al puente (209). La última base de la primera sonda incluye opcionalmente un resto fosfato para ligamiento enzimático o una modificación que permita el ligamiento químico al extremo 5' de la sonda (205) adyacente.

La primera base de la segunda sonda incluye opcionalmente un resto fosfato para ligamiento enzimático o una modificación que permita el ligamiento químico al extremo 5' de la sonda (206) adyacente. Las 15-30 bases del extremo 5' de la segunda sonda incluyen una parte de la segunda sonda que se une a la diana (204) genética. La segunda sonda incluye además un sitio de unión para un cebador (223) de amplificación, una secuencia (224) adaptadora de secuenciación, un sitio de reconocimiento para una endonucleasa de restricción tal como EcoRI o una endonucleasa (225) de migración y otra secuencia (207) adaptadora de secuenciación. Las últimas 15-25 bases de la segunda sonda (207) son complementarias inversas al oligo (208) de puente. Algunos o todos los nucleótidos de 203, 204, 207 o 210 pueden incluir modificaciones químicas que aumenten la afinidad de las sondas a la diana o al oligo de puente.

Las primeras 15-25 bases desde el extremo 5' del oligo (209) de puente, son complementarias inversas a la secuencia específica de oligo de puente de la primera sonda (210) y opcionalmente incluyen nucleótidos químicamente modificados para una unión aumentada. Las últimas 15-25 bases del oligo (208) de puente, son complementarias inversas a la secuencia de la segunda sonda (207) y opcionalmente incluyen nucleótidos químicamente modificados para una unión aumentada. La parte del oligo de puente no complementaria inversa con ninguno de los extremos del oligo (220) de bucle de código de barras contiene opcionalmente un sitio de reconocimiento para una endonucleasa de restricción para permitir la degradación enzimática de las estructuras sin bucle. El extremo 5' del oligo de puente incluye opcionalmente un resto (211) de captura usado para capturar los complejos de ligamiento. Es más, el oligo de puente comprende las secuencias 214 y 216, complementarias a las secuencias 213 y 215 del oligo de bucle de código de barras. El extremo 3' del oligo de puente incluye opcionalmente un resto (212) fosfato (u otro elemento escindible) para evitar la extensión durante el llenado de huecos.

Las primeras 15-25 bases del extremo 5' del oligo (215) de bucle de código de barras son complementarias inversas a la secuencia 216 de oligo de puente. El oligo de bucle de código de barras comprende una región de bucle que comprende un código de barras (218). Las últimas 15-25 bases del oligo (213) de bucle de código de barras son complementarias inversas a la secuencia 214 de oligo de puente.

REIVINDICACIONES

1. Un método para la detección de alto rendimiento de una o más secuencias de nucleótidos diana en una pluralidad de muestras, comprendiendo el método las etapas de:

(i) proporcionar para cada secuencia de nucleótidos diana en cada una de las muestras:

una primera sonda, una segunda sonda, un oligo de bucle de código de barras y un oligo de puente o una pluralidad de oligonucleótidos de puente capaces de hibridarse con el oligo de bucle de código de barras para formar un complejo de oligo de puente, en donde la primera sonda comprende una primera secuencia específica de oligo de puente en el extremo 5' de la primera sonda y una primera porción específica de diana en el extremo 3' de la primera sonda;

en donde la segunda sonda comprende una segunda porción específica de diana en el extremo 5' de la segunda sonda y una segunda secuencia específica de oligo de puente en el extremo 3' de la segunda sonda;

en donde el oligo de bucle de código de barras comprende, a partir del extremo 5' de la molécula, una tercera secuencia específica de oligo de puente, una secuencia de bucle con código de barras y una cuarta secuencia específica de oligo de puente, en donde el oligo de puente o la pluralidad de oligonucleótidos de puente contiene secuencias complementarias a la primera secuencia específica de oligo de puente y a la segunda secuencia específica de oligo de puente en la primera sonda y la segunda sonda, respectivamente, y secuencias complementarias a la tercera secuencia específica de oligo de puente y a la cuarta secuencia específica de oligo de puente en el oligo de bucle de código de barras;

y en donde, opcionalmente, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende una secuencia de reconocimiento para una endonucleasa;

(ii) poner en contacto, para cada una de las una o más secuencias de nucleótidos diana, la primera sonda y la segunda sonda con, preferiblemente para cada una de las muestras en un tubo distinto, el oligo de bucle de código de barras y el oligo de puente o la pluralidad de oligonucleótidos de puente y permitir autohibridarse en complejos de ligamiento;

(iii) poner en contacto los ácidos nucleicos presentes en cada una de las muestras que va a someterse a prueba para las secuencias de nucleótidos diana, con los complejos de ligamiento;

(iv) permitir que la primera porción específica de diana y la segunda porción específica de diana de la respectiva primera sonda y segunda sonda se hibriden con secciones esencialmente adyacentes en la secuencia diana, formando de este modo complejos de hibridación;

(v) agrupar opcionalmente los complejos de hibridación de la pluralidad de muestras;

(vi) ligar las sondas en los complejos de hibridación para proporcionar complejos de ligamiento ligados;

(vii) amplificar ácidos nucleicos de los uno o más complejos de ligamiento ligados usando amplificación por círculo rodante con una polimerasa de desplazamiento de hebra, obteniendo de este modo secuencias concatémicas monocatenarias;

(viii) opcionalmente, siempre que esté presente una secuencia de reconocimiento como la especificada en la etapa (i), realizar una etapa para obtener fragmentos de ácido nucleico mediante:

(a) escindir las secuencias concatémicas monocatenarias obtenidas en la etapa (vii), o

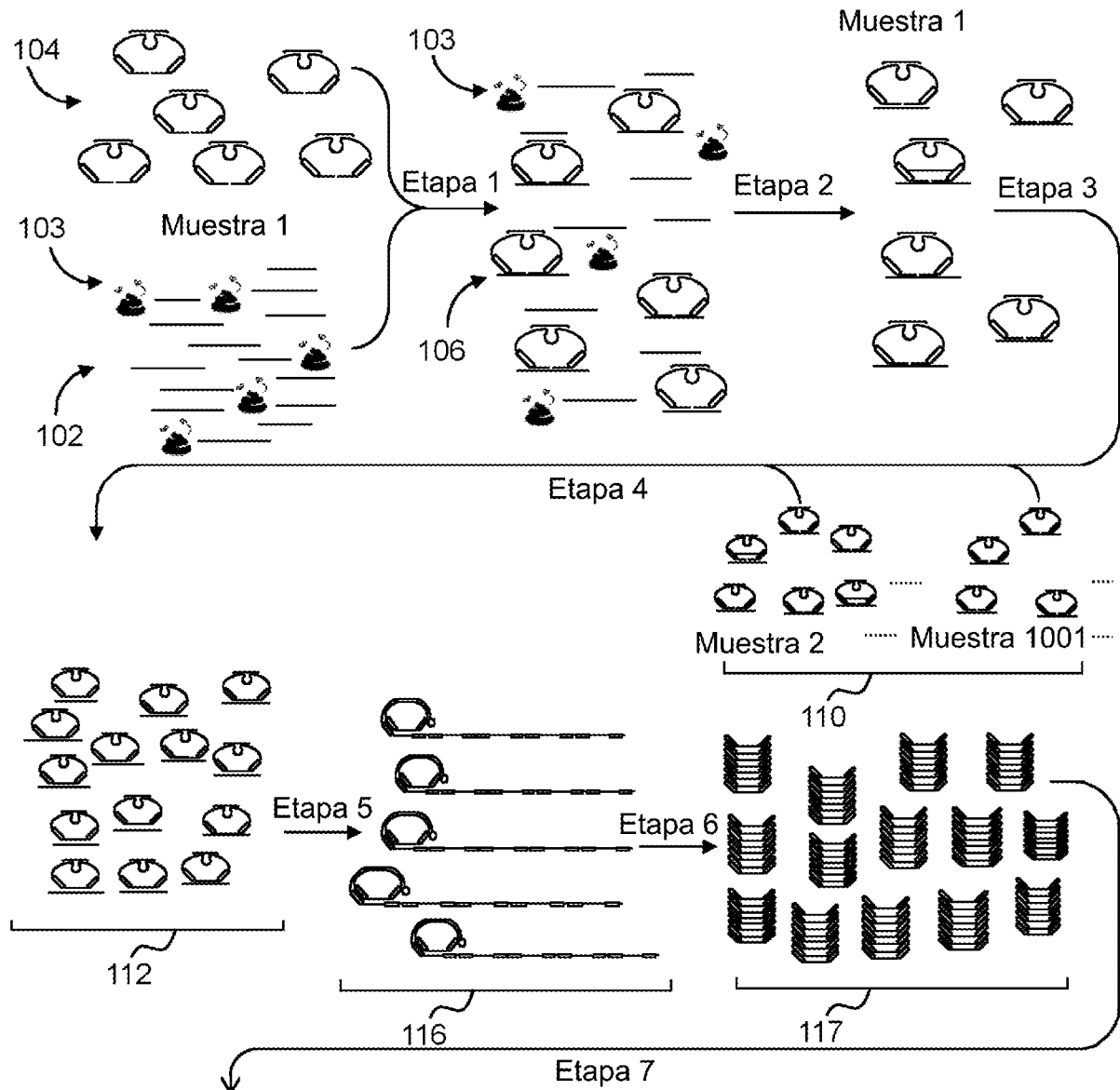
(b) someter las secuencias concatémicas monocatenarias obtenidas en la etapa (vii) para hibridarse con un oligonucleótido específico que contiene una secuencia de reconocimiento para una endonucleasa en donde el oligonucleótido se hibrida con la secuencia de reconocimiento especificada en la etapa (i) de modo que se obtiene un sitio de reconocimiento para la endonucleasa y escindir los complejos hibridados con dicha endonucleasa;

(ix) someter la secuencia concatémica obtenida en la etapa (vii) o los fragmentos de ácido nucleico obtenidos en la etapa (viii) a una tecnología de secuenciación de alto rendimiento para determinar la secuencia o secuencias de códigos de barras; y

(x) identificar la presencia y/o el número de la secuencia de nucleótidos diana en la pluralidad de muestras mediante la determinación de al menos parte de la primera porción específica de diana y/o la segunda porción específica de diana, y/o al menos parte de la secuencia de código de barras correspondiente al código de barras en el oligo de bucle de código de barras, en donde las etapas (v) y (vi) pueden realizarse en cualquier orden.

2. Método según la reivindicación 1, en donde el oligo de bucle de código de barras contiene una o más secuencias de reconocimiento para una endonucleasa, tal como una endonucleasa de corte.
- 5 3. Método según la reivindicación 2, en donde dicho oligo de bucle de código de barras contiene dos secuencias de reconocimiento que son capaces de hibridarse entre sí de tal modo que se obtiene un sitio de reconocimiento de endonucleasas bicatenarias.
- 10 4. Método según la reivindicación 3, en donde el método no comprende la etapa (viii) como se ha especificado en la reivindicación 1, sino que comprende, entre las etapas (vii) y (ix), una etapa para permitir que dichas dos secuencias de reconocimiento en el oligo de bucle de código de barras se hibriden y escindir el sitio de reconocimiento de la endonucleasa bicatenaria resultante con una endonucleasa, tal como una endonucleasa de corte, que tiene especificidad para dicho sitio de reconocimiento.
- 15 5. Método según una cualquiera de las reivindicaciones anteriores, en donde el oligo de puente, o uno o más oligonucleótidos de la pluralidad de oligonucleótidos de puente, comprende, en una región no complementaria a la primera sonda, la segunda sonda o el oligo de bucle de código de barras, una pluralidad de análogos de bases universales para permitir la incorporación de secuencias aleatorias adecuadas para usar como código de barras molecular para la enumeración de dianas, y en donde, como parte de la etapa (vi), se realiza una etapa de llenado de huecos usando polimerasa y nucleótidos con el fin de generar tales secuencias aleatorias.
- 20 6. Método según la reivindicación 5, en donde dicha pluralidad de análogos de bases universales es una pluralidad de 5-nitroindoles.
- 25 7. Método según una cualquiera de las reivindicaciones anteriores, en donde la primera sonda, la segunda sonda o el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende además un código de barras de secuencia.
- 30 8. Método según una cualquiera de las reivindicaciones anteriores, en donde,
después de la etapa (v), pero antes de la etapa (vii), se realizan una etapa (a) y una etapa (b), en donde la etapa (a) comprende permitir que los complejos de ligamiento ligados se disocien de la secuencia de nucleótidos diana y la etapa (b) comprende añadir una sonda específica de diana que comprende una secuencia correspondiente a la secuencia de nucleótidos diana, en donde dicha sonda específica de diana es capaz de hibridarse con los complejos de ligamiento ligados y permitir que la sonda específica de diana se hibride con los complejos de ligamientos ligados, formando de este modo moldes de amplificación, y
35 en donde, en la etapa (vii), dichos moldes de amplificación se amplifican mediante amplificación por círculo rodante con una polimerasa de desplazamiento de hebra.
- 40 9. Método según una cualquiera de las reivindicaciones anteriores, en donde al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos puente, comprende un primer resto de captura, y en donde entre las etapas (iv) y (v) se realiza una etapa intermedia (iv)(a) que comprende poner el complejo de hibridación en contacto con un soporte sólido que comprende un segundo resto de captura, lo que permite que el primer resto de captura y el segundo resto de captura interactúen de tal modo que los complejos de hibridación se ligan al soporte sólido y separar los complejos de hibridación ligados a un soporte sólido de los componentes de las muestras que no están ligados al soporte sólido.
- 45 10. Método según una cualquiera de las reivindicaciones anteriores, en donde la pluralidad de muestras incluye una muestra de sangre, una muestra de saliva, una muestra de orina o una muestra de heces.
- 50 11. Método según una cualquiera de las reivindicaciones anteriores, en donde el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende:
55 (i) de una a cinco bases protuyentes en 3', y/o
(ii) 3' fosfato, y/o
(iii) una o más modificaciones de fosforotioato dentro de tres posiciones desde el extremo 3'.
- 60 12. Método según una cualquiera de las reivindicaciones anteriores, en donde el extremo 3' de la primera sonda o el extremo 5' de la segunda sonda, o ambos, se modifican para permitir el ligamiento químico de la primera sonda a la segunda sonda.
- 65 13. Método según una cualquiera de las reivindicaciones anteriores, en donde la porción de puente de la primera sonda o la segunda sonda, o ambas, o el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprenden bases modificadas químicamente para permitir una mejor unión al oligo de puente o complejo de oligo de puente.

- 5 14. Método según una cualquiera de las reivindicaciones anteriores, en donde la primera porción específica de diana, la segunda porción específica de diana, las primeras secuencias específicas de oligo de puente y/o las segundas secuencias específicas de oligo de puente, contienen independientemente entre sí, uno o más nucleótidos químicamente modificados.
15. Método según una cualquiera de las reivindicaciones anteriores, en donde el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente, comprende uno o más nucleótidos químicamente modificados.
- 10 16. Método según una cualquiera de las reivindicaciones anteriores, en donde la etapa (vii) se realiza usando una polimerasa phi29 o una polimerasa Bst.
- 15 17. Método según una cualquiera de las reivindicaciones anteriores, en donde inmediatamente antes de la etapa (ix) se realiza una amplificación por PCR usando cebadores que se unen a partes universales de la primera y segunda sonda, en donde dichos cebadores incluyen opcionalmente adaptadores para la posterior secuenciación en la etapa (ix).
- 20 18. Método según una cualquiera de las reivindicaciones anteriores, en donde la secuenciación en la etapa (ix) se realiza usando secuenciación por nanoporos, en donde opcionalmente la secuencia concatémica obtenida en la etapa (vii) se fragmenta usando complejos de transposición.
- 25 19. Método según una cualquiera de las reivindicaciones anteriores, en donde la enumeración de diana genética se permite contando el número de códigos de barra moleculares por diana y por muestra.
- 20 20. Método según una cualquiera de las reivindicaciones anteriores, en donde para dos o más muestras o para dos o más combinaciones de locus/alelo, se usan secuencias de código de barras para genotipar la muestra o muestras para una o más secuencias y/o polimorfismos, tales como SNP y/o indeles.
- 30 21. Kit de partes que comprende una pluralidad de recipientes, en donde al menos un recipiente comprende uno o más conjuntos de primera sonda y segunda sonda, al menos un recipiente comprende un oligo de bucle de código de barras y al menos un recipiente comprende uno o más oligos de puente o una pluralidad de oligonucleótidos de puente capaces de formar un complejo de oligo de puente con el oligo de bucle de código de barras,
- 35 en donde la primera sonda comprende una primera secuencia específica de oligo de puente en el extremo 5' de la primera sonda y una primera porción específica de diana en el extremo 3' de la primera sonda;
- en donde la segunda sonda comprende una segunda porción específica de diana en el extremo 5' de la segunda sonda y una segunda secuencia específica de oligo de puente en el extremo 3' de la segunda sonda;
- 40 en donde el oligo de bucle de código de barras comprende, a partir del extremo 5' de la molécula, una tercera secuencia específica de oligo de puente, una secuencia de bucle de código de barras y una cuarta secuencia específica de oligo de puente,
- 45 en donde el oligo de puente o la pluralidad de oligonucleótidos de puente contiene secuencias complementarias a la primera secuencia específica de oligo de puente y a la segunda secuencia específica de oligo de puente en la primera sonda y la segunda sonda, respectivamente, y secuencias complementarias a la tercera secuencia específica de oligo de puente y a la cuarta secuencia específica de oligo de puente en el oligo de bucle de código de barras;
- 50 y en donde, opcionalmente, al menos uno de: la primera sonda, la segunda sonda, el oligo de bucle de código de barras, el oligo de puente o un oligonucleótido de la pluralidad de oligonucleótidos de puente comprende una secuencia de reconocimiento para una endonucleasa; y en donde opcionalmente el kit de partes comprende además un oligonucleótido capaz de hibridarse con dicha secuencia de reconocimiento de tal modo que se obtiene un sitio de reconocimiento para dicha endonucleasa.
- 55
- 60
- 65



```

GGGGCCCCGCGTCGATCGGAGCCGTTAGGAT
TTAAGGTGCCGTCGATCGGAGCCGACGTACG
TTAAGGTGCCGTCGATCGGAGCCGACGTACG
TATAATAGAGGTTCGTGCAGTCACGACCCGGT
ACCAGGTGCCGTCGATCGGAGCCGACCCGGT
GGGGCCGGAGGTCGTGCAGTCACGTTAGGAT
TCCAGGTGCAGTCGATCCGTCACGTACGTACG
AAAATTTTAGCGTACGTCGTACGTTTAGGAT
TATAATAGAGGTTCGTGCAGTCACGACCCGGT
AGGACCTTGAGTCGATCCGCACGTACCCGGT
AGCGACCGAGGTCGTGCAGTCACGACGTACG
TATAATAGAGGTTCGTGCAGTCACGACCCGGT
GGAAAAAGCCGTCGATCGGAGCCGTTAGGAT
ATATACAGAGGTTCGTGCAGTCACGTTAGGAT
GAGAGCCGAGGTCGTGCAGTCACGACCCGGT
CGCAGCGGAGGTCGTGCAGTCACGACGTACG
ATTACAGCCGTCGATCGGAGCCGACCCGGT
TATAATAGAGGTTCGTGCAGTCACGACCCGGT
GGGCAATTAGCGTACGTCGTACGTACGTACG
TATCCGAGCCGTCGATCGGAGCCGACGTACG
AGGACCTTGAGTCGATCCGCACGTACCCGGT
AATTACAAGCCGTCGATCGGAGCCGACCCGGT
GAAGAATTAGCGTACGTCGTACGTTTAGGAT
GGGCAATTAGCGTACGTCGTACGTACGTACG
GGGCAATTAGCGTACGTCGTACGTACGTACG
GAGCACTTAGCGTACGTCGTACGTACCCGGT
AAAGGGGCGAGTCGATCCGCACGTTTAGGAT
AGCGCGCCGTCGATCGGAGCCGTTAGGAT
    
```

Resultado final:
 Cuantificación absoluta, precisa de diana para
 múltiples dianas y muestras genéticas

	Diana 1	Diana 2	...	Diana 50
Muestra 1	1007 copias	58 copias	...	102 copias
Muestra 2	300 copias	550 copias	...	10 copias
⋮	⋮	⋮	⋮	⋮
Muestra 1001	225 copias	37 copias	...	1003 copias
⋮	⋮	⋮	⋮	⋮

Figura 1

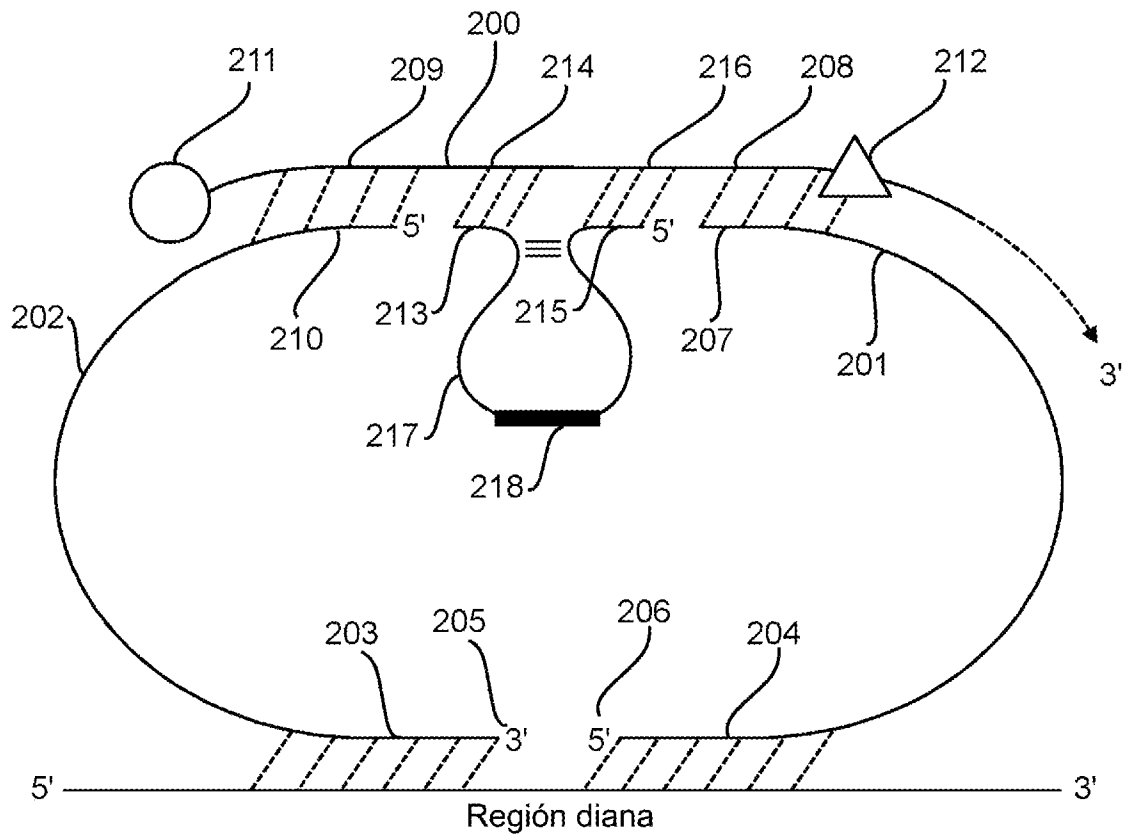


Figura 2A

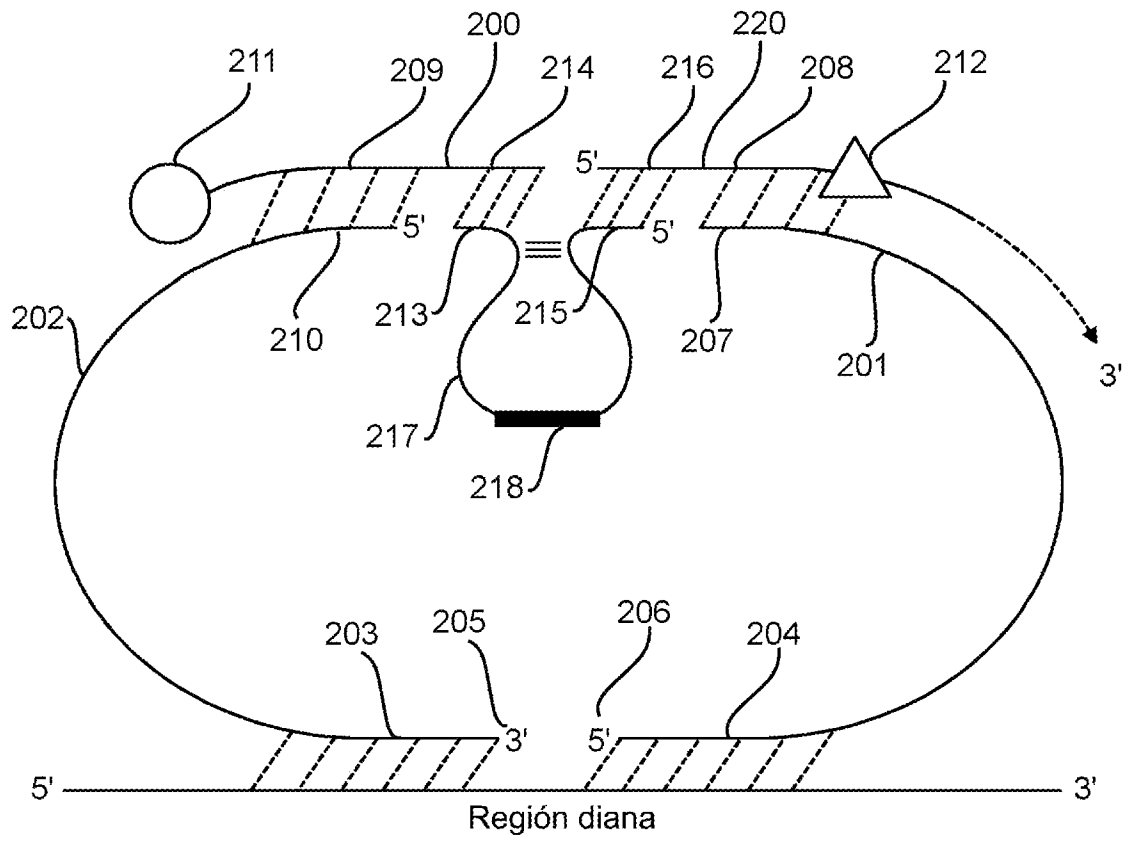


Figura 2B

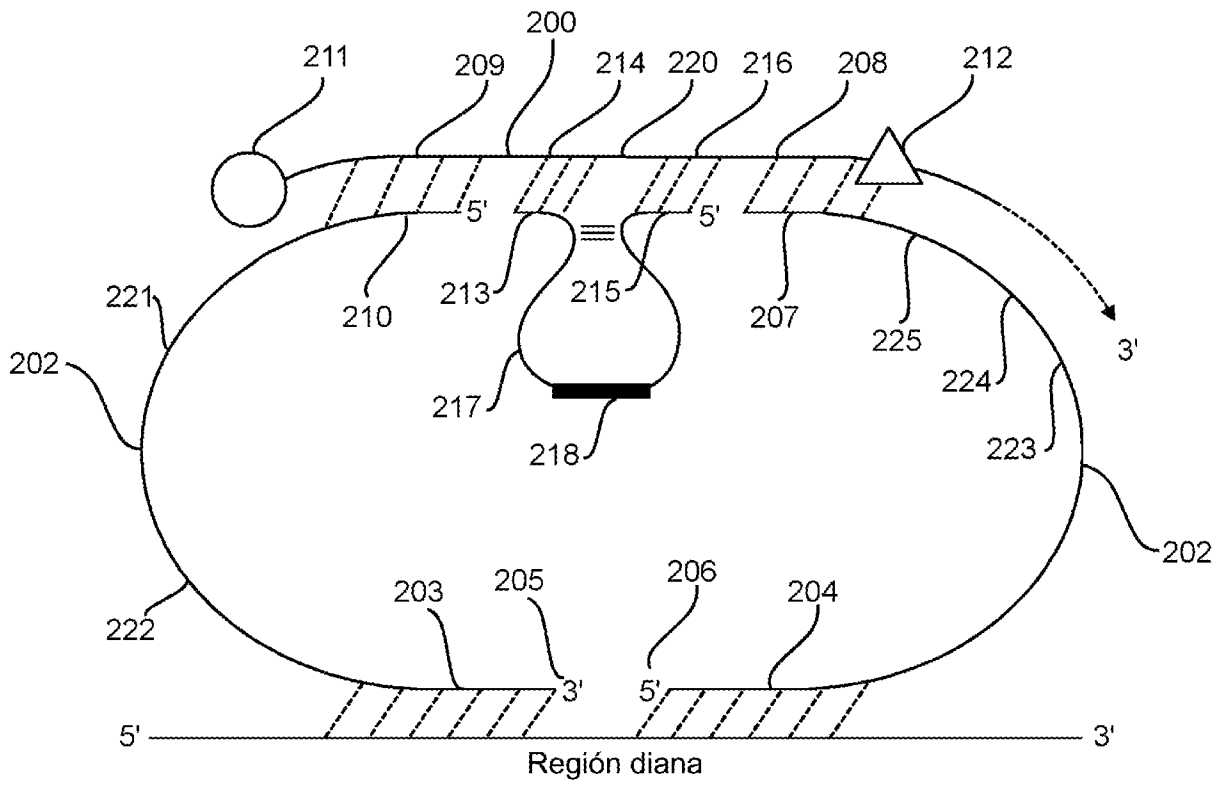


Figura 2D

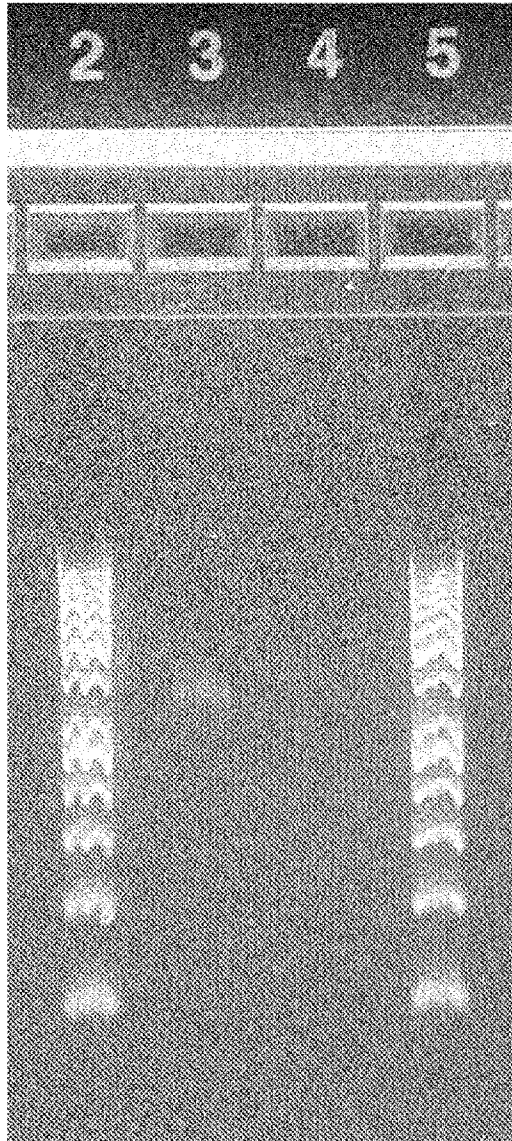


Figura 3

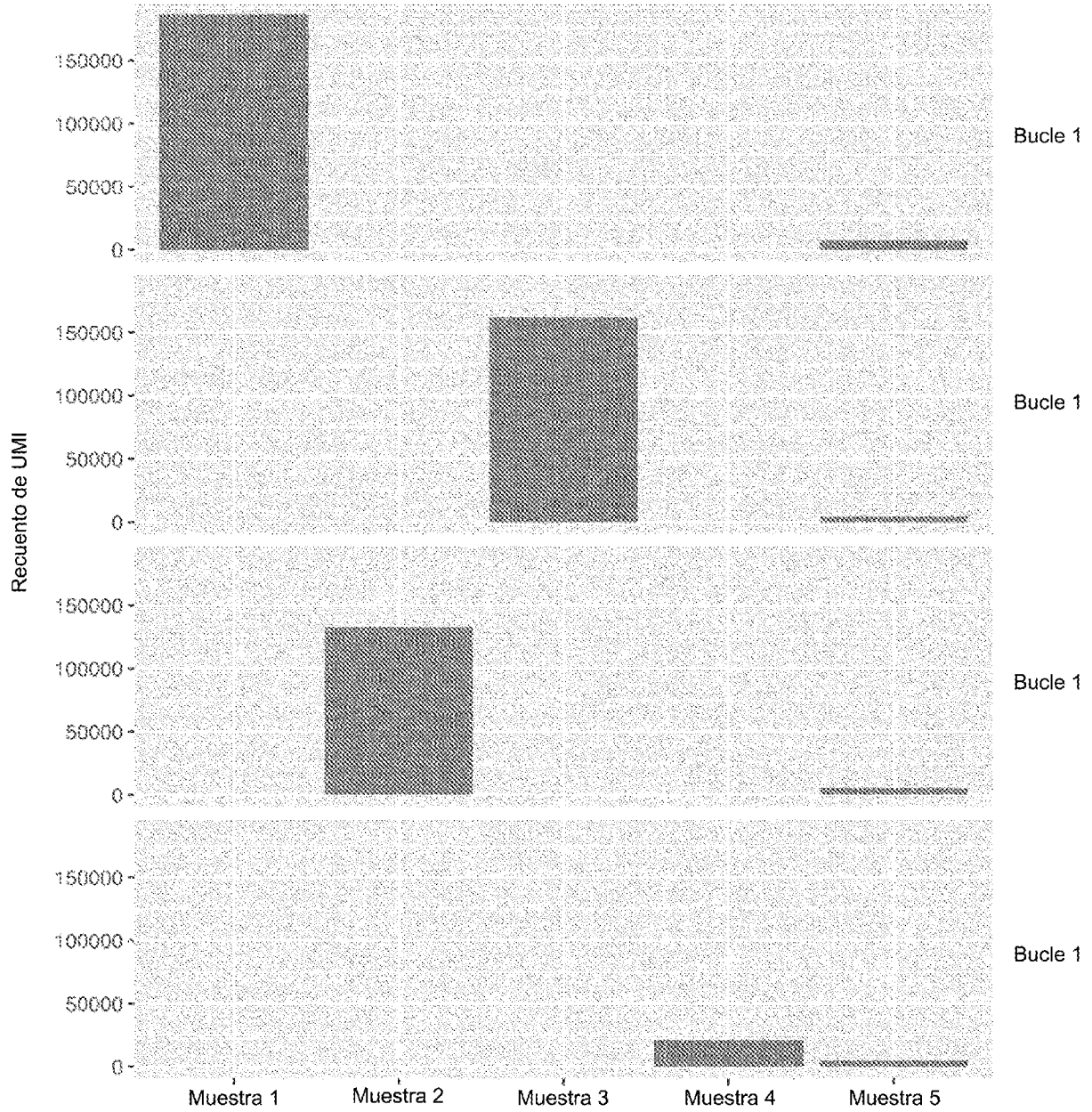


Figura 4