

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 979 286**

51 Int. Cl.:

**C12Q 1/6869** (2008.01)

**C12Q 1/6834** (2008.01)

**C12N 15/10** (2006.01)

**C40B 50/14** (2006.01)

**C40B 50/18** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.01.2014** **E 18198695 (1)**

97 Fecha y número de publicación de la concesión europea: **17.04.2024** **EP 3486331**

54 Título: **Preparación de muestras sobre un soporte sólido**

30 Prioridad:

**09.01.2013 US 201361750682 P**

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

**25.09.2024**

73 Titular/es:

**ILLUMINA CAMBRIDGE LIMITED (100.0%)**  
**Illumina Centre19, Granta Park, Great Abington**  
**Cambridge, Cambridgeshire CB21 6DF, GB**

72 Inventor/es:

**GORMLEY, NIALL, ANTHONY y**  
**SMITH, GEOFFREY, PAUL**

74 Agente/Representante:

**DEL VALLE VALIENTE, Sonia**

ES 2 979 286 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

Preparación de muestras sobre un soporte sólido

## 5 Solicitudes relacionadas

Esta solicitud reivindica prioridad sobre la Solicitud Provisional de EE. UU. n.º 61/750.682, presentada el 9 de enero de 2013.

## 10 Antecedentes

Existe una variedad de métodos y aplicaciones para los que es deseable generar una biblioteca de moléculas de ADN fragmentadas y marcadas a partir de moléculas de ADN bicatenario (ADNbc) diana. A menudo, el objetivo es generar moléculas de ADN más pequeñas (p. ej., fragmentos de ADN) a partir de moléculas de ADNbc más grandes para usarlas como moldes en reacciones de secuenciación de ADN.

Muchos de los métodos utilizados actualmente para la fragmentación y el marcaje del ADN bicatenario para su uso en la secuenciación de próxima generación desperdician el ADN, requieren instrumentos costosos para la fragmentación y los procedimientos para fragmentar, marcar y recuperar fragmentos de ADN marcados son difíciles, tediosos, laboriosos, lentos, ineficaces, costosos y requieren cantidades relativamente grandes de muestras de ácidos nucleicos. Además, muchos de estos métodos generan fragmentos de ADN marcados que no son completamente representativos de las secuencias contenidas en las muestras de ácidos nucleicos a partir de los cuales se generaron. Por lo tanto, en la técnica se necesitan métodos que proporcionen velocidad y facilidad de uso a la hora de generar bibliotecas de fragmentos de ADN marcados a partir del ADN diana y que puedan aplicarse fácilmente a los métodos de análisis de ácidos nucleicos, tales como los métodos de secuenciación y amplificación de próxima generación.

## 25 Breve resumen

En el presente documento se presentan métodos y composiciones para la preparación de muestras de ácido nucleico sobre un soporte sólido. La invención se define mediante las reivindicaciones adjuntas. Los métodos y composiciones se refieren en particular a métodos y composiciones para fragmentar y marcar el ADN usando composiciones de transposones inmovilizadas en un soporte sólido. Los métodos y composiciones presentados en el presente documento son útiles, por ejemplo, para generar bibliotecas de fragmentos de ADN marcados para su uso, por ejemplo, en métodos de secuenciación de próxima generación y similares. La preparación de fragmentos de ADNmc lineales sobre un soporte sólido a partir de ADN diana puede comprender cualquier ADNbc de interés (incluido el ADNc bicatenario preparado a partir de ARN), de cualquier fuente, para el análisis genómico, subgenómico, transcriptómico o metagenómico, o el análisis de la expresión del ARN.

En el presente documento se presentan métodos para preparar una biblioteca inmovilizada de fragmentos de ADN marcados que comprenden: (a) proporcionar un soporte sólido que tenga complejos transposómicos inmovilizados en el mismo, en donde los complejos transposómicos comprenden una transposasa unida a un primer polinucleótido, comprendiendo el primer polinucleótido (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) un primer marcador que comprende un primer dominio de marcaje; y (b) aplicar un ADN diana al soporte sólido en condiciones donde el ADN diana se fragmente mediante los complejos transposómicos y la secuencia terminal del transposón 3' del primer polinucleótido se transfiera a un extremo 5' de al menos una cadena de los fragmentos; produciendo de esta manera una biblioteca inmovilizada de fragmentos bicatenarios en donde al menos una cadena está marcada en 5' con el primer marcador. En algunas realizaciones, los complejos transposómicos comprenden un segundo polinucleótido que comprende una región complementaria de dicha secuencia terminal del transposón. Los métodos pueden comprender además (c) proporcionar complejos transposómicos en solución y poner en contacto los complejos transposómicos con los fragmentos inmovilizados en condiciones donde el ADN diana sea fragmentado por los complejos transposómicos en solución; obteniendo de esta manera fragmentos de ácido nucleico inmovilizados que tienen un extremo en solución. En algunas realizaciones, los complejos transposómicos en solución pueden comprender un segundo marcador, de modo que el método genere fragmentos de ácido nucleico inmovilizados que tienen un segundo marcador, el segundo marcador en solución. El primer y el segundo marcadores pueden ser diferentes o iguales.

También se presentan en el presente documento soportes sólidos que tienen una biblioteca de fragmentos de ADN marcados inmovilizados en los mismos preparados según los métodos anteriores u otros métodos. Por ejemplo, en el presente documento se presentan soportes sólidos que tienen complejos transposómicos inmovilizados en los mismos, en donde los complejos transposómicos comprenden una transposasa unida a un primer polinucleótido, comprendiendo el polinucleótido (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) un primer marcador que comprende un primer dominio de marcaje.

En el presente documento también se presentan métodos para generar una cubeta de lectura, que comprenden inmovilizar una pluralidad de complejos transposómicos en un soporte sólido, comprendiendo los complejos transposómicos una transposasa unida a un primer polinucleótido, comprendiendo el primer polinucleótido (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) un primer marcador que comprende un primer dominio de marcaje.

Los métodos pueden comprender además proporcionar un soporte sólido que tenga una pluralidad de los primeros polinucleótidos inmovilizados en el mismo y poner en contacto el soporte sólido con la holoenzima transposasa y un segundo polinucleótido, comprendiendo el segundo polinucleótido una región complementaria de la secuencia terminal de transposón. En algunas realizaciones de los métodos, la inmovilización comprende (a) proporcionar un soporte sólido que tiene cebadores de amplificación acoplados al mismo; (b) hibridar un segundo polinucleótido con uno de los cebadores de amplificación, comprendiendo el segundo oligonucleótido una región complementaria de una secuencia terminal de transposón y una región complementaria del primer marcador; (c) extender el cebador de amplificación usando una polimerasa para generar una doble cadena que comprende el primer polinucleótido hibridado con el segundo polinucleótido, el primer polinucleótido inmovilizado directamente en el soporte sólido; y (d) poner en contacto el soporte sólido con la holoenzima transposasa, ensamblando así un complejo transposómico sobre el soporte sólido.

También se presenta en el presente documento una población de micropartículas que tienen complejos transposómicos inmovilizados en las mismas, comprendiendo los complejos transposómicos una transposasa unida a un primer polinucleótido y a un segundo polinucleótido; en donde el primer polinucleótido se inmoviliza en su extremo 5' en la superficie de la micropartícula y el segundo polinucleótido se hibrida con el extremo 3' del primer polinucleótido; y en donde el primer polinucleótido comprende: (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) un primer marcador que comprende un primer dominio de marcaje. En el presente documento también se presentan métodos para producir una biblioteca inmovilizada de fragmentos de ADN marcados que comprenden poner en contacto un ADN diana con la población de micropartículas anterior para generar fragmentos de ADN marcados inmovilizados.

En el presente documento también se presentan métodos para generar una biblioteca de fragmentos de ADN marcados para su ensamblaje dirigido por el índice en una lectura de secuencia más larga, comprendiendo los métodos: proporcionar una población de micropartículas que tienen complejos transposómicos inmovilizados en las mismas, comprendiendo los complejos transposómicos una transposasa unida a un primer polinucleótido que comprende un dominio de índice asociado a la micropartícula y un segundo polinucleótido; aplicar un ADN diana a la población de micropartículas, generando de esta manera fragmentos de ADN inmovilizados que están marcados con el dominio índice. En determinadas realizaciones de los métodos anteriores, el primer polinucleótido se inmoviliza en su extremo 5' en la superficie de la micropartícula y el segundo polinucleótido se hibrida con el extremo 3' del primer polinucleótido; y en donde el primer polinucleótido comprende: (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) el dominio de índice; y en donde la población de micropartículas comprende al menos una pluralidad de dominios de índice y en donde los primeros polinucleótidos de una micropartícula individual comparten el mismo dominio de índice.

También se presenta en el presente documento un método para secuenciar una pluralidad de moléculas de ADN diana, que comprende: aplicar una pluralidad de ADN diana a un soporte sólido que tiene complejos transposómicos inmovilizados en el mismo en condiciones donde el ADN diana se fragmenta mediante los complejos transposómicos; produciendo de esta manera una biblioteca inmovilizada de fragmentos bicatenarios, en donde una primera parte de cada ADN diana está unida a dicho soporte sólido en una primera ubicación de dicho soporte sólido y una segunda parte de cada ADN diana está unida a dicho soporte sólido en una segunda ubicación de dicho soporte sólido; y cartografiar dicha biblioteca inmovilizada de fragmentos bicatenarios para generar un conjunto de ubicaciones que están unidas por cada ADN diana; determinar las secuencias de dichas primera y segunda partes del ADN diana; y correlacionar dicho conjunto de ubicaciones para determinar qué primera y segunda partes están unidas por dicho ADN diana y para determinar la secuencia de las moléculas de ADN diana.

En algunas realizaciones de los métodos y composiciones presentados en el presente documento, los complejos transposómicos están presentes en el soporte sólido a una densidad de al menos  $10^3$ ,  $10^4$ ,  $10^5$ ,  $10^6$  complejos por  $\text{mm}^2$ . En algunas realizaciones, el complejo transposómico comprende una transposasa hiperactiva, tal como la transposasa Tn5.

En algunas realizaciones de los métodos y composiciones presentados en el presente documento, el dominio de marcaje puede comprender, por ejemplo, una región para la amplificación por grupos. En algunas realizaciones, el dominio de marcaje puede comprender una región para cebar una reacción de secuenciación.

En algunas realizaciones de los métodos y composiciones presentados en el presente documento, el soporte sólido puede comprender, por ejemplo, micropartículas, una superficie con patrones, pocillos y similares. En algunas realizaciones, los complejos transposómicos se distribuyen aleatoriamente sobre el soporte sólido. En algunas realizaciones, los complejos transposómicos se distribuyen en una superficie con patrones.

Los detalles de una o más realizaciones se exponen en los dibujos adjuntos y la siguiente descripción. Otras características, objetos y ventajas serán evidentes a partir de la descripción y los dibujos y a partir de las reivindicaciones.

## Breve descripción de los dibujos

La figura 1a muestra el concepto general según una realización. La figura 1b muestra el concepto general según una realización.

La figura 2a es un esquema que ilustra una reacción de tagmentación.

La figura 2b es un esquema con el puente resultante redibujado para aclarar la naturaleza del puente resultante.

La figura 3a ilustra una realización donde dos formas de transposoma se ensamblan en la superficie de una cubeta de lectura. La adición de ADN a la cubeta de lectura da lugar a la tagmentación y el acoplamiento del ADN a los transposomas. También se muestran en la figura 3a diferentes tipos de grupos resultantes: grupos P7:P7, P5:P5 y P5:P7. La 3b ilustra una realización donde dos formas de transposoma se ensamblan en la superficie de una cubeta de lectura.

La figura 4a ilustra otra realización. En la figura 4a, solo está presente una forma de transposoma unido a la superficie (p. ej., el transposoma P5), lo que da como lugar a puentes que tienen la misma secuencia de marcaje en cada extremo. Se añaden transposomas adicionales para fragmentar aún más las estructuras de puente e incorporar una secuencia marcadora adicional (P7).

La figura 4b muestra una amplificación para generar pares de grupos resultantes de cada tocón de transposoma que representan dos fragmentos adyacentes en la muestra de ADN intacto original (figura 4b).

Las figuras 5a, 5b, 5c y 5d ilustran diferentes métodos para ensamblar complejos transposómicos unidos a la superficie. La figura 5a muestra una realización de este método para ensamblar los complejos transposómicos unidos a la superficie. La figura 5b muestra que los complejos transposómicos se ensamblan en solución y la inmovilización comprende una etapa adicional de ligar el primer polinucleótido a un oligonucleótido férula acoplado al soporte sólido. La figura 5c muestra un dímero de transposoma que se ensambla hibridando un oligonucleótido en bucle con un primer polinucleótido inmovilizado. La figura 5d muestra que los complejos transposómicos se pueden ensamblar en una cubeta de lectura estándar de extremos apareados con los cebadores de amplificación inmovilizados en la misma.

La figura 6a expone el diseño de un experimento expuesto como ejemplo 1. La figura 6b expone el diseño de un experimento expuesto como ejemplo 1.

La figura 7 expone datos representativos obtenidos en un experimento realizado según el diseño expuesto en el ejemplo 1.

La figura 8 expone datos representativos obtenidos en un experimento realizado según el diseño expuesto en el ejemplo 1.

La figura 9 es una ilustración del ensamblaje de transposomas unidos a perlas y la posterior tagmentación según una realización.

La figura 10 es una ilustración del ensamblaje de transposomas unidos a perlas y la posterior tagmentación según una realización.

La figura 11 es una ilustración del ensamblaje de transposomas unidos a perlas y la posterior tagmentación según una realización.

La figura 12 es una ilustración del ensamblaje de transposomas unidos a perlas y la posterior tagmentación según una realización.

La figura 13 es una ilustración de transposomas unidos a la superficie según una realización.

La figura 14 es una ilustración de la tagmentación realizada en transposomas unidos a perlas según una realización.

La figura 15 es una ilustración de la tagmentación y generación de código de barras realizada en transposomas unidos a perlas según una realización.

La figura 16 es una ilustración de la tagmentación y la generación de código de barras de subfragmentos realizadas en transposomas unidos a perlas según una realización.

La figura 17 es una ilustración del ensamblaje de transposomas como se describe en el ejemplo 3.

La figura 18a es una ilustración del ensamblaje de transposomas como se describe en el ejemplo 3.

La figura 18b expone los resultados usando el ensamblaje de transposomas como se describe en el ejemplo 3.

#### Descripción detallada

Los protocolos actuales para secuenciar muestras de ácidos nucleicos emplean rutinariamente un proceso de preparación de muestras que convierte el ADN o el ARN en una biblioteca de moldes. Estos métodos pueden provocar la pérdida de la muestra de ADN y, a menudo, requieren instrumentos costosos para la fragmentación. Además, los métodos de preparación de muestras suelen ser difíciles, tediosos e ineficaces.

En los métodos de preparación de muestras estándar, cada cadena molde contiene un adaptador en cualquiera de los extremos del inserto y normalmente son necesarias varias etapas tanto para modificar el ADN o el ARN como para purificar los productos deseados de las reacciones de modificación. Estas etapas se realizan en solución antes de la adición de los fragmentos adaptados a una celda de flujo, donde se acoplan a la superficie mediante una reacción de extensión de cebador que copia el fragmento hibridado al extremo de un cebador unido covalentemente a la superficie. Estas cadenas molde “semilla” dan lugar a continuación a grupos monoclonales de cadenas molde copiadas mediante varios ciclos de amplificación.

El número de etapas necesario para transformar el ADN en cadenas molde modificadas con adaptador en solución listas para la formación de grupos y la secuenciación puede minimizarse utilizando fragmentación y etiquetado mediado por transposasas. Este proceso, denominado en el presente documento “tagmentación” frecuentemente supone la modificación del ADN mediante un complejo transposómico que comprende la enzima transposasa en complejo con adaptadores que comprenden la secuencia terminal transposónica. El documento WO201206546 describe transposomas inmovilizados en perlas usados en la transposición en emulsión que incluyen transposasa libre.

La tagmentación da como resultado la fragmentación simultánea del ADN y el ligamiento de los adaptadores en los extremos 5' de ambas cadenas de los fragmentos dúplex. Tras una etapa de purificación para eliminar la enzima transposasa, se añaden secuencias adicionales a los extremos de los fragmentos adaptados mediante PCR.

La tagmentación en solución tiene inconvenientes y requiere varias etapas laboriosas. Además, se puede introducir un sesgo durante las etapas de amplificación por PCR. Los métodos y las composiciones presentados en el presente documento superan esos inconvenientes y permiten realizar una preparación sin sesgos de la muestra, formación de grupos y secuenciación en un único soporte sólido con requisitos mínimos para la manipulación o transferencia de la muestra.

La presente descripción se refiere al sorprendente descubrimiento de que los complejos transposómicos preacoplados a la superficie de una cubeta de lectura pueden fragmentar, marcar e inmovilizar eficazmente el ADN intacto dentro de la cubeta de lectura. En realizaciones específicas, una o más de las cadenas que comprenden los adaptadores transposómicos se unen a la superficie de la cubeta de lectura a través de su extremo 5'. Cuando se bombea ADN intacto a la cubeta de lectura, la reacción de tagmentación se produce de la misma manera que en las reacciones de tagmentación en solución, pero los fragmentos del producto resultante se adhieren físicamente a la superficie de la cubeta de lectura por sus extremos. Las secuencias adaptadoras transposómicas pueden contener secuencias que permiten la posterior generación y secuenciación de grupos.

Los métodos y composiciones presentados en el presente documento proporcionan varias ventajas frente a los métodos de tagmentación en solución. Por ejemplo, el molde de ADN intacto purificado, parcialmente purificado o incluso sin purificar se puede cargar directamente en una cubeta de lectura para la generación de grupos, sin preparación previa de la muestra. Además, la contigüidad de la información de secuencia en el ADN intacto original puede conservarse físicamente mediante la yuxtaposición de fragmentos tagmentados en la superficie de la cubeta de lectura. Como ventaja adicional, el ADN está unido físicamente a la superficie de la cubeta de lectura, por lo que se puede lograr la purificación de los reactivos después de una manipulación adicional del ADN mediante el intercambio de tampones de flujo continuo en el canal de la cubeta de lectura.

#### Tagmentación sobre un soporte sólido

Según lo anterior, en el presente documento se presentan métodos para preparar una biblioteca inmovilizada de fragmentos de ADN marcados. En algunas realizaciones, los métodos pueden comprender: (a) proporcionar un soporte sólido que tenga complejos transposómicos inmovilizados en el mismo, en donde los complejos transposómicos comprenden una transposasa unida a un primer polinucleótido, comprendiendo el primer polinucleótido (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) un primer marcador que comprende un primer dominio de marcaje; y (b) aplicar un ADN diana al soporte sólido en condiciones donde el ADN diana se fragmente mediante los complejos transposómicos y la secuencia terminal del transposón 3' del primer polinucleótido se transfiera a un extremo 5' de al menos una cadena de los fragmentos; produciendo de esta manera una biblioteca inmovilizada de fragmentos bicatenarios en donde al menos una cadena está marcada en 5' con el primer marcador.

Como se usa en el presente documento, la expresión “complejo transposómico” se refiere generalmente a una enzima transposasa unida de forma no covalente a un ácido nucleico bicatenario. Por ejemplo, el complejo puede ser una enzima transposasa preincubada con ADN de transposón bicatenario en condiciones que ayudan a la formación de complejos no covalentes. El ADN de transposón bicatenario puede incluir, sin limitación, ADN de Tn5, una parte del ADN de Tn5, una composición terminal de transposón, una mezcla de composiciones terminales de transposón u otros ADN bicatenarios capaces de interactuar con una transposasa tal como la transposasa Tn5 hiperactiva.

Una “transposasa” significa una enzima que es capaz de formar un complejo funcional con una composición que contiene los extremos de un transposón (p. ej., transposones, extremos de transposones, composiciones terminales de transposones) y catalizar la inserción o transposición de la composición que contiene los extremos del transposón en el ADN diana bicatenario con el que se incuba, por ejemplo, en una reacción de transposición *in vitro*. Una transposasa tal como se presenta en el presente documento también puede incluir integrasas de retrotransposones y retrovirus. Los expertos en la

técnica conocen generalmente las transposasas, los transposomas y los complejos transposómicos, como se ilustran en la descripción del documento US 2010/0120098. Aunque muchas realizaciones descritas en el presente documento se refieren a la transposasa Tn5 y/o a la transposasa Tn5 hiperactiva, se apreciará que en la presente invención se puede usar cualquier sistema de transposición que sea capaz de insertar un extremo de transposón con suficiente eficiencia para marcar en 5' y fragmentar un ADN diana para el propósito previsto. En casos particulares, un sistema de transposición preferido es capaz de insertar el extremo del transposón de forma aleatoria o casi aleatoria para marcar en 5' y fragmentar el ADN diana.

La expresión “extremo del transposón” se refiere a un ADN de ácido nucleico bicatenario que presenta solo las secuencias de nucleótidos (las “secuencias terminales del transposón”) que son necesarias para formar el complejo con la enzima transposasa o integrasa que es funcional en una reacción de transposición *in vitro*. En algunas realizaciones, el extremo de un transposón es capaz de formar un complejo funcional con la transposasa en una reacción de transposición. Como ejemplos no limitativos, los extremos del transposón pueden incluir el extremo exterior (“OE”) de 19 pb del transposón, el extremo interior (“IE”) del transposón o el extremo del transposón “extremo en mosaico” (“ME”) reconocido por una transposasa Tn5 mutante o natural, o el extremo del transposón R1 y R2 como se expone en la descripción del documento US 2010/0120098. Los extremos del transposón pueden comprender cualquier ácido nucleico o análogo de ácido nucleico adecuado para formar un complejo funcional con la enzima transposasa o integrasa en una reacción de transposición *in vitro*. Por ejemplo, el extremo del transposón puede comprender ADN, ARN, bases modificadas, bases no naturales, cadena principal modificada y puede comprender muescas en una o ambas cadenas. Aunque el término “ADN” se usa a lo largo de la presente descripción en relación con la composición de los extremos del transposón, debe entenderse que se puede utilizar cualquier ácido nucleico o análogo de ácido nucleico adecuado en un extremo del transposón.

La expresión “cadena transferida” se refiere a la parte transferida de ambos extremos del transposón. Similarmente, la expresión “cadena no transferida” se refiere a la parte no transferida de ambos “extremos del transposón”. El extremo 3' de una cadena transferida se une o se transfiere al ADN diana en una reacción de transposición *in vitro*. La cadena no transferida, que presenta una secuencia terminal del transposón que es complementaria de la secuencia terminal del transposón transferida, no se une ni se transfiere al ADN diana en una reacción de transposición *in vitro*.

En algunas realizaciones, la cadena transferida y la cadena no transferida se unen covalentemente. Por ejemplo, en algunas realizaciones, las secuencias de cadena transferidas y no transferidas se proporcionan en un único oligonucleótido, p. ej., en una configuración de horquilla. De este modo, aunque el extremo libre de la cadena no transferida no se une directamente al ADN diana mediante la reacción de transposición, la cadena no transferida se une al fragmento de ADN de forma indirecta, porque la cadena no transferida está unida a la cadena transferida mediante el bucle de la estructura en horquilla. Se pueden encontrar ejemplos adicionales de estructura de transposomas y métodos para preparar y usar transposomas en la descripción del documento US 2010/0120098.

Las expresiones “marcador” y “dominio de marcaje”, como se utilizan en la presente memoria, se refieren a una parte o un dominio de un polinucleótido que presenta una secuencia para un propósito o aplicación previstos deseados. Algunas realizaciones presentadas en el presente documento incluyen un complejo transposómico que comprende un polinucleótido que tiene una parte 3' que comprende una secuencia terminal del transposón y un marcador que comprende un dominio de marcaje. Los dominios de marcaje pueden comprender cualquier secuencia proporcionada para cualquier propósito deseado. Por ejemplo, en algunas realizaciones, un dominio de marcaje comprende uno o más sitios de reconocimiento de endonucleasas de restricción. En algunas realizaciones, un dominio de marcaje comprende una o más regiones adecuadas para la hibridación con un cebador para una reacción de amplificación por grupos. En algunas realizaciones, un dominio de marcaje comprende una o más regiones adecuadas para la hibridación con un cebador para una reacción de secuenciación. Se apreciará que se puede incorporar cualquier otra característica adecuada en un dominio de marcaje. En algunas realizaciones, el dominio de marcaje comprende una secuencia que tiene una longitud entre 5 y 200 pb. En algunas realizaciones, el dominio de marcaje comprende una secuencia que tiene una longitud entre 10 y 100 pb. En algunas realizaciones, el dominio de marcaje comprende una secuencia que tiene una longitud entre 20 y 50 pb. En algunas realizaciones, el dominio de marcaje comprende una secuencia que tiene una longitud entre 5, 6, 7, 8, 9, 10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 150 y 200 pb.

En los métodos y composiciones presentados en el presente documento, los complejos transposómicos se inmovilizan en el soporte sólido. En algunas realizaciones, los complejos transposómicos se inmovilizan en el soporte mediante uno o más polinucleótidos, tales como un polinucleótido que comprende una secuencia terminal del transposón. En algunas realizaciones, el complejo transposómico puede inmovilizarse mediante una molécula conectora que acopla la enzima transposasa al soporte sólido. En algunas realizaciones, tanto la enzima transposasa como el polinucleótido están inmovilizados en el soporte sólido. Cuando se hace referencia a la inmovilización de moléculas (p. ej., ácidos nucleicos) a un soporte sólido, los términos “inmovilizado” y “unido” se usan indistintamente en el presente documento y se pretende que ambos términos abarquen la unión directa o indirecta, covalente o no covalente, salvo que se indique lo contrario, ya sea de forma explícita o por el contexto. En determinadas realizaciones de la invención, se puede preferir una unión covalente, pero generalmente todo lo que se requiere es que las moléculas (p. ej., ácidos nucleicos) permanezcan inmovilizadas o unidas al soporte en las condiciones en las que se pretende usar el soporte, por ejemplo, en aplicaciones que requieren la amplificación y/o secuenciación de ácidos nucleicos.

Determinadas realizaciones de la invención pueden utilizar soportes sólidos compuestos por un sustrato o matriz inerte (p. ej., portaobjetos de vidrio, perlas de polímero, etc.) que se han “funcionalizado”, por ejemplo, mediante la aplicación de una

capa o recubrimiento de un material intermedio que comprende grupos reactivos que permiten la unión covalente a biomoléculas, tales como polinucleótidos. Los ejemplos de dichos soportes incluyen, pero sin limitación, hidrogeles de poli(acrilamida) sostenidos sobre un sustrato inerte tal como vidrio, particularmente hidrogeles de poli(acrilamida) como se describe en los documentos WO 2005/065814 y US 2008/0280773. En dichas realizaciones, las biomoléculas (p. ej., polinucleótidos) pueden estar unidas directamente de manera covalente al material intermedio (p. ej., al hidrogel), pero el propio material intermedio puede estar unido de manera no covalente al sustrato o matriz (p. ej., al sustrato de vidrio). Por consiguiente, la expresión “unión covalente a un soporte sólido” debe interpretarse como que abarca este tipo de disposición.

Las expresiones “superficie sólida”, “soporte sólido” y otros equivalentes gramaticales en la presente descripción se refieren a cualquier material que sea adecuado o pueda modificarse de modo que sea adecuado para la unión de los complejos transposómicos. Como apreciarán los expertos en la técnica, la cantidad de sustratos posibles es muy grande. Posibles sustratos incluyen, pero sin limitarse a, vidrio y vidrio modificado o funcionalizado, plásticos (incluidos acrílicos, poliestireno y copolímeros de estireno y otros materiales, polipropileno, polietileno, polibutileno, poliuretanos, Teflon™, etc.), polisacáridos, nailon o nitrocelulosa, cerámicas, resinas, sílice o materiales basados en sílice que incluyen silicio y silicio modificado, carbono, metales, vidrios inorgánicos, plásticos, haces de fibras ópticas y una variedad de otros polímeros. Los soportes sólidos y superficies sólidas especialmente útiles para algunas realizaciones están situados dentro de un aparato de celda de flujo. A continuación se exponen celdas de flujo ilustrativas con mayor detalle.

En algunas realizaciones, el soporte sólido comprende una superficie con patrones adecuada para la inmovilización de complejos transposómicos en un patrón ordenado. Una “superficie con patrones” se refiere a una disposición de diferentes regiones en o sobre una capa expuesta de un soporte sólido. Por ejemplo, una o más de las regiones pueden ser características donde haya uno o más complejos transposómicos. Las características pueden estar separadas por regiones intersticiales donde no haya complejos transposómicos. En algunas realizaciones, el patrón puede ser un formato de casillas x-y que estén en filas y columnas. En algunas realizaciones, el patrón puede ser una disposición repetitiva de casillas y/o regiones intersticiales. En algunas realizaciones, el patrón puede ser una disposición aleatoria de casillas y/o regiones intersticiales. En algunas realizaciones, los complejos transposómicos se distribuyen aleatoriamente sobre el soporte sólido. En algunas realizaciones, los complejos transposómicos se distribuyen en una superficie con patrones. Se describen superficies con patrones ilustrativas que pueden utilizarse en los métodos y composiciones expuestos en la presente memoria en la publicación de solicitud de patente estadounidense n.º 2013/0116153 o la publicación de solicitud de patente estadounidense n.º 2012/0316086 A1.

En algunas realizaciones, el soporte sólido comprende una matriz de pocillos o surcos en una superficie. Este puede fabricarse, como se sabe generalmente en la técnica, utilizando una variedad de técnicas, incluyendo, pero sin limitación, fotolitografía, técnicas de estampado, técnicas de moldeo y técnicas de micrograbado. Como apreciarán los expertos en la técnica, la técnica utilizada dependerá de la composición y forma del sustrato de la matriz.

La composición y la geometría del soporte sólido pueden variar con su uso. En algunas realizaciones, el soporte sólido es una estructura plana, tal como un portaobjetos, un chip, un microchip y/o una matriz. Como tal, la superficie de un sustrato puede tener la forma de una capa plana. En algunas realizaciones, el soporte sólido comprende una o más superficies de una cubeta de lectura. El término “cubeta de lectura”, como se utiliza en la presente memoria, se refiere a una cámara que comprende una superficie sólida a través de la cual pueden fluir uno o más reactivos líquidos. Se describen ejemplos de cubetas de lectura y de sistemas fluidos y plataformas de detección relacionados que pueden utilizarse fácilmente en los métodos de la presente descripción, por ejemplo, en Bentley y col., *Nature* 456:53-59 (2008), documentos WO 04/018497; US-7.057.026; WO 91/06678; WO 07/123744; US-7.329.492; US-7.211.414; US-7.315.019; US-7.405.281, y US 2008/0108082.

En algunas realizaciones, el soporte sólido o su superficie no forma un plano, tal como la superficie interna o externa de un tubo o un recipiente. En algunas realizaciones, el soporte sólido comprende microesferas o perlas. Por “microesferas” o “perlas” o “partículas” o equivalentes gramaticales se entiende en la presente memoria partículas discretas pequeñas. Las composiciones de perlas adecuadas incluyen, pero sin limitación, plástico, cerámica, vidrio, poliestireno, metilmetilacrilato, polímeros acrílicos, materiales paramagnéticos, sol de toria, grafito de carbono, dióxido de titanio, látex o dextranos reticulados, tales como Sepharose, celulosa, nailon, micelas reticuladas y teflón, así como cualquier otro material indicado en la presente descripción para soportes sólidos. La “Microsphere Detection Guide” de Bangs Laboratories, Fishers Ind. es una guía útil. En determinadas realizaciones, las microesferas son microesferas o perlas magnéticas.

Las perlas no tienen por qué ser esféricas; pueden utilizarse partículas irregulares. De forma alternativa o adicional, las perlas pueden ser porosas. Los tamaños de las perlas varían de nanómetros, es decir, 100 nm, a milímetros, es decir, 1 mm, prefiriéndose las perlas de aproximadamente 0,2 micrómetros a aproximadamente 200 micrómetros, prefiriéndose particularmente las de aproximadamente 0,5 a aproximadamente 5 micrómetros, aunque en algunas realizaciones pueden utilizarse perlas más pequeñas o más grandes.

Las figuras 1a y 1b ilustran generalmente el método según una realización. Se muestra que un soporte sólido recubierto con oligonucleótidos injertados, algunos de los cuales contienen las secuencias ME, formará complejos transposómicos activos que se acoplan físicamente al soporte sólido en presencia de Tn5. La densidad de estos transposomas unidos a la superficie se puede modular variando la densidad de los oligonucleótidos injertados que contienen la secuencia ME o según la cantidad

de transposasa añadida al soporte sólido. Por ejemplo, en algunas realizaciones, los complejos transposómicos están presentes en el soporte sólido a una densidad de al menos  $10^3$ ,  $10^4$ ,  $10^5$  o al menos  $10^6$  complejos por  $\text{mm}^2$ .

Cuando se añade ADN bicatenario al soporte sólido, los complejos transposómicos tagmentarán el ADN añadido, generando por lo tanto fragmentos bc acoplados en ambos extremos a la superficie. En algunas realizaciones, la longitud de los fragmentos con puente se puede variar cambiando la densidad de los complejos transposómicos en la superficie. En determinadas realizaciones, la longitud de los fragmentos con puente resultantes es inferior a 100 pb, 200 pb, 300 pb, 400 pb, 500 pb, 600 pb, 700 pb, 800 pb, 900 pb, 1000 pb, 1100 pb, 1200 pb, 1300 pb, 1400 pb, 1500 pb, 1600 pb, 1700 pb, 1800 pb, 1900 pb, 2000 pb, 2100 pb, 2200 pb, 2300 pb, 2400 pb, 2500 pb, 2600 pb, 2700 pb, 2800 pb, 2900 pb, 3000 pb, 3100 pb, 3200 pb, 3300 pb, 3400 pb, 3500 pb, 3600 pb, 3700 pb, 3800 pb, 3900 pb, 4000 pb, 4100 pb, 4200 pb, 4300 pb, 4400 pb, 4500 pb, 4600 pb, 4700 pb, 4800 pb, 4900 pb, 5000 pb, 10.000 pb, 30.000 pb o menos de 100.000 pb. En tales realizaciones, los fragmentos con puente pueden amplificarse a continuación en grupos usando la química de grupos estándar, como se ilustra en la descripción de las patentes US-7.985.565 y US-7.115.400.

En algunas realizaciones, la longitud de los moldes es mayor que la que se puede amplificar adecuadamente usando la química de grupos estándar. Por ejemplo, en algunas realizaciones, la longitud de los moldes es superior a 100 pb, 200 pb, 300 pb, 400 pb, 500 pb, 600 pb, 700 pb, 800 pb, 900 pb, 1000 pb, 1100 pb, 1200 pb, 1300 pb, 1400 pb, 1500 pb, 1600 pb, 1700 pb, 1800 pb, 1900 pb, 2000 pb, 2100 pb, 2200 pb, 2300 pb, 2400 pb, 2500 pb, 2600 pb, 2700 pb, 2800 pb, 2900 pb, 3000 pb, 3100 pb, 3200 pb, 3300 pb, 3400 pb, 3500 pb, 3600 pb, 3700 pb, 3800 pb, 3900 pb, 4000 pb, 4100 pb, 4200 pb, 4300 pb, 4400 pb, 4500 pb, 4600 pb, 4700 pb, 4800 pb, 4900 pb, 5000 pb, 10.000 pb, 30.000 pb o superior a 100.000 pb. En tales realizaciones, se puede realizar a continuación una segunda reacción de tagmentación añadiendo transposomas de la solución que fragmentan adicionalmente los puentes, como se ilustra, por ejemplo, en la figura 4a. Por lo tanto, la segunda reacción de tagmentación puede eliminar el tramo interno de los puentes, dejando tocones cortos anclados a la superficie que pueden convertirse en grupos listos para posteriores etapas de secuenciación. En realizaciones particulares, la longitud del molde puede estar dentro de un intervalo definido por un límite superior e inferior seleccionado entre los ilustrados anteriormente.

En determinadas realizaciones, antes de la generación de grupos, se pueden obtener imágenes del ADN inmovilizado por la tagmentación en superficie. Por ejemplo, el ADN inmovilizado se puede tefir con un colorante interquelante y se pueden obtener imágenes para conservar un registro de la posición de la cadena principal de la molécula de ADN en la superficie. Tras la generación y secuenciación de los grupos, las coordenadas de los grupos se pueden asociar con su posición en la columna vertebral original, lo que ayuda a alinear las lecturas a lo largo del ensamblaje de una molécula y un genoma.

En algunas realizaciones, la etapa de aplicar un ADN diana comprende añadir una muestra biológica a dicho soporte sólido. La muestra biológica puede ser de cualquier tipo que comprenda ADN y que pueda depositarse sobre la superficie sólida para su tagmentación. Por ejemplo, la muestra puede comprender ADN en diversos estados de purificación, incluido el ADN purificado. Sin embargo, no es necesario que la muestra se purifique por completo y puede comprender, por ejemplo, ADN mezclado con proteínas, otras especies de ácidos nucleicos, otros componentes celulares y/o cualquier otro contaminante. Como se demuestra en el ejemplo 2 siguiente, en algunas realizaciones, la muestra biológica comprende una mezcla de ADN, proteínas, otras especies de ácidos nucleicos, otros componentes celulares y/o cualquier otro contaminante presente en aproximadamente la misma proporción que la encontrada *in vivo*. Por ejemplo, en algunas realizaciones, los componentes se encuentran en la misma proporción que en una célula intacta. En algunas realizaciones, la muestra biológica tiene una relación 260/280 inferior a 2,0, 1,9, 1,8, 1,7, 1,6, 1,5, 1,4, 1,3, 1,2, 1,1, 1,0, 0,9, 0,8, 0,7 o inferior a 0,60. En algunas realizaciones, la muestra biológica tiene una relación 260/280 de al menos 2,0, 1,9, 1,8, 1,7, 1,6, 1,5, 1,4, 1,3, 1,2, 1,1, 1,0, 0,9, 0,8, 0,7 o al menos 0,60. Debido a que los métodos proporcionados en el presente documento permiten que el ADN se una a un soporte sólido, otros contaminantes pueden eliminarse simplemente lavando el soporte sólido después de que se produzca la tagmentación unida a la superficie. La muestra biológica puede comprender, por ejemplo, un lisado celular bruto o células completas. Por ejemplo, no es necesario que un lisado celular bruto que se aplica a un soporte sólido en un método expuesto en el presente documento se haya sometido a una o más de las etapas de separación que se usan tradicionalmente para aislar los ácidos nucleicos de otros componentes celulares. Se exponen etapas de separación ilustrativas en Maniatis y col., Molecular Cloning: A Laboratory Manual, 2.<sup>a</sup> Edición, 1989, y Short Protocols in Molecular Biology, ed. Ausubel, y col.

Por lo tanto, en algunas realizaciones, la muestra biológica puede comprender, por ejemplo, sangre, plasma, suero, linfa, moco, esputo, orina, semen, líquido cefalorraquídeo, aspirado bronquial, heces y tejido macerado, o un lisado de los mismos, o cualquier otra muestra biológica que comprenda ADN. Una ventaja de los métodos y composiciones presentados en el presente documento es que se puede añadir una muestra biológica a la cubeta de lectura y las etapas posteriores de lisis y purificación pueden tener lugar en la cubeta de lectura sin etapas adicionales de transferencia o manipulación, simplemente haciendo fluir los reactivos necesarios a la cubeta de lectura. Los ejemplos 1 y 2 siguientes demuestran la aplicación exitosa de lisados celulares brutos a los métodos y composiciones proporcionados en el presente documento.

Las figuras 2a y 2b ilustran además la reacción de tagmentación según una realización. Como se muestra en la figura 2a, los transposomas comprenden un dímero de Tn5 y cada monómero se une a una molécula bicatenaria, es decir, el adaptador ME está unida covalentemente a la superficie de una cubeta de lectura. Los transposomas se unen al ADN diana y generan dos muescas en la cadena principal del ADN, a una distancia de 9 bases en cada cadena. Aunque la figura 2a muestra un hueco de 9 pb entre las muescas, en otras



realizaciones, el transposoma puede generar un hueco de 7, 8, 9, 10, 11 o 12 pb entre las muescas. Solo una de las dos cadenas de cada adaptador ME está ligada a la cadena 5' en cada posición de muesca. Esta cadena, "la cadena transferida", se injerta en la superficie de la cubeta de lectura a través de su extremo 5'. El puente resultante de la tagmentación en superficie se vuelve a dibujar en la figura 2b para aclarar la naturaleza del puente resultante.

Las figuras 3a y 3b ilustran un ejemplo de la invención puesta en práctica. Se ensamblan dos formas de transposoma en la superficie de una cubeta de lectura. La primera forma comprende un transposoma P7 en el que la "cadena de transferencia" del adaptador ME comprende una secuencia oligonucleotídica extendida que une el transposoma a la superficie que incluye un dominio de amplificación (P7) y un cebador de secuenciación (S2). La segunda forma comprende un transposoma P5 en el que la "cadena de transferencia" del adaptador ME comprende una secuencia oligonucleotídica extendida que une el transposoma a la superficie que incluye un dominio de amplificación (P5) y un cebador de secuenciación (S1). La adición de ADN a la cubeta de lectura da lugar a la tagmentación y el acoplamiento del ADN a los transposomas. El resultado son tres tipos de puentes: P5-P5, P7-P7 y P5-P7. Tras la formación y linealización de los grupos (figura 3b), se eliminan los grupos P5-P5 o P7-P7 (dependiendo de la linealización). Como se muestra en la figura 3b, si la linealización ocurre a través de P5, solo quedan los puentes P5-P7 y P7-P7. Los grupos P7-P7 no se linealizan mediante esta reacción y, por lo tanto, no se secuenciarán. Los grupos P7-P7 se eliminan posteriormente durante la linealización de la lectura 2 para la secuenciación.

El método presentado en el presente documento puede comprender además una etapa adicional de proporcionar complejos transposómicos en solución y poner en contacto los complejos transposómicos en fase de solución con los fragmentos inmovilizados en condiciones donde el ADN diana sea fragmentado por los complejos transposómicos en solución; obteniendo de esta manera fragmentos de ácido nucleico inmovilizados que tienen un extremo en solución. En algunas realizaciones, los complejos transposómicos en solución pueden comprender un segundo marcador, de modo que el método genere fragmentos de ácido nucleico inmovilizados que tienen un segundo marcador, el segundo marcador en solución. El primer y el segundo marcadores pueden ser diferentes o iguales.

En algunas realizaciones, una forma de transposoma unido a la superficie está presente predominantemente en el soporte sólido. Por ejemplo, en algunas realizaciones, al menos el 50 %, el 55 %, el 60 %, el 65 %, el 70 %, el 75 %, el 80 %, el 85 %, el 90 %, el 95 %, el 96 %, el 97 %, el 98 % o al menos el 99 % de los marcadores presentes en dicho soporte sólido comprenden el mismo dominio de marcaje. En tales realizaciones, después de una reacción de tagmentación inicial con transposomas unidos a la superficie, al menos el 50 %, el 55 %, el 60 %, el 65 %, el 70 %, el 75 %, el 80 %, el 85 %, el 90 %, el 95 %, el 96 %, el 97 %, el 98 % o al menos el 99 % de las estructuras de puente comprenden el mismo dominio de marcaje en cada extremo del puente. Se puede realizar una segunda reacción de tagmentación añadiendo transposomas de la solución que fragmentan adicionalmente los puentes. En algunas realizaciones, la mayoría o todos los transposomas en fase de solución comprenden un dominio de marcaje que difiere del dominio de marcaje presente en las estructuras puente generadas en la primera reacción de tagmentación. Por ejemplo, en algunas realizaciones, al menos el 50 %, el 55 %, el 60 %, el 65 %, el 70 %, el 75 %, el 80 %, el 85 %, el 90 %, el 95 %, el 96 %, el 97 %, el 98 % o al menos el 99 % de los marcadores presentes en los transposomas en fase de solución comprenden un dominio de marcaje que difiere del dominio de marcaje presente en las estructuras de puente generadas en la primera reacción de tagmentación.

Las figuras 4a y 4b ilustran esta realización. El soporte sólido mostrado en la figura 4a comprende solo un transposoma unido a la superficie que comprende un único tipo de secuencia marcadora (p. ej., el transposoma P5). En este caso, todos los puentes tagmentados son puentes P5-P5. Se añade un transposoma P7 de la solución y los tagmentos a los puentes, convirtiendo todas las moléculas unidas a la superficie en moldes P5-P7. Los fragmentos unidos a la superficie se pueden convertir posteriormente en grupos (véase, por ejemplo, la figura 4b). Además, de cada tocón de transposoma se generará un par de grupos que representan dos fragmentos adyacentes en la muestra de ADN intacto original (figura 4b).

También se presentan en el presente documento soportes sólidos que tienen una biblioteca de fragmentos de ADN marcados inmovilizados en los mismos preparados según los métodos anteriores.

#### Mapas físicos de moléculas polinucleotídicas inmovilizadas

En la presente descripción también se presentan métodos para generar un mapa físico de polinucleótidos inmovilizados. Los métodos se pueden aprovechar de forma ventajosa para identificar grupos que probablemente contengan secuencias enlazadas (es decir, la primera y la segunda partes de la misma molécula polinucleotídica diana). La proximidad relativa de dos grupos que resultan de un polinucleótido inmovilizado proporciona, por lo tanto, información útil para la alineación de la información de secuencia obtenida de los dos grupos. Específicamente, la distancia entre dos grupos determinados en una superficie sólida se correlaciona positivamente con la probabilidad de que las dos agrupaciones provengan de la misma molécula polinucleotídica diana, como se describe con mayor detalle en el documento WO 2012/025250.

Como ejemplo, en algunas realizaciones, las moléculas largas de ADNbc que se extienden sobre la superficie de una cubeta de lectura se tagmentan en el sitio, lo que da lugar a una línea de puentes de ADNbc conectados a lo largo de la superficie de la cubeta de lectura. Además, puede generarse a continuación un mapa físico del ADN inmovilizado. El mapa físico correlaciona así la relación física de los grupos después de amplificarse el ADN inmovilizado. Específicamente, el mapa físico se usa para calcular la probabilidad de que los datos de secuencia obtenidos de dos grupos cualesquiera estén ligados, como se describe en el documento WO 2012/025250.

En algunas realizaciones, el mapa físico se genera mediante la obtención de imágenes del ADN para establecer la ubicación de las moléculas de ADN inmovilizadas en una superficie sólida. En algunas realizaciones, se obtienen imágenes del ADN inmovilizado añadiendo un agente de obtención de imágenes al soporte sólido y detectando una señal del agente de obtención de imágenes. En algunas realizaciones, el agente de obtención de imágenes es un marcador detectable. Los marcadores adecuados incluyen, pero sin limitación, protones, haptenos, radionucleótidos, enzimas, marcadores fluorescentes, marcadores quimioluminiscentes y agentes cromógenos. Por ejemplo, en algunas realizaciones, el agente de obtención de imágenes es un colorante intercalante o un agente de unión al ADN no intercalante. Se puede usar cualquier colorante intercalante o agente de unión al ADN no intercalante adecuado como se conoce en la técnica, incluidos, pero sin limitación, los expuestos en el documento US 2012/0282617.

En algunas realizaciones, los fragmentos bicatenarios inmovilizados se fragmentan aún más para liberar un extremo libre (véase la figura 4a) antes de la generación del grupo (figura 4b). La escisión de estructuras con puente se puede realizar usando cualquier metodología adecuada conocida en la técnica, como se ilustra en el documento WO 2012/025250. Por ejemplo, la escisión puede producirse mediante la incorporación de un nucleótido modificado, tal como el uracilo, como se describe en el documento WO 2012/025250, mediante la incorporación de un sitio de endonucleasa de restricción o mediante la aplicación de complejos transposómicos en fase de solución a las estructuras de ADN con puente, como se describe en otra parte de la presente descripción.

En determinadas realizaciones, una pluralidad de moléculas de ADN diana fluyen hacia una cubeta de lectura que comprende una pluralidad de nanocanales, teniendo el nanocanal una pluralidad de complejos transposómicos inmovilizados en el mismo. Como se utiliza en la presente memoria, el término nanocanal se refiere a un canal estrecho en el que fluye una molécula de ADN lineal larga. En algunas realizaciones, no fluyen más de 1, 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 30, 40, 50, 60, 70, 80, 90, 100, 200, 300, 400, 500, 600, 700, 800, 900 o no más de 1000 cadenas largas individuales de ADN diana en cada nanocanal. En algunas realizaciones, los nanocanales individuales están separados por una barrera física que evita que las cadenas largas individuales del ADN diana interactúen con múltiples nanocanales. En algunas realizaciones, el soporte sólido comprende al menos 10, 50, 100, 200, 500, 1000, 3000, 5000, 10.000, 30.000, 50.000, 80.000 o al menos 100.000 nanocanales. En algunas realizaciones, los transposomas unidos a la superficie de un nanocanal tagmentan el ADN. A continuación, se puede realizar una cartografía de contigüidad, por ejemplo, siguiendo los grupos a lo largo de uno de estos canales. En algunas realizaciones, la cadena larga del ADN diana puede tener al menos 0,1 kb, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 15 kb, 20 kb, 25 kb, 30 kb, 35 kb, 40 kb, 45 kb, 50 kb, 55 kb, 60 kb, 65 kb, 70 kb, 75 kb, 80 kb, 85 kb, 90 kb, 95 kb, 100 kb, 150 kb, 200 kb, 250 kb, 300 kb, 350 kb, 400 kb, 450 kb, 500 kb, 550 kb, 600 kb, 650 kb, 700 kb, 750 kb, 800 kb, 850 kb, 900 kb, 950 kb, 1000 kb, 5000 kb, 10.000 kb, 20.000 kb, 30.000 kb o al menos 50.000 kb de longitud. En algunas realizaciones, la cadena larga del ADN diana no tiene más de 0,1 kb, 1 kb, 2 kb, 3 kb, 4 kb, 5 kb, 6 kb, 7 kb, 8 kb, 9 kb, 10 kb, 15 kb, 20 kb, 25 kb, 30 kb, 35 kb, 40 kb, 45 kb, 50 kb, 55 kb, 60 kb, 65 kb, 70 kb, 75 kb, 80 kb, 85 kb, 90 kb, 95 kb, 100 kb, 150 kb, 200 kb, 250 kb, 300 kb, 350 kb, 400 kb, 450 kb, 500 kb, 550 kb, 600 kb, 650 kb, 700 kb, 750 kb, 800 kb, 850 kb, 900 kb, 950 kb o no más de 1000 kb de longitud. Como ejemplo, se puede usar una cubeta de lectura que tenga 1000 o más nanocanales con productos de tagmentación inmovilizados cartografiados en los nanocanales para secuenciar el genoma de un organismo con lecturas cortas "colocadas". En algunas realizaciones, los productos de tagmentación inmovilizados cartografiados en los nanocanales pueden usarse para resolver haplotipos. En algunas realizaciones, los productos de tagmentación inmovilizados cartografiados en los nanocanales pueden usarse para resolver problemas de fase.

#### Amplificación y secuenciación de fragmentos de ADN inmovilizados

Amplificación. La presente descripción se refiere además a la amplificación de los fragmentos de ADN inmovilizados producidos según los métodos proporcionados en la presente descripción. Los fragmentos de ADN inmovilizados producidos por la tagmentación mediada por transposomas unidos a la superficie se pueden amplificar según cualquier metodología de amplificación adecuada conocida en la técnica. En algunas realizaciones, los fragmentos de ADN inmovilizados se amplifican sobre un soporte sólido. En algunas realizaciones, el soporte sólido es el mismo soporte sólido sobre el que se produce la tagmentación unida a la superficie. En tales realizaciones, los métodos y composiciones proporcionados en el presente documento permiten que la preparación de la muestra continúe sobre el mismo soporte sólido desde la etapa inicial de introducción de la muestra hasta la amplificación y, opcionalmente, hasta una etapa de secuenciación.

Por ejemplo, en algunas realizaciones, los fragmentos de ADN inmovilizados se amplifican usando metodologías de amplificación por grupos, como se ejemplifica en las descripciones de las patentes US-7.985.565 y US-7.115.400. Las patentes US-7.985.565 y US-7.115.400 describen métodos de amplificación de ácido nucleico en fase sólida, que permiten que los productos de amplificación se inmovilicen sobre un soporte sólido para formar matrices compuestas por grupos o "colonias" de moléculas de ácido nucleico inmovilizadas. Cada grupo o colonia en dicha matriz se forma a partir de una pluralidad de cadenas de polinucleótidos idénticas inmovilizadas y una pluralidad de cadenas de polinucleótidos complementarias inmovilizadas idénticas. Las matrices así formadas se denominan generalmente en el presente documento "matrices agrupadas". Los productos de reacciones de amplificación en fase sólida tales como los descritos en las patentes US-7.985.565 y US-7.115.400 se denominan estructuras "con puente" formadas por la hibridación de pares de cadenas polinucleotídicas inmovilizadas y cadenas complementarias inmovilizadas, estando ambas cadenas inmovilizadas en el soporte sólido en el extremo 5', preferiblemente mediante una unión covalente. Las metodologías de amplificación de grupos son ejemplos de métodos en donde para producir amplicones inmovilizados se utiliza un molde de ácido nucleico

inmovilizado. También pueden utilizarse otras metodologías adecuadas para producir amplicones inmovilizados a partir de fragmentos de ADN inmovilizados producidos según los métodos proporcionados en el presente documento. Por ejemplo, pueden formarse uno o más grupos o colonias mediante PCR en fase sólida si uno o ambos cebadores de cada par de cebadores de amplificación se inmovilizan.

En otras realizaciones, los fragmentos de ADN inmovilizados se amplifican en solución. Por ejemplo, en algunas realizaciones, los fragmentos de ADN inmovilizados se escinden o se liberan de otro modo del soporte sólido y los cebadores de amplificación se hibridan a continuación en solución con las moléculas liberadas. En otras realizaciones, los cebadores de amplificación se hibridan con los fragmentos de ADN inmovilizados para una o más etapas de amplificación iniciales, seguidas de etapas de amplificación posteriores en solución. Por lo tanto, en algunas realizaciones, se puede usar un molde de ácido nucleico inmovilizado para producir amplicones en fase de solución.

Se apreciará que cualquiera de las metodologías de amplificación descritas en la presente descripción o generalmente conocidas en la técnica pueden utilizarse con cebadores universales o específicos de la diana para amplificar fragmentos de ADN inmovilizados. Como métodos de amplificación adecuados se incluyen, pero sin limitación, la reacción en cadena de la polimerasa (PCR, *polymerase chain reaction*), la amplificación por desplazamiento de cadena (SDA, *strand displacement amplification*), la amplificación mediada por transcripción (TMA, *transcription mediated amplification*) y la amplificación basada en secuencias de ácidos nucleicos (NASBA, *nucleic acid sequence based amplification*), como se describe en la patente US-8.003.354. Los métodos de amplificación anteriores pueden emplearse para amplificar uno o más ácidos nucleicos de interés. Por ejemplo, la PCR, que incluye PCR múltiple, SDA, TMA, NASBA y similares, puede utilizarse para amplificar fragmentos de ADN inmovilizados. En algunas realizaciones, en la reacción de amplificación se incluyen cebadores dirigidos específicamente al ácido nucleico de interés.

Otros métodos adecuados para la amplificación de ácidos nucleicos pueden incluir tecnologías de extensión y ligamiento de oligonucleótidos, amplificación por círculo rodante (RCA) (Lizardi y col., Nat. Genet. 19:225-232 (1998)) y ensayo de ligamiento de oligonucleótidos (OLA) (véase en general las patentes US-7.582.420, US-5.185.243, US-5.679.524 y US-5.573.907; los documentos EP 0 320 308 B1; EP 0 336 731 B1; EP 0 439 182 B1; WO 90/01069; WO 89/12696; y WO 89/09835). Se apreciará que estas metodologías de amplificación se pueden diseñar para amplificar fragmentos de ADN inmovilizados. Por ejemplo, en algunas realizaciones, el método de amplificación puede incluir amplificación de sonda de ligamiento o reacciones de ensayo de ligamiento de oligonucleótidos (OLA) que contienen cebadores dirigidos específicamente al ácido nucleico de interés. En algunas realizaciones, el método de amplificación puede incluir una reacción de extensión de cebador por ligamiento que contiene cebadores dirigidos específicamente al ácido nucleico de interés. Como ejemplo no limitativo de extensión de cebadores y cebadores de ligamiento que pueden diseñarse específicamente para amplificar un ácido nucleico de interés, la amplificación puede incluir cebadores utilizados para el ensayo GoldenGate (Illumina, Inc., San Diego, CA) como se ilustra en las patentes US-7.582.420 y US-7.611.869.

Como ejemplos de métodos de amplificación isotérmica que pueden utilizarse en un método de la presente descripción se incluyen, pero sin limitación, amplificación de desplazamiento múltiple (MDA) como se ilustra, por ejemplo, en Dean y col., Proc. Natl. Acad. Sci. EE. UU. 99:5261-66 (2002) o amplificación isotérmica de ácido nucleico por desplazamiento de cadena ilustrada, por ejemplo, en la patente US-6.214.587. Otros métodos no basados en PCR que pueden utilizarse en la presente divulgación incluyen, por ejemplo, amplificación por desplazamiento de cadena (SDA, *strand displacement amplification*) que se describe, por ejemplo, en Walker y col., Molecular Methods for Virus Detection, Academic Press, Inc., 1995; patentes US-5.455.166 y US-5.130.238, y en Walker y col., Nucl. Acids Res. 20:1691-96 (1992) o amplificación por desplazamiento de cadena hiperramificada que se describe, por ejemplo, en Lage y col., Genome Research 13:294-307 (2003). Para la amplificación aleatoria de cebadores de ADN genómico, pueden utilizarse métodos de amplificación isotérmica con la polimerasa Phi 29 de desplazamiento de cadena o el fragmento grande de la ADN polimerasa Bst, *exo*<sup>5'→3'</sup>. El uso de estas polimerasas aprovecha su alta procesividad y actividad de desplazamiento de cadenas. La alta procesividad permite que las polimerasas produzcan fragmentos que tienen una longitud de 10-20 kb. Como se ha expuesto anteriormente, pueden producirse fragmentos más pequeños en condiciones isotérmicas utilizando polimerasas que tienen baja procesividad y actividad de desplazamiento de cadena tal como la polimerasa Klenow. Se expone una descripción adicional en detalle de las reacciones, condiciones y componentes de amplificación en la descripción de la patente US-7.670.810.

Otro método de amplificación de ácidos nucleicos que es útil en la presente descripción es la PCR con marcaje que utiliza una población de cebadores de dos dominios que tienen una región 5' constante seguida por una región 3' aleatoria como se describe, por ejemplo, en Grothues y col. Nucleic Acids Res. 21(5):1321-2 (1993). Las primeras rondas de amplificación se llevan a cabo para permitir una multitud de iniciaciones en ADN desnaturalizado por calor basadas en hibridación individual de la región 3' sintetizada al azar. Debido a la naturaleza de la región 3', se contempla que los sitios de iniciación sean aleatorios en todo el genoma. Posteriormente, los cebadores no unidos pueden retirarse y puede producirse una replicación adicional utilizando cebadores complementarios a la región 5' constante.

Secuenciación. La presente descripción se refiere además a la secuenciación de los fragmentos de ADN inmovilizados producidos según los métodos proporcionados en la presente descripción. Los fragmentos de ADN inmovilizados producidos por tagmentación mediada por transposoma unido a la superficie pueden secuenciarse según cualquier metodología de secuenciación adecuada, tal como secuenciación directa, incluyendo secuenciación por síntesis, secuenciación por ligamiento, secuenciación por hibridación, secuenciación de nanoporos y similares. En algunas realizaciones, los fragmentos

de ADN inmovilizados se secuencian sobre un soporte sólido. En algunas realizaciones, el soporte sólido para la secuenciación es el mismo soporte sólido sobre el que se produce la tagmentación unida a la superficie. En algunas realizaciones, el soporte sólido para la secuenciación es el mismo soporte sólido sobre el que se produce la amplificación.

Una metodología de secuenciación preferida es la secuenciación por síntesis (SBS, *sequencing-by-synthesis*). En la SBS, para determinar la secuencia de nucleótidos en el molde, se controla la extensión de un cebador de ácido nucleico a lo largo de un molde de ácido nucleico (p. ej., un ácido nucleico o amplicón del mismo). El proceso químico subyacente puede ser la polimerización (p. ej., catalizada por una enzima polimerasa). En una realización particular de SBS basada en polimerasa, se añaden nucleótidos etiquetados con fluorescencia a un cebador (extendiéndose de ese modo el cebador) de una manera dependiente del molde, de modo que, para determinar la secuencia del molde, pueda utilizarse la detección del orden y tipo de nucleótidos añadidos al cebador.

Las cubetas de lectura proporcionan un soporte sólido conveniente para alojar fragmentos de ADN amplificados producidos mediante los métodos de la presente descripción. Uno o más fragmentos de ADN amplificados en dicho formato pueden someterse a una SBS o a otra técnica de detección que implique el suministro repetido de reactivos en ciclos. Por ejemplo, para iniciar un primer ciclo de SBS, uno o más nucleótidos marcados, una ADN polimerasa, etc., pueden fluir hacia/a través de una cubeta de lectura que contenga una o más moléculas de ácido nucleico amplificado. Pueden detectarse aquellos sitios en los que la extensión del cebador provoque la incorporación de un nucleótido marcado. Opcionalmente, los nucleótidos pueden incluir además una propiedad de terminación reversible que finalice la extensión adicional del cebador una vez que se haya añadido un nucleótido a un cebador. Por ejemplo, a un cebador se le puede añadir un análogo de nucleótido que tenga una fracción de terminación reversible, de modo que no se pueda producir la extensión posterior hasta que se suministre un agente desbloqueante para eliminar la fracción. Por lo tanto, para realizaciones que usan terminación reversible, puede suministrarse un reactivo de desbloqueo a la célula de flujo (antes o después de que se produzca la detección). Los lavados pueden llevarse a cabo entre las diversas etapas de suministro. Después, el ciclo puede repetirse n veces para extender el cebador en n nucleótidos, detectando de este modo una secuencia de longitud n. Se describen procedimientos de SBS, sistemas de fluidos y plataformas de detección ilustrativos que pueden adaptarse fácilmente para su uso con amplicones producidos mediante los métodos de la presente divulgación, por ejemplo, en Bentley y col., *Nature* 456:53-59 (2008), en el documento WO 04/018497; US-7.057.026; WO 91/06678; WO 07/123744; US-7.329.492; US-7.211.414; US-7.315.019; US-7.405.281, y US 2008/0108082.

Pueden utilizarse otros procedimientos de secuenciación que utilizan reacciones cíclicas, como la pirosecuenciación. La pirosecuenciación detecta la liberación de pirofosfato inorgánico (PPi) a medida que se incorporan nucleótidos particulares en una cadena de ácido nucleico incipiente (Ronaghi, y col., *Analytical Biochemistry* 242(1), 84-9 (1996); Ronaghi, *Genome Res.* 11(1), 3-11 (2001); Ronaghi y col. *Science* 281(5375), 363 (1998); US-6.210.891; US-6.258.568 y US-6.274.320). En la pirosecuenciación, el PPi liberado se puede detectar convirtiéndolo inmediatamente en trifosfato de adenosina (ATP) por la ATP sulfurilasa, y el nivel de ATP generado se puede detectar mediante fotones producidos por luciferasa. Por lo tanto, la reacción de secuenciación se puede controlar a través de un sistema de detección de luminiscencia. Las fuentes de radiación de excitación utilizadas para los sistemas de detección a base de fluorescencia no son necesarias para los procedimientos de pirosecuenciación. Los sistemas fluidicos, detectores y procedimientos útiles que se pueden usar para la aplicación de pirosecuenciación a amplicones producidos según la presente descripción se describen, por ejemplo, en los documentos WO 2012/058096, US 2005/0191698 A1, US-7.595.883 y US-7.244.559.

Algunas realizaciones pueden utilizar métodos que suponen la monitorización en tiempo real de la actividad de la ADN polimerasa. Por ejemplo, las incorporaciones de nucleótidos pueden detectarse a través de interacciones de transferencia de energía por resonancia de fluorescencia (FRET) entre una polimerasa portadora de fluoróforo y nucleótidos marcados con  $\gamma$ -fosfato o con guías de ondas de modo cero (ZMW). Se describen técnicas y reactivos de secuenciación basada en FRET, por ejemplo, en Levene y col. *Science* 299, 682-686 (2003); Lundquist y col. *Opt. Lett.* 33, 1026-1028 (2008); Korch y col. *Proc. Natl. Acad. Sci. EE. UU.* 105, 1176-1181 (2008).

Algunas realizaciones de SBS (secuenciación por síntesis) incluyen la detección de un protón liberado tras la incorporación de un nucleótido en un producto de extensión. Por ejemplo, la secuenciación basada en la detección de protones liberados puede usar un detector eléctrico y técnicas asociadas que son comercializados por Ion Torrent (Guilford, CT, una subsidiaria de Life Technologies) o métodos y sistemas de secuenciación descritos en los documentos US 2009/0026082 A1; US 2009/0127589 A1; US 2010/0137143 A1; o US 2010/0282617 A1. Los métodos expuestos en el presente documento para amplificar ácidos nucleicos diana utilizando exclusión cinética pueden aplicarse fácilmente a sustratos utilizados para detectar protones. Más específicamente, los métodos expuestos en el presente documento pueden utilizarse para producir poblaciones clonales de amplicones que se utilizan para detectar protones.

Otra técnica de secuenciación útil es la secuenciación de nanoporos (véase, por ejemplo, Deamer y col. *Trends Biotechnol.* 18, 147-151 (2000); Deamer y col. *Acc. Chem. Res.* 35:817-825 (2002); Li y col. *Nat. Mater.* 2:611-615 (2003)). En algunas realizaciones de secuenciación de nanoporos, el ácido nucleico diana o los nucleótidos individuales eliminados de un ácido nucleico diana pasan a través de un nanoporo. A medida que el ácido nucleico o el nucleótido pasa a través del nanoporo, cada tipo de nucleótido puede identificarse midiendo las fluctuaciones en la

conductancia eléctrica del poro. (Patente US-7.001.792; Soni y col. Clin. Chem. 53, 1996-2001 (2007); Healy, Nanomed. 2, 459-481 (2007); Cockroft y col. J. Am. Chem. Soc. 130, 818-820 [2008]).

Los métodos ilustrativos para los análisis de expresión y genotipado basados en matriz que pueden aplicarse a la detección según la presente divulgación se describen en las patentes US-7.582.420; US-6.890.741; US-6.913.884 o US-6.355.431 o publicación de patente estadounidense n.º 2005/0053980 A1; 2009/0186349 A1 o US 2005/0181440 A1.

Una ventaja de los métodos expuestos en la presente descripción es que proporcionan una detección rápida y eficiente de una pluralidad de ácidos nucleicos diana en paralelo. Por consiguiente, la presente descripción proporciona sistemas integrados capaces de preparar y detectar ácidos nucleicos usando técnicas conocidas en la técnica tales como las ilustradas anteriormente. Por tanto, un sistema integrado de la presente descripción puede incluir componentes fluidicos capaces de administrar reactivos de amplificación y/o reactivos de secuenciación a uno o más fragmentos de ADN inmovilizados, comprendiendo el sistema componentes tales como bombas, válvulas, depósitos, líneas fluidicas y similares. En un sistema integrado para la detección de ácidos nucleicos diana puede configurarse y/o utilizarse una cubeta de lectura. Por ejemplo, se describen cubetas de lectura ilustrativas en los documentos US 2010/0111768 A1 y US con n.º de serie 13/273.666. Como se ilustra en las cubetas de lectura, puede utilizarse uno o más de los componentes fluidicos de un sistema integrado para un método de amplificación y para un método de detección. Tomando como ejemplo una realización de secuenciación de ácidos nucleicos, puede utilizarse uno o más de los componentes fluidicos de un sistema integrado para un método de amplificación expuesto en la presente memoria y para el suministro de reactivos de secuenciación en un método de secuenciación como en los ilustrados anteriormente. Como alternativa, un sistema integrado puede incluir sistemas fluidicos distintos para llevar a cabo métodos de amplificación y para llevar a cabo métodos de detección. Como ejemplos de sistemas de secuenciación integrados que son capaces de crear ácidos nucleicos amplificados y también de determinar la secuencia de los ácidos nucleicos se incluyen, sin limitación, la plataforma MiSeq™ (Illumina, Inc., San Diego, CA) y los dispositivos descritos en la publicación de solicitud de patente estadounidense n.º 2012/0270305.

#### Soportes sólidos con transposomas inmovilizados y métodos de preparación

Otras realizaciones presentadas en el presente documento incluyen soportes sólidos, tales como cubetas de lectura, que tienen complejos transposómicos inmovilizados sobre los mismos. En determinadas realizaciones, los complejos transposómicos comprenden una transposasa unida a un primer polinucleótido, comprendiendo el polinucleótido (i) una parte 3' que comprende una secuencia terminal del transposón y (ii) un primer marcador que comprende un primer dominio de marcaje. La densidad de estos transposomas unidos a la superficie puede variar. Por ejemplo, en algunas realizaciones, los complejos transposómicos están presentes en el soporte sólido a una densidad de al menos  $10^3$ ,  $10^4$ ,  $10^5$  o al menos  $10^6$  complejos por  $\text{mm}^2$ .

En la presente descripción también se presentan métodos para generar una cubeta de lectura para la tagmentación. Los métodos pueden comprender, por ejemplo, inmovilizar una pluralidad de complejos transposómicos en un soporte sólido, comprendiendo los complejos transposómicos una transposasa unida a un primer polinucleótido, comprendiendo el primer polinucleótido (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) un primer marcador que comprende un primer dominio de marcaje.

Los complejos transposómicos pueden inmovilizarse en un soporte sólido mediante diversos métodos, lo que apreciará un experto en la técnica. En una realización, el método comprende proporcionar un soporte sólido que tenga una pluralidad de los primeros polinucleótidos inmovilizados en el mismo y poner en contacto el soporte sólido con la holoenzima transposasa y un segundo polinucleótido, comprendiendo el segundo polinucleótido una región complementaria de la secuencia terminal de transposón. En algunas realizaciones, el segundo polinucleótido se hibrida con el primer polinucleótido inmovilizado antes de añadir la holoenzima transposasa. En algunas realizaciones, el segundo polinucleótido y la holoenzima transposasa se proporcionan juntos. La figura 5a muestra una realización de este método para ensamblar los complejos transposómicos unidos a la superficie. En un método, la holoenzima transposasa se añade junto con la cadena no transferida del adaptador ME a una cubeta de lectura que contiene oligonucleótidos que comprenden, desde el extremo injertado en la superficie 5': un cebador de amplificación por grupos (A.pr) y un cebador de secuenciación (S.pr) y la secuencia ME (figura 5a). Alternativamente, la cadena ME' no transferida se puede hibridar primero con la cadena ME injertada en la superficie y luego añadir la transposasa a la cubeta de lectura.

En algunas realizaciones, los complejos transposómicos se ensamblan en solución y la inmovilización comprende una etapa adicional de ligar el primer polinucleótido a un oligonucleótido férula acoplado al soporte sólido. Esta realización se ilustra en la figura 5b. En algunas realizaciones, el oligonucleótido férula se puede extender usando una polimerasa antes de que se produzca una etapa de ligamiento.

En algunas realizaciones, el dímero de transposoma se ensambla hibridando un oligonucleótido en bucle con un primer polinucleótido inmovilizado. Por ejemplo, un oligonucleótido en bucle puede comprender un primer extremo y un segundo extremo, siendo el primer extremo complementario de la secuencia terminal del transposón del primer polinucleótido. El segundo extremo del oligonucleótido en bucle puede ser complementario de una segunda secuencia terminal del transposón. La segunda secuencia terminal del transposón puede ser, por ejemplo, parte de un primer polinucleótido en fase de solución. En algunas realizaciones, el primer polinucleótido inmovilizado y el primer polinucleótido en fase de solución pueden comprender secuencias terminales de transposones diferentes o complementos de las mismas. En algunas de estas

realizaciones, el oligonucleótido en bucle comprende secuencias complementarias de cada una de las secuencias terminales de transposones diferentes en los extremos primero y segundo. En la figura 5c se muestra una ilustración de esta realización. Como se muestra en la figura 5c, se genera un par de adaptadores contiguos hibridando dos oligonucleótidos con el oligonucleótido unido a la superficie. Esto se puede lograr empleando dos secuencias terminales de transposones diferentes (ME1 y ME2). La adición de la holoenzima transposasa reconstituye un complejo transposómico “en bucle” activo, donde solo uno de los dos adaptadores de cada transposoma se acopla a la superficie (figura 5c).

En otra realización, los complejos transposómicos se pueden ensamblar en una cubeta de lectura estándar de extremos apareados con cebadores de amplificación inmovilizados en la misma (p. ej., una cubeta de lectura HiSeq o una cubeta de lectura MiSeq vendidas por Illumina Inc, San Diego, CA). Esto se puede lograr, por ejemplo, mediante la hibridación de un oligonucleótido “férula” que se hibrida con una o ambas especies de cebadores de amplificación injertados en superficie. La férula actúa como molde para a continuación extender el cebador de superficie injertado con una polimerasa y dNTP para formar una doble cadena oligonucleotídica que contiene un cebador de amplificación de superficie, un cebador de secuenciación y las secuencias terminales de los transposones de la transposasa. La adición de transposasa ensambla un transposoma de la superficie. Esta realización se ilustra en la figura 5d.

En cualquiera de las realizaciones proporcionadas en el presente documento, los complejos transposómicos pueden ser homodímeros o heterodímeros. Por ejemplo, como se ilustra en la figura 11, un transposoma homodimérico comprendería dos adaptadores P5-ME en ambos sitios o, alternativamente, dos adaptadores P7-ME. De manera similar, un complejo transposómico heterodimérico comprendería adaptadores P5-ME y P7-ME, como se muestra en la figura 11.

#### Tagmentación mediante perlas de transposomas

Una realización presentada en el presente documento es una población de micropartículas que tienen complejos transposómicos inmovilizados en las mismas. El uso de un soporte sólido, tal como perlas, puede proporcionar varias ventajas sobre la tagmentación en solución. Por ejemplo, en la tagmentación estándar en solución, es difícil controlar el tamaño final del fragmento de la reacción de tagmentación. El tamaño del fragmento depende de la relación entre los transposomas y la cantidad y el tamaño del ADN y la duración de la reacción. Incluso si se controlan estos parámetros, el fraccionamiento por selección de tamaño es necesario habitualmente como una etapa adicional para eliminar el exceso de fragmentos pequeños más cortos que las longitudes combinadas de lectura en ambos extremos. Los métodos proporcionados en el presente documento evitan esas desventajas. Específicamente, los transposomas inmovilizados en perlas permiten la selección del tamaño final del fragmento en función de la separación espacial de los transposomas unidos, independientemente de la cantidad de perlas de transposomas añadidas a la reacción de tagmentación. Una limitación adicional de la tagmentación basada en solución es que típicamente es necesario realizar alguna forma de purificación de los productos de la reacción de tagmentación tanto antes como después de la amplificación por PCR. Esto requiere típicamente cierta transferencia de reacciones de un tubo a otro. Por el contrario, los productos de tagmentación en los transposomas basados en perlas pueden lavarse y luego liberarse para su amplificación u otro procesamiento posterior, evitando así la necesidad de transferir muestras. Por ejemplo, en las realizaciones donde los transposomas se ensamblan en perlas paramagnéticas, la purificación de los productos de la reacción de tagmentación se puede lograr fácilmente inmovilizando las perlas con imanes y lavándolas. Por lo tanto, en algunas realizaciones, la tagmentación y otros procesos posteriores, tales como la amplificación por PCR, pueden realizarse en un solo tubo, recipiente, gotícula u otro envase. En algunas realizaciones, la tagmentación y el procesamiento posterior de las muestras tienen lugar en un dispositivo basado en gotículas microfluídicas. Por ejemplo, en un dispositivo basado en gotículas microfluídicas, puede hacerse pasar una gotícula que contiene el ácido nucleico diana, el tampón de lavado u otros reactivos sobre una superficie que comprende complejos transposómicos inmovilizados. Del mismo modo, una gotícula que comprende perlas que tienen transposomas inmovilizados en las mismas puede ponerse en contacto con el ácido nucleico diana, el tampón de lavado u otros reactivos en un dispositivo basado en gotículas microfluídicas.

En algunas realizaciones, los complejos transposómicos inmovilizados comprenden una transposasa unida a un primer polinucleótido y a un segundo polinucleótido; en donde el primer polinucleótido se inmoviliza en su extremo 5' en la superficie de la micropartícula y el segundo polinucleótido se hibrida con el extremo 3' del primer polinucleótido; y en donde el primer polinucleótido comprende: (i) una parte 3' que comprende una secuencia terminal de transposón y (ii) un primer marcador que comprende un primer dominio de marcaje.

Las figuras 9-12 proporcionan una ilustración de los transposomas ensamblados en la superficie de las perlas paramagnéticas. La figura 9 muestra una superficie de perla con dos primeros polinucleótidos diferentes inmovilizados en la misma. Los primeros polinucleótidos que se muestran comprenden una secuencia terminal de transposón (ME). Uno de los primeros polinucleótidos mostrados comprende un dominio de marcaje que comprende una secuencia cebadora de amplificación (P5) y una secuencia cebadora de secuenciación (lectura 1). El otro polinucleótido mostrado en la figura 9 comprende una secuencia cebadora de amplificación diferente (P7) y una secuencia cebadora de secuenciación (lectura 2). Los primeros polinucleótidos también pueden comprender un marcador índice. El marcador índice puede ser exclusivo de la perla o puede compartirse con una o más perlas de la población. Una sola perla puede tener solo un marcador índice inmovilizado en la misma o puede tener una pluralidad de marcadores índice inmovilizados en la misma.

La figura 10 muestra un segundo polinucleótido hibridado con cada primer polinucleótido. El segundo polinucleótido comprende una región complementaria de la secuencia terminal del transposón del primer polinucleótido. En algunas

realizaciones, el segundo polinucleótido tiene una longitud de 15 pb. En algunas realizaciones, el segundo polinucleótido puede tener una longitud de aproximadamente 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 pb o más de 20 pb. En algunas realizaciones, el segundo polinucleótido está fosforilado en su extremo 5'.

5 La figura 11 ilustra el ensamblaje de transposomas en la superficie de la perla, donde la enzima transposasa se pone en contacto con los oligonucleótidos inmovilizados. Como se muestra en la figura 11, cuando la transposasa se añade a las perlas, se forman tres tipos de complejos transposómicos: un complejo P5:P5, un complejo P7:P7 y un complejo P5:P7 (figura 11).

10 Cuando las perlas del transposoma se añaden a una solución de ADN diana en un tampón de tagmentación, tiene lugar la tagmentación, que une el ADN a la superficie de las perlas. Se genera una biblioteca inmovilizada de fragmentos de ADN marcados.

15 La figura 12 ilustra una realización donde la tagmentación sucesiva en el ADN da lugar a moléculas con puentes entre los transposomas. Cuando están presentes tres tipos de transposomas (como en la figura 11), se obtienen tres tipos de complejos puente: un complejo P5:P5, un complejo P7:P7 y un complejo P5:P7 en una relación de 25:25:50, respectivamente.

20 En algunas realizaciones, la longitud de los fragmentos con puente puede estar dictada por la densidad de los transposomas en la superficie de la perla. Esta densidad se puede ajustar mediante la cantidad de oligonucleótido en la superficie, la cantidad de complejos terminales de transposones bicatenarios en la superficie y la cantidad de enzima transposasa añadida durante el ensamblaje del transposoma. Una vez que se completa la tagmentación, los productos de tagmentación P5:P7 pueden liberarse de la superficie de la perla usando cualquier método adecuado. En algunas realizaciones, los productos de tagmentación se liberan de las perlas usando un método de amplificación tal como la PCR de supresión, la PCR de desplazamiento exterior ("*step-out*") y similares. En algunas realizaciones, los productos de tagmentación se liberan de las perlas por escisión. La escisión puede ser, por ejemplo, química, enzimática, fotoquímica o una combinación de las mismas. Se apreciará que se puede utilizar cualquier método adecuado para liberar uno o más productos de tagmentación de un soporte sólido en los métodos proporcionados en el presente documento.

30 El ADN se puede poner en contacto eficientemente con los transposomas unidos a la superficie usando cualquier método adecuado para aumentar la probabilidad de contacto. Por ejemplo, en algunas realizaciones, la precipitación de ADN sobre la superficie sólida se puede utilizar para aumentar el contacto entre el ADN diana y los complejos transposómicos en la superficie sólida. Se puede utilizar cualquiera de los diversos métodos que se conocen en la técnica para poner en contacto el ADN con un soporte sólido, como se ilustra en la descripción del documento WO 2010/115122. Como apreciará un experto en la técnica, el ADN puede precipitarse sobre un soporte sólido mediante la adición de PEG, etanol o cualquiera de diversos otros agentes conocidos por precipitar el ADN sobre las superficies, incluyendo, por ejemplo, cualquiera de los diversos tampones utilizados en la tecnología de inmovilización reversible en fase sólida (SPRI).

40 En algunas realizaciones, una población de perlas que contienen complejos transposómicos inmovilizados se puede mezclar con un exceso de perlas que no contienen transposomas u oligonucleótidos, lo que reduce la probabilidad de tagmentación entre dos o más perlas diferentes. Otro método para reducir la probabilidad de tagmentación entre dos o más perlas diferentes incluye inmovilizar las perlas para minimizar el contacto entre ellas. La inmovilización de las perlas se puede lograr mediante cualquiera de las diversas técnicas conocidas en la técnica, que incluyen, por ejemplo, la inmovilización de las perlas mediante magnetismo en los lados de una superficie sólida, tal como un tubo de microcentrífuga, o cualquier otra técnica de inmovilización, como se ilustra en el documento WO 2010/115122.

50 En algunas realizaciones, las perlas de transposomas se pueden usar para aislar e identificar ácidos nucleicos de una sola célula, tal como una célula procariota o eucariota. Por ejemplo, en algunas realizaciones, las partículas, tales como las perlas, se recubren con transposomas indexados que comparten el mismo índice (todos los transposomas presentes en una perla particular tienen el mismo índice, que es diferente del índice presente en otra perla). A continuación, las perlas se pueden colocar dentro de las células mediante cualquiera de las diversas metodologías conocidas en la técnica. Por ejemplo, los métodos para administrar perlas dentro de las células incluyen, pero sin limitación, pistolas génicas, hojas nanométricas fototérmicas (Wu y col. Anal Chem. (2011) 4:1321-7) y péptidos usados junto con sustancias permeabilizadoras de células (Nitin y col Ann Biomed Eng. (2009) 37:2018-2027) y similares. Se apreciará que puede usarse cualquier método adecuado para asociar el ADN de una sola célula con una partícula que porta transposomas indexados en los métodos presentados en el presente documento.

60 En algunas realizaciones, los transposomas se pueden unir covalentemente a las perlas como se ha descrito en detalle anteriormente. De forma adicional o alternativamente, los transposomas pueden liberarse de las perlas tras la aplicación de un estímulo químico o físico. Algunos ejemplos de estímulos que pueden desencadenar la liberación del transposoma desde un soporte sólido incluyen cambios de luz y/o temperatura. En algunas realizaciones, los transposomas se liberan del soporte sólido usando la actividad de una enzima tal como una endonucleasa de restricción. En determinadas realizaciones, los transposomas pueden desprenderse de las perlas y moverse libremente dentro de la célula. Una vez que las perlas (o, alternativamente, los transposomas liberados) entran en contacto con la cromatina o el ADN, puede tener lugar la tagmentación. Se entenderá que, en los sistemas eucariotas y procariotas, no todo el ADN genómico estará siempre accesible y/o disponible para la tagmentación.



En algunas realizaciones, hasta el 0,001 %, el 0,01 %, el 0,1 %, el 1 %, el 2 %, el 3 %, el 4 %, el 5 %, el 6 %, el 7 %, el 8 %, el 9 %, el 10 %, el 15 %, el 20 %, el 25 %, el 30 %, el 35 %, el 40 %, el 45 %, el 50 %, el 55 %, el 60 %, el 65 %, el 70 %, el 75 %, el 80 %, el 85 %, el 90 %, el 95 % o el 99 % o más del 99 % del ADN total de una célula es tagmentado por los transposomas. El uso de perlas marcadas de forma única permite identificar las lecturas de la misma célula agrupando las lecturas que comparten los mismos índices. Estas lecturas pueden considerarse derivadas de la misma perla (y por lo tanto de la misma célula). De forma adicional o alternativamente, en algunas realizaciones, se puede realizar una etapa de entrecruzamiento para entrecruzar el ADN celular mientras aún está dentro del núcleo de la célula. Esta etapa garantiza que el ADN de una sola célula se mantenga unido. Después de extraer el ADN de las células, puede mezclarse con las perlas indexadas y someterse a tagmentación como se ha descrito en detalle anteriormente. De forma adicional o alternativamente, las variaciones del método incluyen la tagmentación de los cromosomas clasificados por flujo. Los cromosomas clasificados por flujo se pueden mezclar con las perlas indexadas descritas anteriormente. Un cromosoma en particular se adheriría a una perla y los fragmentos generados mediante la tagmentación contendrían el mismo índice. Esto podría permitir la eliminación gradual de los SNP de cromosomas completos (haplotipado).

En algunas de estas realizaciones, el número total de índices es mayor que el número total de células que se han tagmentado satisfactoriamente. En algunas realizaciones, se puede implementar un enfoque de doble índice, por ejemplo, como una forma de aumentar el número de posibles combinaciones de índices. Como ejemplo, el uso de dos índices de 8 bases proporcionaría un número teórico de combinaciones de  $4^8 \times 4^8 = 4,3 \times 10^9$ .

En algunas realizaciones, se puede usar un enfoque para garantizar que una célula individual no sea tagmentada por múltiples perlas. Por ejemplo, un enfoque es usar perlas de un tamaño similar al de la célula. Esto garantizaría que una célula no pudiera alojar múltiples perlas. De forma adicional o alternativamente, otro enfoque es hacer uso de una proporción de células con respecto a perlas que favorezca el direccionamiento a una sola célula. Por ejemplo, si hay muchas más células que perlas, la distribución de Poisson de las perlas dentro de las células significa que las células que han captado una sola perla superan con creces en número a las células que han captado dos o más perlas.

En algunas realizaciones, el enfoque de célula única descrito anteriormente se puede usar para determinar si dos SNP (o reordenamientos estructurales) están presentes en la misma célula. Por ejemplo, en el caso de poblaciones heterogéneas de células cancerosas, saber si dos SNP están presentes en la misma célula o en células diferentes podría ayudar a implementar la terapia adecuada contra el cáncer.

En algunas realizaciones, el enfoque de célula única descrito anteriormente se puede usar para estudiar el ARN. Por ejemplo, al recubrir las perlas con enzimas adecuadas (es decir, transcriptasa inversa) y oligonucleótidos para una etapa de transcripción inversa, se puede analizar la expresión génica a nivel de una sola célula. En una realización, después de introducir las perlas recubiertas con transcriptasa inversa, oligonucleótidos y transposomas, el ARN citoplasmático puede convertirse en ADNc, marcarse y prepararse dentro de las células.

Métodos para ensamblar lecturas largas utilizando códigos de barras en transposomas inmovilizados

El ensamblaje asistido por código de barras de fragmentos de ADN permite el aislamiento de moléculas de ADN largas individuales dentro de una población de moléculas de ADN y la conversión de cada molécula en una biblioteca de subfragmentos con códigos de barras únicos. Cuando se secuencian toda la población de moléculas de ADN subfragmentadas, los subfragmentos se pueden volver a ensamblar en su molécula larga original haciendo referencia a los códigos de barras que contienen.

Se conocen diversos métodos para generar códigos de barras para moléculas de ADN individuales. Por ejemplo, el “método de dilución” aísla moléculas largas individuales mediante dilución extrema y las divide en alícuotas en compartimentos separados (p. ej., los pocillos de una placa), de modo que cada pocillo contenga solo una o unas pocas moléculas de ADN. Debido a que cada pocillo está físicamente separado, se puede hacer una preparación de biblioteca en cada pocillo con un código de barras único. Posteriormente, los contenidos de los pocillos se agrupan y se secuencian. Otro método emplea una emulsión en donde cada gotícula contiene moléculas de ADN largas, reactivos de preparación de bibliotecas y un código de barras único para cada gotícula. Otro enfoque utiliza una gran biblioteca de complejos transposómicos en bucle indexados para insertar múltiples códigos de barras gemelos a lo largo del ADN y, al mismo tiempo, conservar la integridad de la molécula. La posterior división entre los códigos de barras “gemelos” produce fragmentos que se pueden secuenciar y volver a ensamblar haciendo coincidir los códigos de barras gemelos. Cada uno de los métodos de generación de códigos de barras mencionados anteriormente conlleva desventajas que se superan con los métodos de generación de códigos de barras presentados en la presente descripción.

En el presente documento se presentan métodos alternativos a las formas descritas anteriormente de aislar moléculas largas individuales y generar códigos de barras para ellas. Los métodos presentados en el presente documento logran ventajas de aislamiento físico similares al método de emulsión sin usar emulsiones y, al mismo tiempo, proporcionan una complejidad que es mucho mayor que la proporcionada por la gran cantidad de “pocillos” utilizados en el método de dilución. Los códigos de barras únicos de los presentes métodos son en cierto modo análogos a los “pocillos” del método de dilución, excepto que el número de perlas en el método basado en perlas a menudo puede ser mucho mayor que el número de pocillos en el método de dilución. Una ventaja adicional sobre los métodos de emulsión es que, en el método



basado en perlas, los códigos de barras se distribuyen de forma determinista (es decir, un código de barras por perla) y no de forma aleatoria (es decir, con distribución de Poisson). Los métodos presentados en el presente documento también logran la misma conservación inicial de la integridad y contigüidad de la molécula, pero sin la necesidad de complejos transposómicos en bucle como los utilizados en algunos otros métodos. De forma adicional, los métodos presentados en la presente descripción no requieren un espacio de código tan grande como el usado en algunos otros métodos.

En consecuencia, en algunas realizaciones presentadas en la presente descripción, el método de generación de códigos de barras comprende proporcionar una población de micropartículas que tienen complejos transposómicos inmovilizados en las mismas, comprendiendo los complejos transposómicos una transposasa unida a un primer polinucleótido y a un segundo polinucleótido. En algunas realizaciones, el primer polinucleótido comprende un dominio índice asociado a la micropartícula. En algunas realizaciones, el dominio índice puede ser exclusivo de la micropartícula. En algunas realizaciones, la población de micropartículas comprende al menos una pluralidad de dominios índice. En algunas realizaciones, el dominio índice está presente en más de una micropartícula de la población de micropartículas. En algunas realizaciones, una micropartícula en la población de micropartículas comprende más de un dominio índice.

Los métodos de generación de códigos de barras presentados en el presente documento comprenden además aplicar un ADN diana a la población de micropartículas, generando de esta manera fragmentos de ADN inmovilizados que están marcados con el dominio índice. El ADN se puede poner en contacto eficientemente con los transposomas unidos a la superficie usando cualquier método adecuado para aumentar la probabilidad de contacto como se ilustra en el documento WO 2010/115122.

Los métodos se pueden realizar usando cualquiera de diversos formatos conocidos, por ejemplo, con una combinación de reactivos de tagmentación y una matriz de perlas para la preparación de la biblioteca, seguida de un ciclo de secuenciación indexada y un análisis de datos personalizado. Se puede usar cualquier otro método adecuado que mantenga las perlas en separación estática entre sí para la tagmentación en la superficie y la indexación de las muestras. Por ejemplo, configuraciones físicas tales como pocillos o pequeñas depresiones en el sustrato que puedan retener las perlas, de manera que una microesfera pueda reposar en el pocillo, o el uso de otras fuerzas (magnéticas o compresivas) o sitios químicamente alterados o activos, tales como sitios funcionalizados químicamente, sitios alterados electrostáticamente, sitios funcionalizados de manera hidrófoba y/o hidrófila o puntos de adhesivo.

En algunas realizaciones, las microesferas están asociadas de manera no covalente en los pocillos, aunque los pocillos pueden funcionalizarse químicamente de forma adicional como se describe generalmente a continuación, se pueden usar agentes de reticulación o se puede usar una barrera física, p. ej., una película o membrana sobre las perlas.

En determinadas realizaciones, la superficie del sustrato se modifica para contener sitios modificados químicamente que pueden usarse para unir, de forma covalente o no covalente, las microesferas de la invención a los sitios o ubicaciones individuales del sustrato. Los “sitios modificados químicamente” en este contexto incluyen, pero no se limitan a, la adición de un patrón de grupos funcionales químicos que incluyen grupos amino, grupos carboxi, grupos oxo y grupos tiol, que pueden usarse para unir covalentemente microesferas, que generalmente también contienen los grupos funcionales reactivos correspondientes; la adición de un patrón de adhesivo que pueda usarse para unir las microesferas (ya sea mediante funcionalización química previa para la adición del adhesivo o mediante la adición directa del adhesivo); la adición de un patrón de grupos con carga (similar a las funcionalidades químicas) para la unión electrostática de las microesferas, p. ej., cuando las microesferas comprenden grupos con carga opuestos a los sitios; la adición de un patrón de grupos funcionales químicos que haga que los sitios sean diferencialmente hidrófobos o hidrófilos, de modo que la adición de microesferas hidrófobas o hidrófilas similares en condiciones experimentales adecuadas dará lugar a la asociación de las microesferas a los sitios en función de la hidroafinidad. Por ejemplo, el uso de sitios hidrófobos con perlas hidrófobas, en un sistema acuoso, impulsa la asociación preferencial de las perlas a los sitios. Como se ha descrito anteriormente, el “patrón” en este sentido incluye el uso de un tratamiento uniforme de la superficie para permitir la unión de las perlas en sitios individuales, así como el tratamiento de la superficie que da lugar a sitios individuales. Como apreciarán los expertos en la técnica, esto puede lograrse de diversas maneras.

En determinadas realizaciones, se generan una multitud de perlas que comprenden transposomas unidos a la superficie, en donde cada perla contiene muchos transposomas, pero todos los transposomas de cualquier perla determinada contienen el mismo código de barras. La generación de una población de oligonucleótidos de transposomas con códigos de barras monoclonales en perlas se puede realizar según una cualquiera de las diversas técnicas conocidas en la técnica, como se ilustra en la descripción del documento US-5.604.097.

La figura 13 ilustra una realización donde un transposoma unido a la superficie contiene oligonucleótidos largos unidos a la superficie a través de su extremo 5'. Como se expone en la figura 13, el extremo 3' del oligonucleótido comprende las secuencias del extremo en mosaico (METS) de la enzima transposasa Tn5. Cadena arriba de esta secuencia (más cerca del extremo 5') hay una secuencia de código de barras de típicamente 6 a 8 bases de longitud. Más cadena arriba hay una secuencia cebadora. De forma adicional, el oligonucleótido puede comprender además un sitio de escisión para permitir que el oligonucleótido se escinda de la superficie: su presencia en el oligonucleótido es opcional para el método general. Un segundo oligonucleótido corto (la cadena no transferida, MENTS) se hibrida con la secuencia METS en el extremo 3' del oligonucleótido injertado en la superficie larga. Finalmente, se ensambla un dímero de transposasa en el extremo 3' de los oligonucleótidos para formar un transposoma funcional.

La figura 14 ilustra el código de barras en las perlas en una matriz de pocillos. Como se muestra en la figura 14, cuando se añaden moléculas largas de ADNbc a una matriz de transposomas inmovilizados en perlas, una molécula dada encuentra una perla y es tagmentada muchas veces por los transposomas de la perla. Cada fragmento se inmoviliza en la perla y se marca con el código de barras asociado a esa perla. En la realización particular que se muestra en la figura 14, la separación física de las perlas en el chip de matriz evita que una molécula de ADN se introduzca entre dos perlas. En otras realizaciones, las perlas están en estrecho contacto y una o más moléculas de ADN pueden estirarse entre dos o más perlas. En algunas realizaciones, se puede tagmentar más de una molécula de ADN por perla. La probabilidad de que dos alelos se tagmenten en la misma perla es baja y depende de la concentración del ADN añadido, del número de perlas y del número de códigos de barras. Por ejemplo, para evitar que dos alelos aparezcan en el mismo pocillo, sería necesario cargar 0,1 veces equivalentes de haplomas (50.000 equivalentes de genomas) en 1 millón de perlas, cada una con un código de barras único.

La figura 15 ilustra la transferencia de moléculas marcadas con códigos de barras a una reacción de secuenciación. Como se muestra en la figura 15, una vez que se completa la tagmentación, el ADN se transfiere desde la superficie de la perla a la solución para que los fragmentos individuales puedan agruparse y prepararse para la secuenciación. La secuenciación y el ensamblaje por referencia a los códigos de barras permiten recrear la secuencia de las moléculas de ADN originales tagmentadas largas, lo que permite lecturas largas o pseudolargas y determinación de fases de los SNP.

La liberación de los fragmentos tagmentados en la superficie con códigos de barras a las soluciones se puede realizar usando cualquier método adecuado, como se conoce en la técnica. Por ejemplo, en algunas realizaciones, las moléculas tagmentadas pueden escindirse de la superficie de las perlas mediante un resto de escisión presente en el extremo 5' de los oligonucleótidos unidos a la superficie (véase la figura 13). El resto de escisión puede ser cualquier resto adecuado para la escisión de una cadena de ácido nucleico de un soporte sólido. Los ejemplos de métodos que utilizan un resto de escisión incluyen, pero no se limitan a, la escisión por endonucleasas de restricción, la escisión química, la escisión por RNasa, la escisión fotoquímica y similares, incluidos los métodos de escisión y fracciones de escisión establecidos en el documento WO 2006/064199.

La escisión usando un resto de escisión produce una molécula que tiene el siguiente formato:

5' - Cebador - Código de barras - ME - Inserto - ME - Código de barras - Cebador - 3'

Las regiones “cebadoras” se pueden usar como puntos de hibridación para hibridar los cebadores de desplazamiento exterior de la PCR que permiten añadir secuencias adicionales, tales como cebadores de amplificación y secuenciación. Por ejemplo, se pueden añadir los cebadores de amplificación P5 y P7. Una vez añadidos, se puede utilizar PCR de supresión, por ejemplo, para enriquecer moléculas que tienen adaptadores P5 en un extremo y P7 en el otro.

En algunas realizaciones, la amplificación se puede realizar directamente a partir de las perlas con cebadores de desplazamiento exterior que añaden secuencias adaptadoras de P5 y P7 mediante PCR de supresión. En otra realización, cada perla puede tener dos tipos de oligonucleótidos injertados en la superficie donde la secuencia cebadora (como en la figura 13) es un cebador de secuenciación P5-Read1 o un cebador de secuenciación P7-Read 2. Esto dará lugar a transposomas P5-P7 mixtos. Estos pueden escindirse de las perlas y seguirse de una PCR de supresión para enriquecer las moléculas P5/P7 o amplificarse directamente de las perlas, como se ha descrito anteriormente.

En algunas realizaciones, puede estar presente un único tipo de transposoma (p. ej., un código de barras P5-Read 1) en la superficie de la perla. Una vez que se completa la tagmentación en la superficie, se puede añadir en solución un segundo transposoma que porte un cebador de amplificación y/o secuenciación diferente para escindir las moléculas con puente. Esto produce todas las moléculas con el mismo formato de adaptador que pueden escindirse o amplificarse fuera de la superficie de la perla. Podría añadirse un código de barras adicional específico para la muestra al transposoma en solución, de modo que se puedan agrupar múltiples muestras mediante el método. La figura 16 es una ilustración de esta realización. Como se muestra en la figura 16, se añaden transposomas en fase de solución que portan una secuencia cebadora de amplificación P7 a una perla con productos de tagmentación con puente inmovilizados. Los fragmentos inmovilizados se marcan con una secuencia polinucleotídica que comprende una secuencia cebadora de amplificación P5 y un marcador índice de código de barras (i9) que es exclusivo de la perla. Como se muestra en la figura 16, después de la segunda reacción de tagmentación, cada fragmento tiene un extremo libre que porta una secuencia cebadora P7 y un extremo inmovilizado que porta una secuencia cebadora P5.

#### Ejemplo 1

##### Tagmentación en superficie en una cubeta de lectura

Este ejemplo describe un experimento que confirma la realización ilustrada en las figuras 4a y 4b (tagmentación en superficie seguida de tagmentación en solución).

Se llevó a cabo un experimento en una cubeta de lectura de 8 carriles usando una lista de condiciones y controles como se muestra en las figuras 6a y 6b. El carril 6 indica la exposición completa de extremo a extremo del método: se añadió ADN genómico de *E. coli* no fragmentado a una cubeta de lectura, tras lo cual se tagmentó mediante los

transposomas unidos a la superficie. A continuación, se aplicó calor a la cubeta de lectura para deshibridar selectivamente las secuencias de ME no marcados, eliminándolas así como objetivo para una segunda reacción de tagmentación cuando a continuación se añadió el transposoma de la solución. Tras la segunda reacción de tagmentación, se generaron grupos y se realizó una secuenciación de “extremos emparejados” (2 x 36 lecturas de bases). Las métricas de secuenciación se muestran en la diapositiva 7, que muestra que en el carril 6, el 73,14 % de los grupos pasaron los filtros y el 74,69 de estos se alinearon. Esto proporciona datos suficientes para obtener una gráfica del tamaño de los huecos de los insertos y una estimación de la diversidad de la biblioteca (véase la figura 8).

Otros carriles incluían controles que se enumeran a continuación:

El **carril 1** incluye ADN de PhiX como control positivo para garantizar que todo se haya bombeado correctamente y que la generación y secuenciación de clústeres funcionen según lo esperado.

El **carril 2** es otro carril de control e ilustra la tagmentación en transposones unidos a la superficie en ausencia de ADN diana. La cubeta de lectura (FC) comprende una FC estándar de extremos apareados en la que se hibridó un oligonucleótido que contenía la secuencia cebadora de tagmentación y la secuencia ME' (cadena ME' no transferida) con el oligonucleótido de superficie P7. Este oligonucleótido se añadió a una concentración de saturación. La primera extensión dio lugar a un transposón bicatenario con extremos ME. Se ensambló un transposoma P5 en solución y se hizo fluir hacia la FC (6,25 nM/carril). El transposoma P5 tagmenta los transposones bicatenarios unidos a la superficie. Los productos de tagmentación se convierten posteriormente en grupos.

El **carril 3** ilustra la tagmentación en transposones unidos a la superficie en ausencia de ADN diana, solo que en este caso la FC se ha calentado a 75 °C para convertir los transposones bc unidos a la superficie en oligonucleótidos monocatenarios antes de la adición del transposoma P5 de la solución para evitar la tagmentación en estas construcciones, ya que esto interferiría con la tagmentación del ADN diana.

El **carril 4** ilustra la adición de un transposoma unido a la superficie a las condiciones del carril 3. En este carril, un oligonucleótido que contenía la secuencia del cebador de tagmentación y la secuencia ME' se hibridó con el oligonucleótido de superficie P7. La primera extensión da lugar a un transposón bicatenario con extremos ME. Después de esto, se ensambló un transposoma P7 en la superficie de la FC añadiendo la enzima Tn5 al carril a una concentración de 50X e incubando a 30 °C durante 30 minutos. A continuación, el FC se calentó a 75 °C antes de añadir el transposoma P5 de la solución.

El **carril 5** comprende las mismas condiciones que el carril 4, solo que en este caso, en lugar de añadir el transposoma P5 de la solución, se añadió una biblioteca de 900 pb de *E. coli* (con extremos P5/P7) para determinar si el transposoma P7 unido a la superficie permanece activo después de la etapa de calentamiento.

El **carril 6** ilustra un ejemplo de la invención puesta en práctica. En este caso, la FC comprende una FC estándar de extremos apareados en la que se hibridó un oligonucleótido que contenía la secuencia cebadora de tagmentación y la secuencia ME' con el oligonucleótido de superficie P7. La primera extensión da lugar a transposones bicatenarios con extremos ME. Los transposomas P7 se ensamblan en la superficie de la FC añadiendo la enzima Tn5 al carril a una concentración de 50X e incubando a 30 °C durante 30 minutos. Se añadió ADN diana (300 ng de ADN genómico de *E. coli* no fragmentado) al carril de FC y se llevó a cabo una etapa de incubación de 15 minutos a 55 °C para que tuviera lugar la tagmentación. El transposoma P7 unido a la superficie se lavó usando PBI (Qiagen) y a continuación, la FC se calentó a 75 °C antes de añadir el transposoma P5 de la solución. Tras la adición del transposoma P5 de la solución y una etapa de incubación de 15 minutos a 55 °C, se llevó a cabo una reacción de extensión por desplazamiento de cadena para rellenar los huecos de 9 pb generados en la cadena principal del ADN por la reacción de transposición. La reacción de extensión por desplazamiento de cadena comprende la adición de una mezcla de Bst y dNTP y la incubación a 65 °C durante 5 min. El transposoma P5 se eliminó mediante lavado en la etapa final. En este punto, todas las moléculas unidas a la superficie deberían ser moldes P5-P7 y, por lo tanto, pueden convertirse en grupos.

El **carril 7** presenta las mismas condiciones que el carril 6, solo que en este caso se ha omitido la etapa de calentamiento para resaltar el efecto del calor en el número de conglomerados y el % de alineamiento.

El **carril 8** comprende un control negativo donde el transposoma P7 se ha ensamblado en la superficie a una concentración de 50X en presencia de una concentración saturada de transposones bc unidos a la superficie, sin embargo, no se ha añadido ningún ADN diana. Esto permite evaluar si el transposoma P7 unido a la superficie se tagmenta en sus transposones bc unidos a la superficie vecinos.

## Ejemplo 2

Preparación de muestras unidas a la superficie a partir de un raspado de *E. coli*

Se raspó una muestra de *E. coli* (raspado de 5 mm por 2 mm de un césped en una placa de agar) y se resuspendió en un tubo que contenía agua y perlas de vidrio. Las células y perlas suspendidas se mezclaron usando un mezclador vortical para romper las células y después se centrifugaron para sedimentar los residuos celulares. El sobrenadante (que contenía

el lisado celular, incluidas las proteínas y los ácidos nucleicos) se retiró y se añadió a una cubeta de lectura del Genome Analyzer (Illumina, Inc., San Diego, CA) que tenía transposomas inmovilizados según el protocolo descrito en el ejemplo 1.

La generación de grupos se realizó en la cubeta de lectura usando un dispositivo de preparación de muestras Cluster Station (Illumina, Inc., San Diego, CA). Después de la generación del grupo, se realizó un ciclo de secuenciación de extremos apareados con lecturas de 36 bases en cada dirección.

Para la lectura 1, el 58,99 % de los grupos pasaron los filtros y el 92,16 de estos se alinearon. Para la lectura 2, el 58,99 % de los grupos pasaron los filtros y el 55,08 de estos se alinearon. Estos datos confirman que se pueden añadir lisados celulares no purificados directamente a una cubeta de lectura que tenga transposomas inmovilizados con resultados de secuenciación sorprendentemente sólidos.

### Ejemplo 3

Este ejemplo describe métodos para evitar la tagmentación de dobles cadenas de oligonucleótidos unidos a la superficie cuando se añaden transposomas en fase de solución a una cubeta de lectura.

Un método para ensamblar transposomas en la superficie de una cubeta de lectura consiste en tomar una cubeta de lectura estándar de extremos apareados e hibridar un oligonucleótido “férula” contra los oligonucleótidos injertados en la superficie P5 y/o P7, formando un saliente extensible que se puede extender con una polimerasa para formar una doble cadena que contenga una secuencia ME bicatenaria. En esta etapa, la enzima transposasa se puede añadir para formar un transposoma funcional unido a la superficie (figura 5d). Si solo una parte de las dobles cadenas forman un transposoma, las dobles cadenas de ME “no marcadas” restantes pueden convertirse posteriormente en objetivos de tagmentación por parte de transposomas cercanos unidos a la superficie o transposomas añadidos de una solución. Además, los transposomas de superficie completamente ensamblados contienen partes de ADNbc cadena arriba de las secuencias ME que también pueden convertirse en objetivos para la tagmentación por parte de transposomas cercanos unidos a la superficie o transposomas añadidos desde la solución. Esta tagmentación no deseada se manifiesta en una reducción del porcentaje de grupos que pasan por los filtros de pureza que se alinean con el genoma objetivo.

Un ejemplo de este efecto se puede ver comparando los carriles 4 y 5 con los carriles 6 y 7 en la figura 18b. En los carriles 6 y 7, los transposomas se ensamblaron como se ha descrito anteriormente con los oligonucleótidos férula largos que, tras la extensión y el ensamblaje del transposoma, produjeron dobles cadenas unidos a la superficie que tienen al menos 50 bases bicatenarias. Tras su uso en una reacción de tagmentación en superficie (figura 18a) y la secuenciación posterior, solo el 22,59 y el 15,52 % de los grupos se alinearon con la *E. coli* diana, respectivamente (figura 18b).

Para evitar la producción de una población de grupos que contienen secuencias que no se alinean con el ADN genómico diana, el ensamblaje del transposoma se realizó como se indica en la figura 17. Primero se hibridó un oligonucleótido “férula” contra los oligonucleótidos injertados en la superficie P5 y/o P7, formando un saliente extensible que podía extenderse con una polimerasa para formar una doble cadena que contenía una secuencia ME bicatenaria. A continuación, el oligonucleótido férula se retiró mediante deshibridación y lavado, y a continuación se reemplazó hibridando un oligonucleótido más corto de tan solo 12 bases de longitud, pero preferiblemente de 16 bases de longitud, hasta el extremo 3' del oligonucleótido ME unido a la superficie. Se añadió transposasa para ensamblar un transposoma que no contenía ningún ADNbc expuesto. Todas las dobles cadenas “no marcadas” que no contenían una transposasa unida se eliminaron incubando la cubeta de lectura a 60 °C y lavándola en condiciones de flujo para deshibridar selectivamente y eliminar el ADN bicatenario corto.

El efecto beneficioso de este enfoque para ensamblar los transposomas se puede observar en los carriles 4 y 5, donde los transposomas unidos a la superficie se ensamblaron mediante este método y se usaron en una reacción de tagmentación en la superficie (figura 18a). El porcentaje de grupos que se alinearon con *E. coli* aumenta del 22,59 y el 15,52 % para los carriles 6 y 7, respectivamente, a 96,47 y 96,41 para los carriles 4 y 5, respectivamente (figura 18b).

El término que comprende se entiende en el presente documento como abierto, incluyendo no sólo los elementos citados, sino abarcando además cualquier elemento adicional.

Se ha descrito un número de realizaciones. No obstante, se entenderá que se podrán realizar diversas modificaciones. En consecuencia, otras realizaciones están dentro del alcance de las siguientes reivindicaciones.

## REIVINDICACIONES

1. Un soporte sólido que tiene complejos transposómicos inmovilizados sobre el mismo, en donde cada complejo transposómico comprende una transposasa unida de forma no covalente a un ácido nucleico bicatenario que comprende un primer y un segundo polinucleótido, en donde:

el primer polinucleótido comprende:

- (i) una parte 3' que comprende una secuencia terminal de transposón, y
- (ii) un primer marcador que comprende un primer dominio de marcaje;

el segundo polinucleótido comprende una región complementaria de la secuencia terminal del transposón;

al menos algunos de los complejos transposómicos son homodímeros, en donde los homodímeros comprenden una primera pluralidad de homodímeros que comprenden un primer polinucleótido de una primera secuencia y una segunda pluralidad de homodímeros que comprenden un primer polinucleótido de una segunda secuencia, en donde la primera secuencia es diferente de la segunda secuencia; y

los complejos transposómicos se inmovilizan en el soporte sólido mediante el primer polinucleótido de la primera secuencia o el primer polinucleótido de la segunda secuencia.

2. El sólido soporte de la reivindicación 1, en donde el primer dominio de marcaje comprende una región para la amplificación por grupos; o en donde el primer dominio de marcaje comprende una región para cebar una reacción de secuenciación; o en donde el complejo transposómico comprende una transposasa Tn5 hiperactiva; o en donde el soporte sólido comprende micropartículas, una superficie con patrones, pocillos, una superficie de una cubeta de lectura o una población de micropartículas distribuidas sobre una superficie sólida de una cubeta de lectura.

3. El soporte sólido de la reivindicación 1 o la reivindicación 2, en donde el soporte sólido es una pluralidad de perlas.

4. El soporte sólido de la reivindicación 3, en donde cada perla tiene inmovilizados sobre la misma cada uno de los dos primeros polinucleótidos.

5. El soporte sólido de la reivindicación 1 o la reivindicación 2, en donde el soporte sólido es una población de micropartículas y el primer polinucleótido se inmoviliza en su extremo 5' en la superficie de la micropartícula y el segundo polinucleótido se hibrida con el extremo 3' del primer polinucleótido.

6. El soporte sólido de la reivindicación 5, en donde cada micropartícula comprende una pluralidad de dichos complejos de transposasas; o en donde el primer polinucleótido comprende además un dominio índice asociado con la micropartícula; o en donde las micropartículas son perlas.

7. El soporte sólido de la reivindicación 6, en donde:

- (a) la población comprende una pluralidad de dominios índice en el primer polinucleótido; o
- (b) las perlas son paramagnéticas.

8. El soporte sólido de una cualquiera de las reivindicaciones 1-7, en donde algunos de los complejos transposómicos son heterodímeros que se inmovilizan en el soporte sólido a través de su primer polinucleótido.

9. Un método para generar una cubeta de lectura que comprende:

inmovilizar una pluralidad de complejos transposómicos en un soporte sólido, en donde el soporte sólido comprende una superficie de una cubeta de lectura o micropartículas distribuidas sobre una superficie de la cubeta de lectura,

en donde cada complejo transposómico comprende una transposasa unida a un primer polinucleótido y a un segundo polinucleótido, en donde:

el primer polinucleótido comprende:

- (i) una parte 3' que comprende una secuencia terminal de transposón, y
- (ii) un primer marcador que comprende un primer dominio de marcaje; y

el segundo polinucleótido comprende una región complementaria de la secuencia terminal del transposón;

al menos algunos de los complejos transposómicos son homodímeros, en donde los homodímeros comprenden una primera pluralidad de homodímeros que comprenden un primer polinucleótido de una primera secuencia y una segunda pluralidad de

- homodímeros que comprenden un primer polinucleótido de una segunda secuencia, en donde la primera secuencia es diferente de la segunda secuencia; y la inmovilización de la pluralidad de complejos transposómicos en el soporte sólido se realiza mediante el primer polinucleótido de la primera secuencia o el primer polinucleótido de la segunda secuencia.
- 5
10. El método de la reivindicación 9, en donde dicho dominio de marcaje comprende una región para la amplificación por grupos o una región para cebar una reacción de secuenciación.
- 10 11. El método de la reivindicación 9, en donde la inmovilización comprende:
- (a)proporcionar un soporte sólido que tiene una pluralidad de los primeros polinucleótidos inmovilizados en el mismo, y
- (b)poner en contacto el soporte sólido con una holoenzima transposasa y el segundo polinucleótido.
- 15
12. El método de la reivindicación 11, en donde la puesta en contacto comprende:
- hibridar el primer polinucleótido con el soporte sólido y después poner en contacto el soporte sólido con la holoenzima transposasa con el soporte sólido; o
- 20 poner en contacto la holoenzima transposasa y el segundo polinucleótido con el soporte sólido simultáneamente.
13. El método de la reivindicación 9, en donde la inmovilización comprende:
- (a)proporcionar un soporte sólido que tiene cebadores de amplificación acoplados al mismo;
- (b)hibridar el segundo polinucleótido con uno de dichos cebadores de amplificación, en donde el segundo oligonucleótido comprende una región complementaria del primer marcador;
- (c)extender el cebador de amplificación usando una polimerasa para generar una doble cadena que comprende el primer polinucleótido hibridado con el segundo polinucleótido, el primer polinucleótido inmovilizado directamente en el soporte sólido; y
- 30 (d)poner en contacto el soporte sólido con una holoenzima transposasa, ensamblando así un complejo transposómico sobre el soporte sólido.
14. Un método para generar una biblioteca de fragmentos de ADN marcados para su ensamblaje dirigido por índices en una lectura de secuencia más larga que comprende:
- 35 aplicar un ADN diana al soporte sólido de la reivindicación 1;
- en donde el soporte sólido es una población de micropartículas; y
- el primer polinucleótido comprende un dominio índice;
- 40 generando de esta manera fragmentos de ADN inmovilizados que están marcados con el dominio índice.
15. El método de la reivindicación 14, en donde el primer polinucleótido se inmoviliza en su extremo 5' en la superficie de la micropartícula, el segundo polinucleótido se hibrida con el extremo 3' del primer polinucleótido, la población de micropartículas comprende una pluralidad de dominios índice y los primeros polinucleótidos de una micropartícula individual comparten el mismo dominio índice; o en donde el primer polinucleótido comprende un dominio de escisión para liberar dicho primer polinucleótido de dicha micropartícula, una región para la amplificación por grupos o una región para cebar una reacción de secuenciación; o en donde la población de micropartículas se distribuye sobre una superficie sólida.
- 45
16. El método de la reivindicación 15, en donde la superficie sólida comprende una matriz de pocillos.
17. El método de la reivindicación 16, en donde una pluralidad de dichos pocillos comprende una micropartícula por pocillo.
- 55
18. El método de la reivindicación 14, que comprende además:
- realizar una reacción de secuenciación para determinar la secuencia del dominio índice y al menos una parte del fragmento de ADN marcado.
- 60
19. El método de la reivindicación 18, que comprende además:
- (a)ensamblar secuencias de ADN que comparten dominios índice comunes, generando de esta manera una lectura de secuencia más larga que comprende una pluralidad de subfragmentos; o
- (b)determinar los SNP de fase en dicha lectura de secuencia.
- 65
20. Un método de secuenciación de una pluralidad de moléculas de ADN diana que comprende:

- 5 aplicar una pluralidad de moléculas de ADN diana al soporte sólido de la reivindicación 1 en condiciones mediante las que las moléculas de ADN diana se fragmentan mediante los complejos transposómicos, produciendo de esta manera una biblioteca inmovilizada de fragmentos bicatenarios, en donde una primera parte de cada molécula de ADN diana se une al soporte sólido en una primera ubicación del soporte sólido y una segunda parte de cada molécula de ADN diana se une al soporte sólido en una segunda ubicación del soporte sólido;
- 10 cartografiar la biblioteca inmovilizada de fragmentos bicatenarios para generar un conjunto de ubicaciones que están unidas por cada molécula de ADN diana;
- determinar las secuencias de la primera y segunda partes de las moléculas de ADN diana; y correlacionar el conjunto de ubicaciones para determinar qué primera y segunda partes están unidas por las moléculas de ADN diana y para determinar la secuencia de la pluralidad de moléculas de ADN diana.
- 15 21. El método de la reivindicación 20, en donde la determinación comprende escindir los fragmentos bicatenarios inmovilizados para liberar un extremo libre; o en donde el cartografiado comprende generar un mapa físico de los fragmentos bicatenarios inmovilizados; o en donde el soporte sólido comprende una cubeta de lectura.
- 20 22. El método de la reivindicación 21, en donde:
- (a) la escisión se realiza escindiendo un sitio escindible o poniendo en contacto los fragmentos bicatenarios inmovilizados con un complejo de transposasa en fase de solución;
- (b) la generación comprende obtener imágenes del ADN para establecer la ubicación de las moléculas de ADN inmovilizadas en la superficie sólida; y/o
- 25 (c) la cubeta de lectura comprende un nanocanal.
23. El método de la reivindicación 22, en donde:
- (a) el sitio escindible comprende un sitio de enzimas de restricción o nucleótidos modificados; o
- 30 (b) la obtención de imágenes comprende añadir un agente de obtención de imágenes, un marcador detectable y/o un colorante intercalante al soporte sólido y detectar una señal del agente de obtención de imágenes.

35

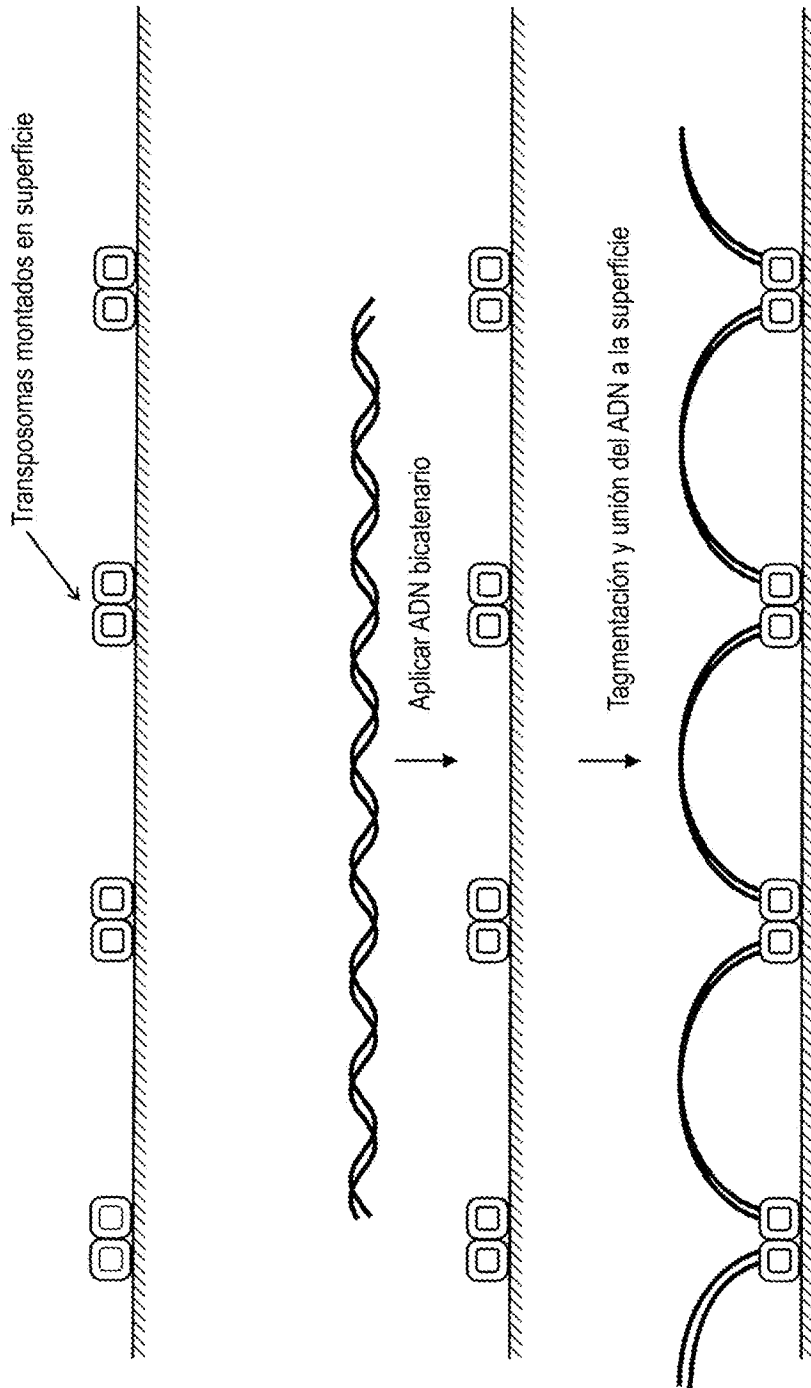


Figura 1a



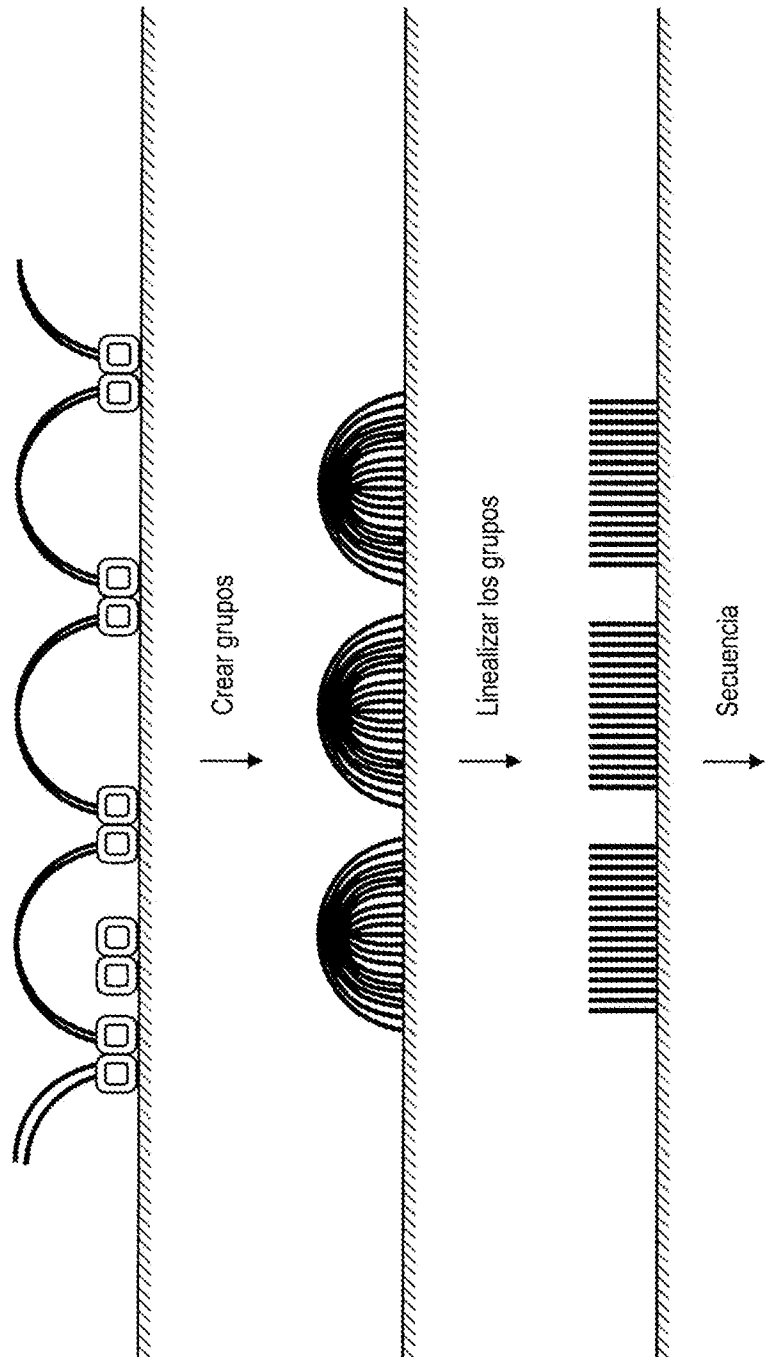


Figura 1b

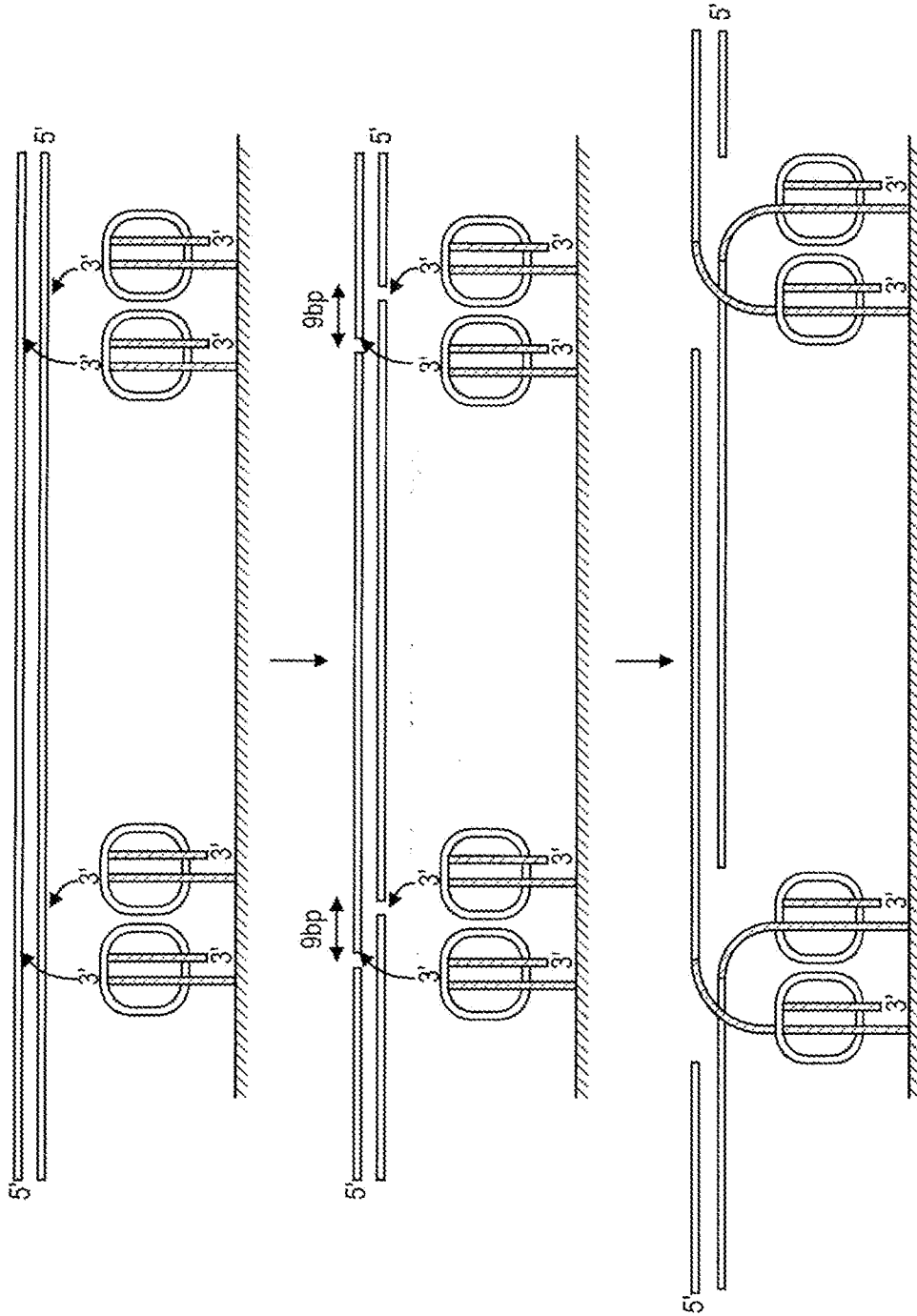


Figura 2a

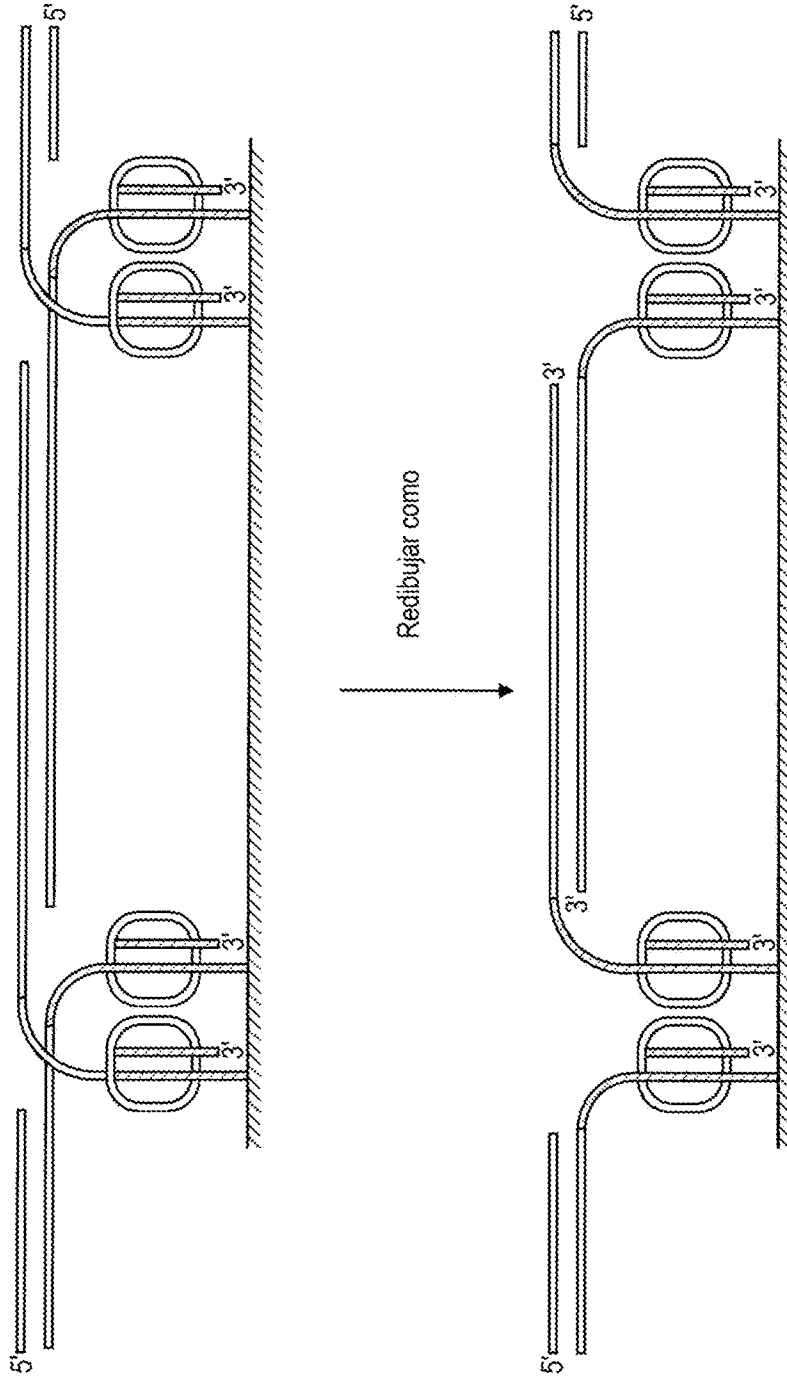


Figura 2b

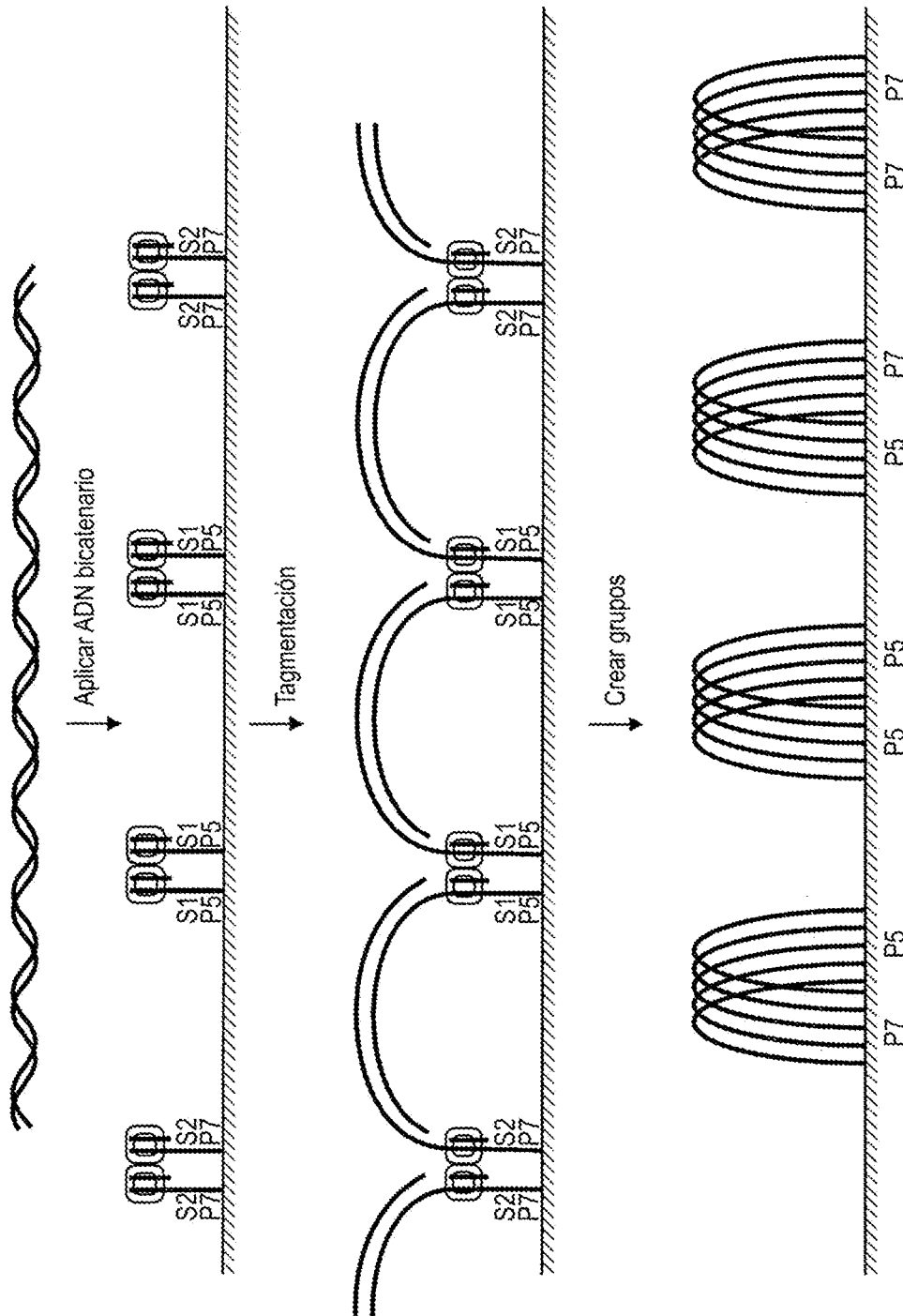


Figura 3a

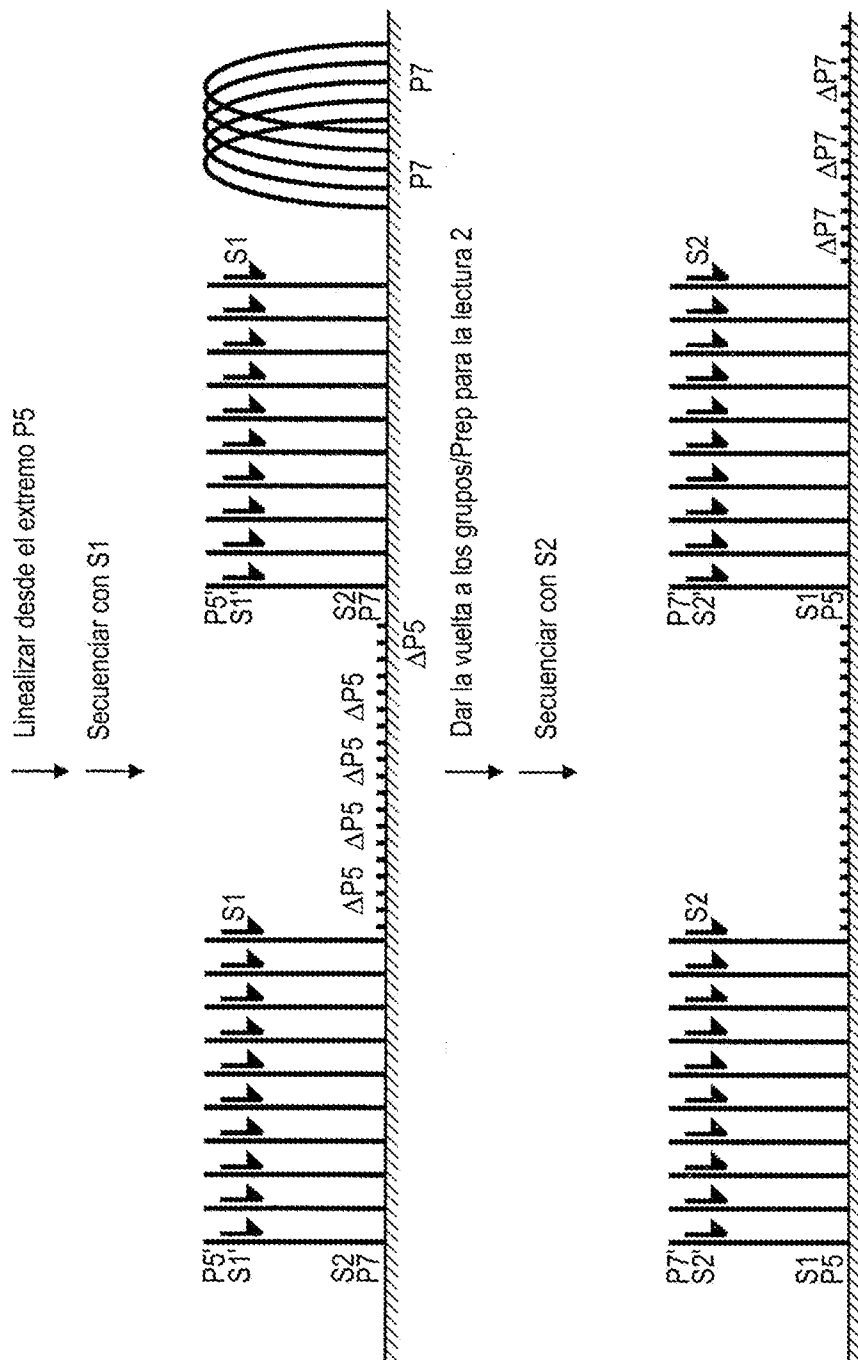


Figura 3b

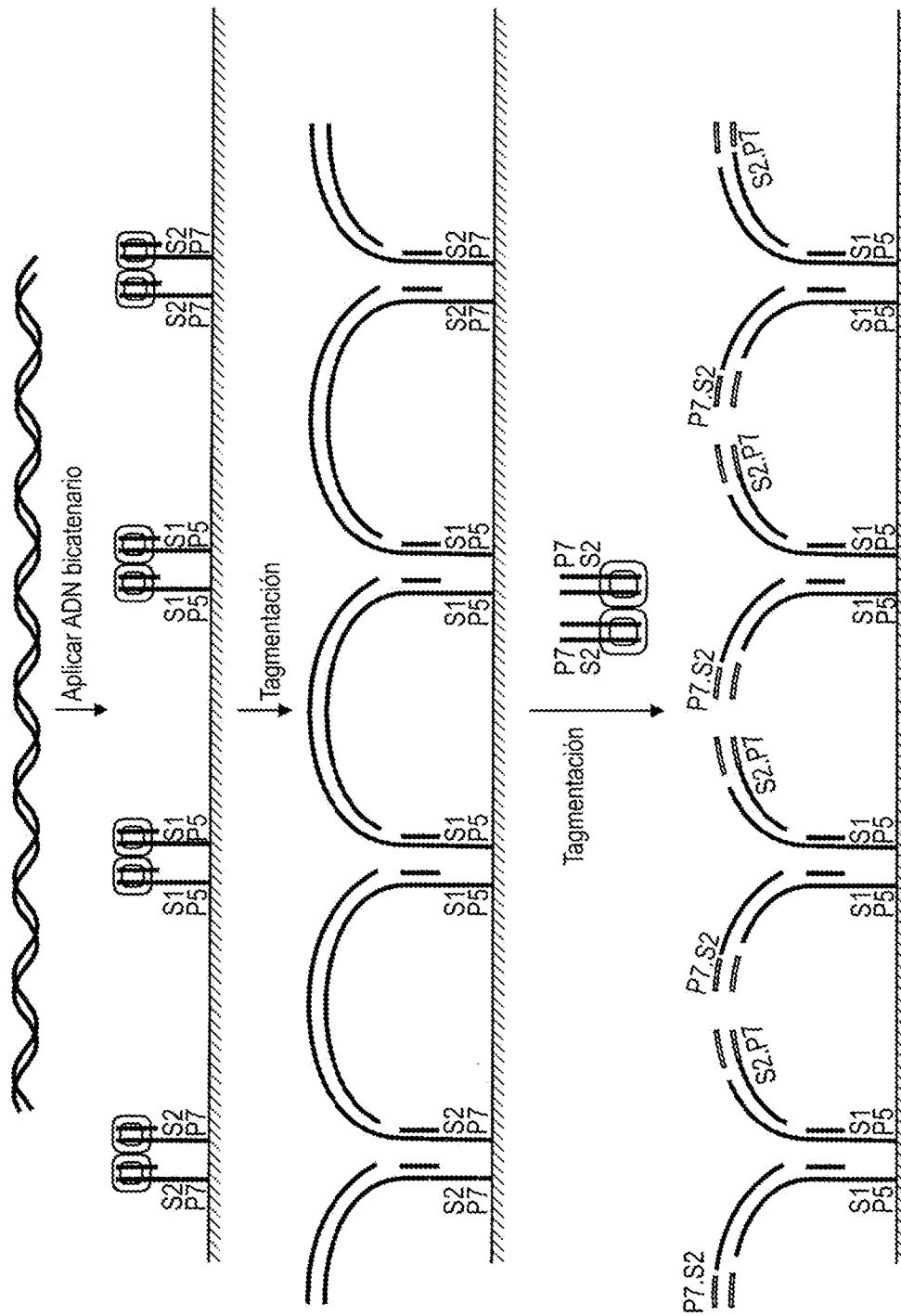
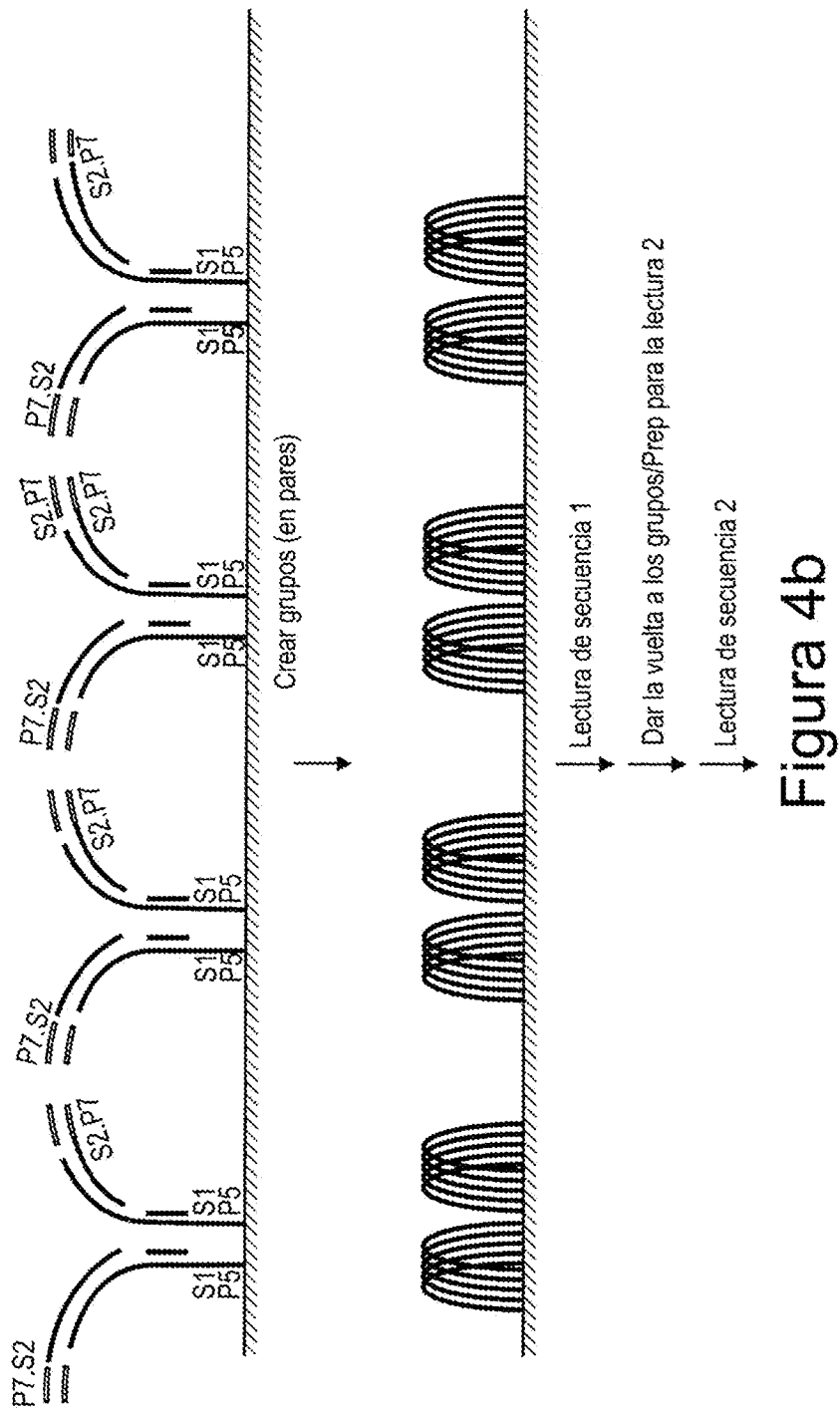


Figura 4a



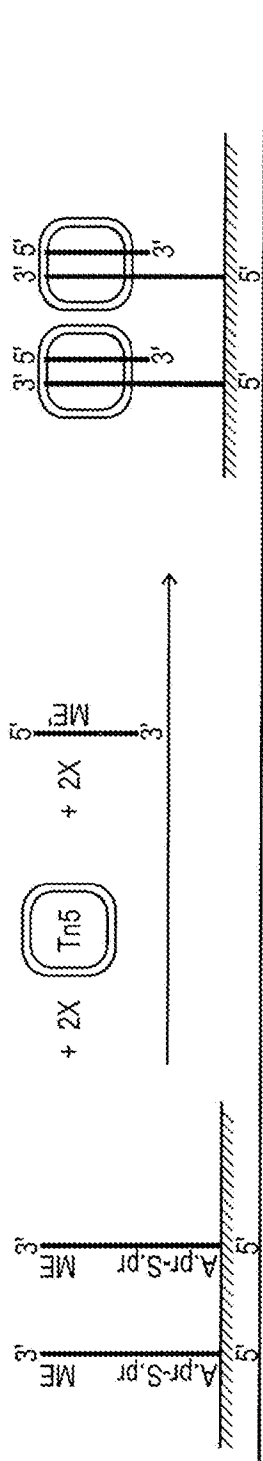


Figura 5a

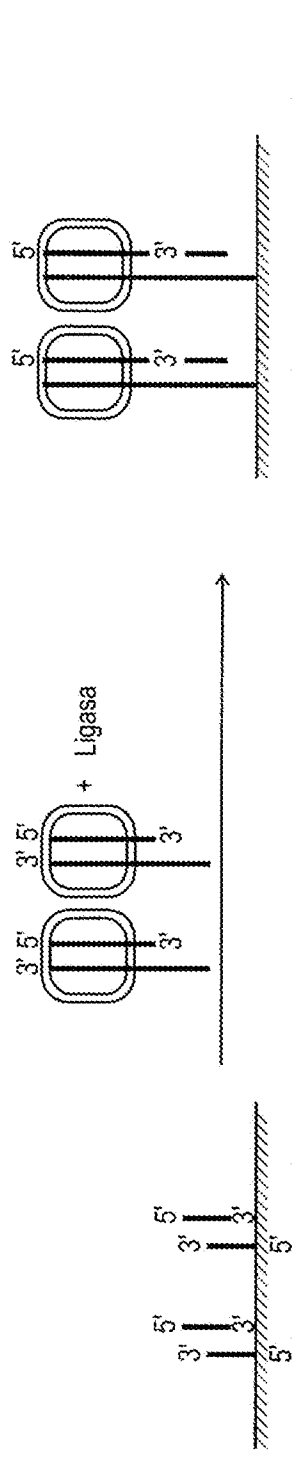


Figura 5b

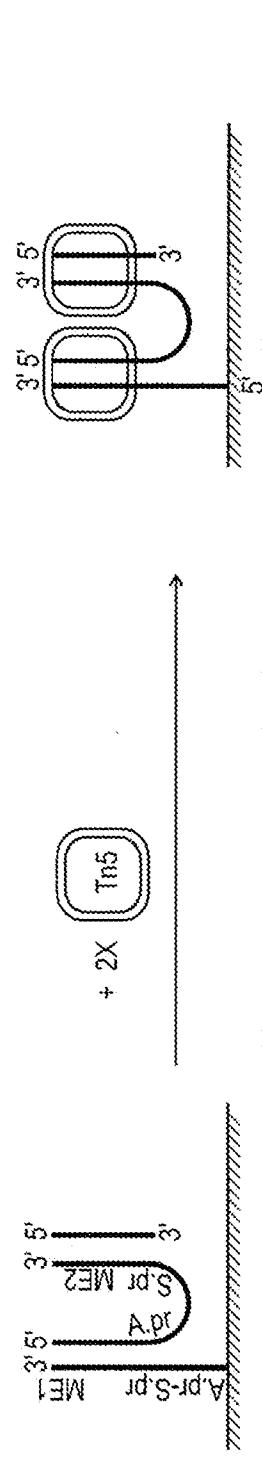


Figura 5c



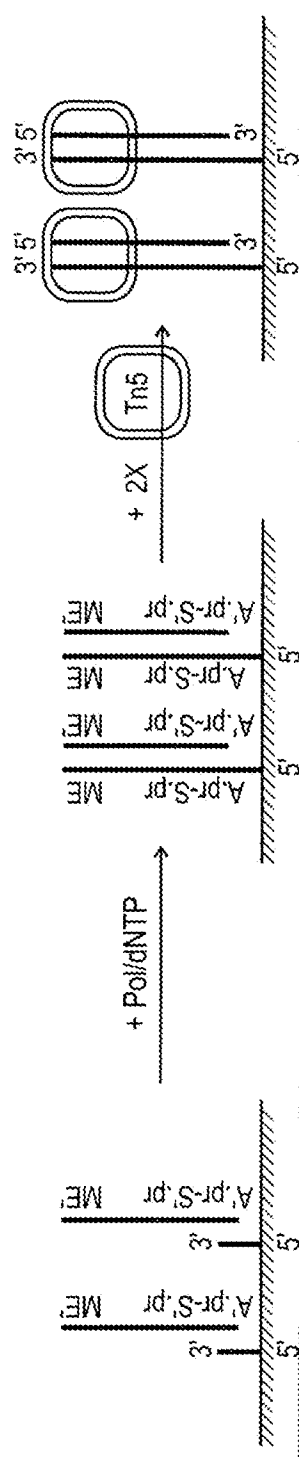


Figura 5d

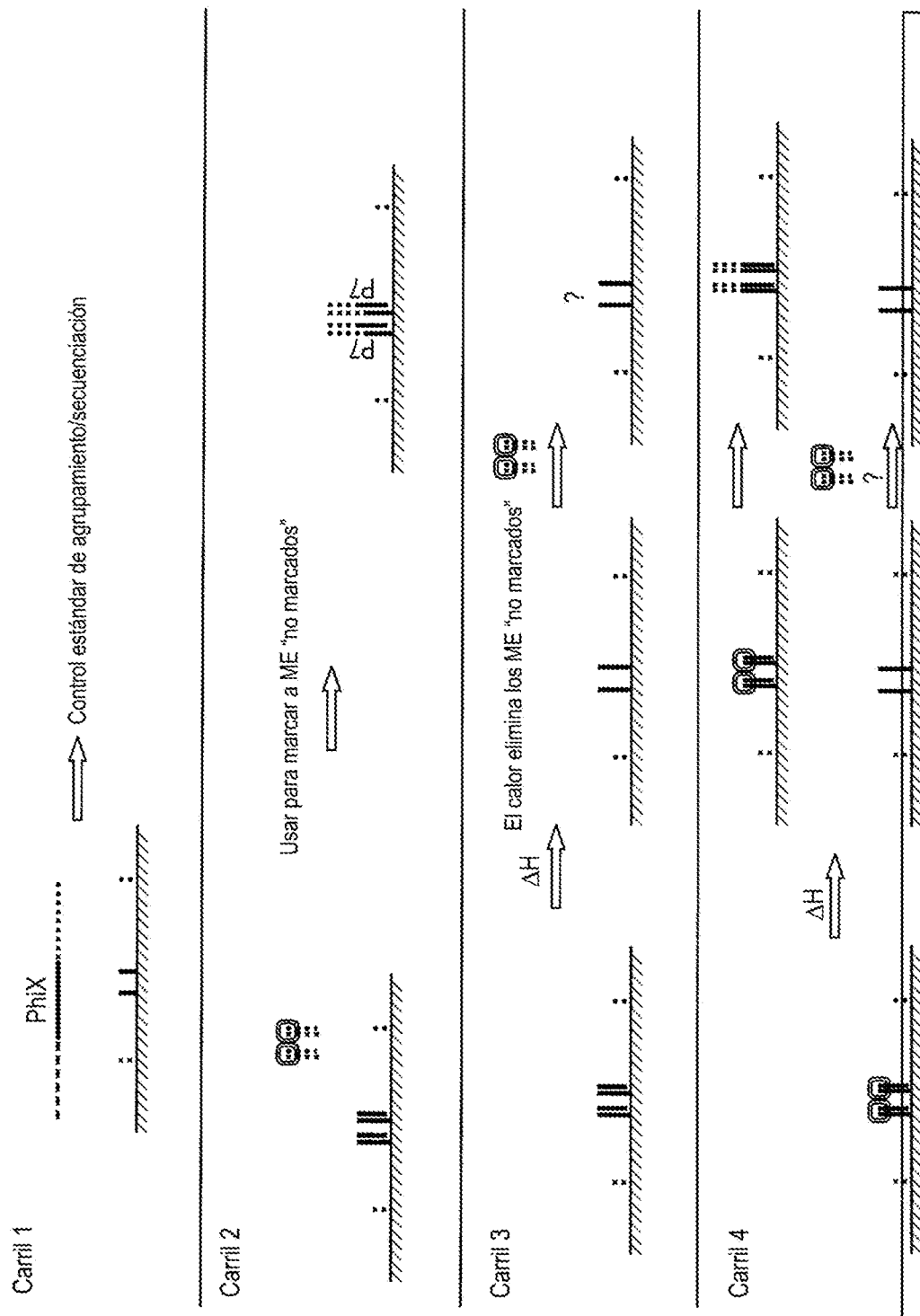
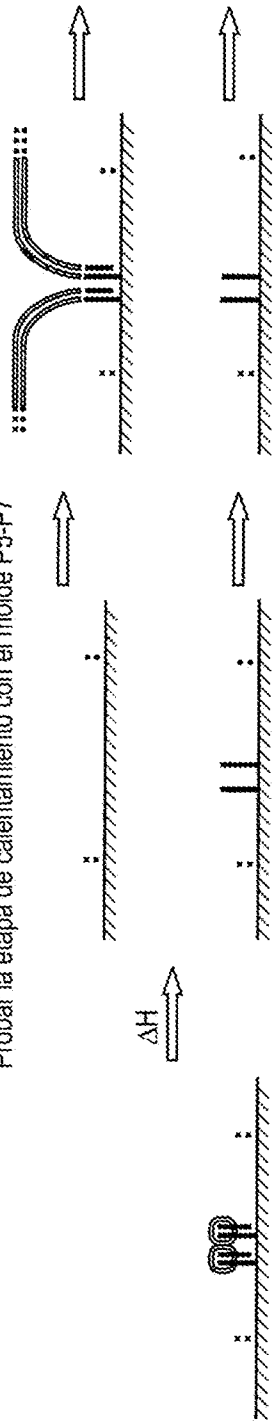


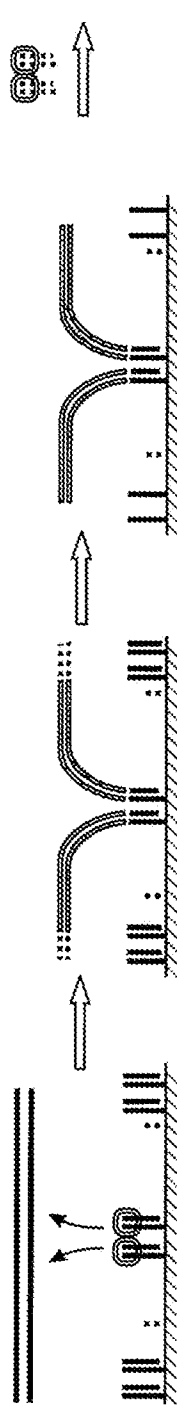
Figura 6a

Carril 5

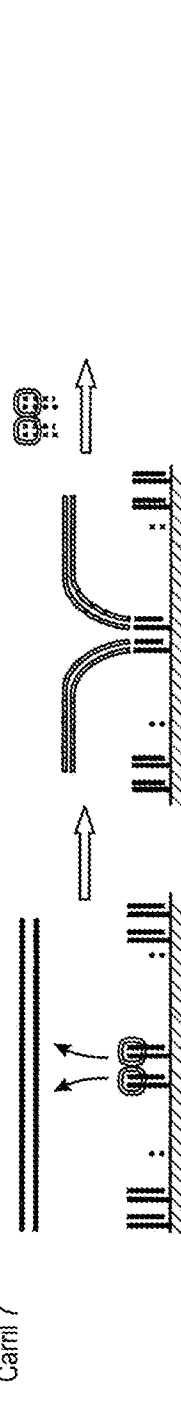
Probar la etapa de calentamiento con el molde P5-P7



Carril 6



Carril 7



Carril 8

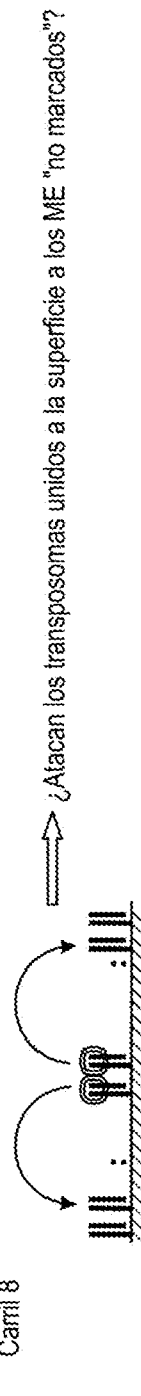
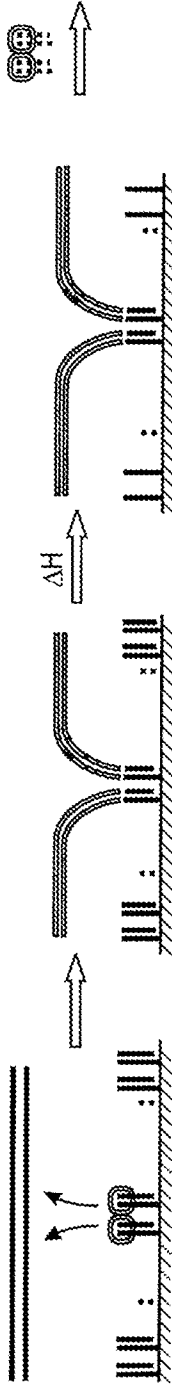


Figura 6b

(Estado del carril 6)



	Rendimiento de la muestra (Mbases)	Grupos (sin procesar)	Grupos (PF)	Int de 1. <sup>er</sup> ciclo (PF)	% de intensidad después de 20 ciclos (PF)	% PF de los grupos	% de alineación (PF)	% tasa de desajuste (PF)	% >=Q30 bases (PF)	Puntuación media de calidad (PF)
carril1	87	2.640.647	2.494.533	335	80,44	94,47	98,29	0,11	99,02	39,38
carril2	142	4.669.765	4.049.565	199	78,47	86,72	0		95,18	37,56
carril3	25	1.012.657	728.288	236	76,56	71,92	0		87,69	35,49
carril4	13	1.660.050	368.331	38	88,69	22,19	0		66,92	28,75
carril5	0	1.651.305	2.790	104	89,81	0,17	4,23	0,97	33,3	18
carril6	51	1.992.375	1.457.197	134	83,8	73,14	74,69	0,35	90,27	36,1
carril7	89	3.508.504	2.545.003	128	80,46	72,54	27,62	0,6	90,51	36,09
carril8	0	1.191.904	6.629	107	73,09	0,56	0		11,23	7,43

Lectura 1

	Rendimiento de la muestra (Mbases)	Grupos (sin procesar)	Grupos (PF)	Int de 1. <sup>er</sup> ciclo (PF)	% de intensidad después de 20 ciclos (PF)	% PF de los grupos	% de alineación (PF)	% tasa de desajuste (PF)	% >=Q30 bases (PF)	Puntuación media de calidad (PF)
carril1	87	2.640.647	2.494.533	96	80,99	94,47	96,86	0,19	95,71	37,93
carril2	142	4.669.765	4.049.565	15	105,29	86,72	0		1,97	2,96
carril3	25	1.012.657	728.288	18	105,22	71,92	0		2,38	3,16
carril4	13	1.660.050	368.331	13	92,27	22,19	0		2,18	3,18
carril5	0	1.651.305	2.790	37	18,39	0,17	0,43	1,43	5,51	5,01
carril6	51	1.992.375	1.457.197	32	119,21	73,14	17,45	2,26	8,55	7,34
carril7	89	3.508.504	2.545.003	26	115,87	72,54	4,96	3,31	6,08	5,77
carril8	0	1.191.904	6.629	81	72,94	0,56	0		5,83	5,40

Lectura 2

Figura 7

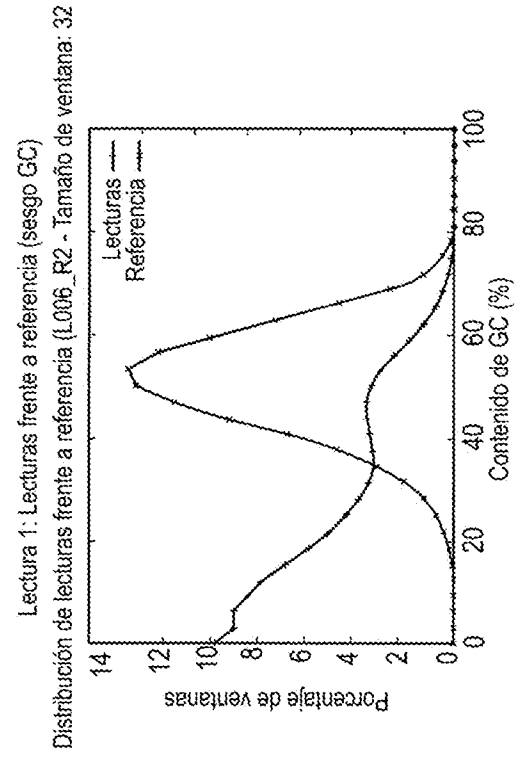
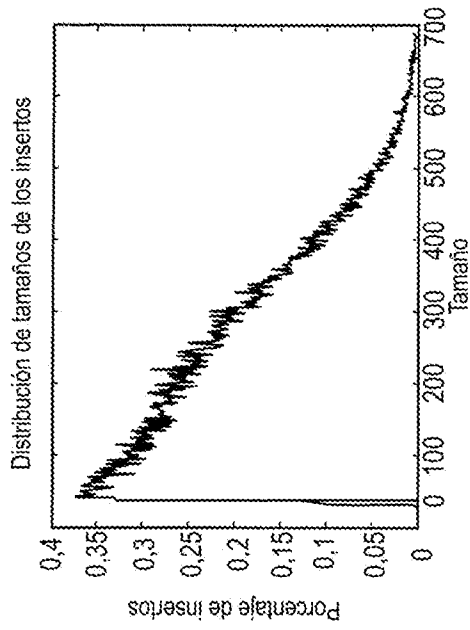
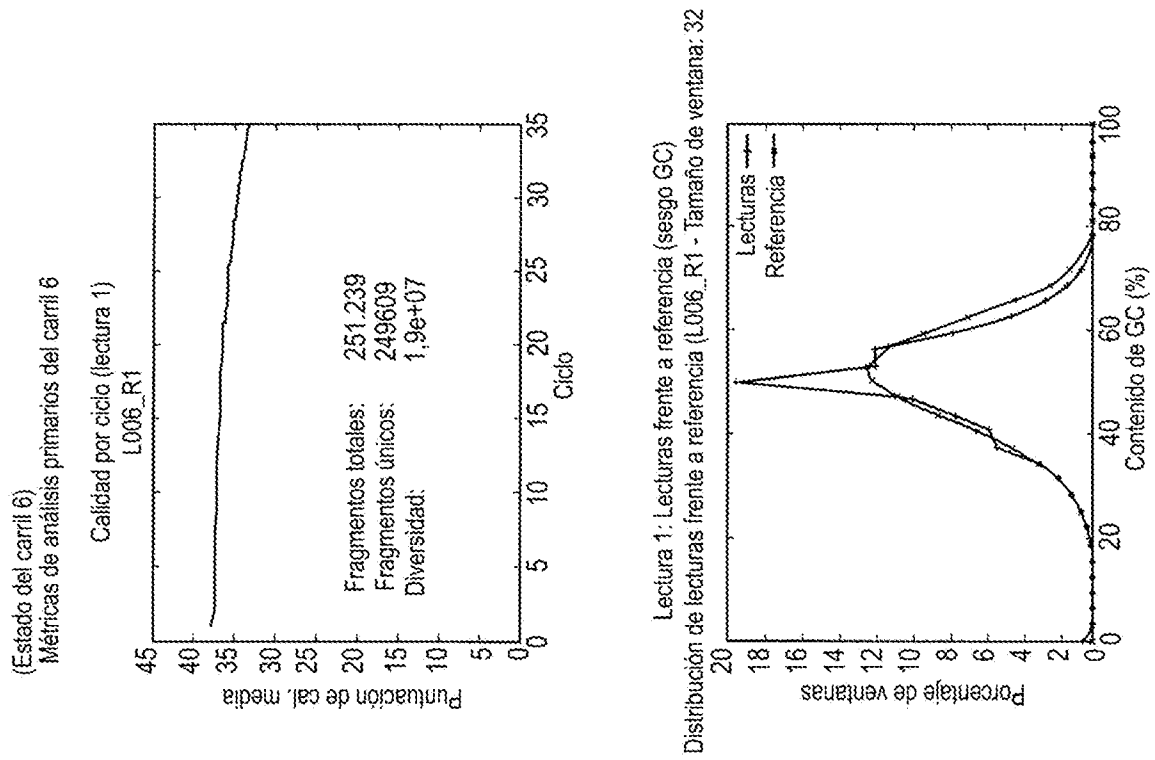


Figura 8

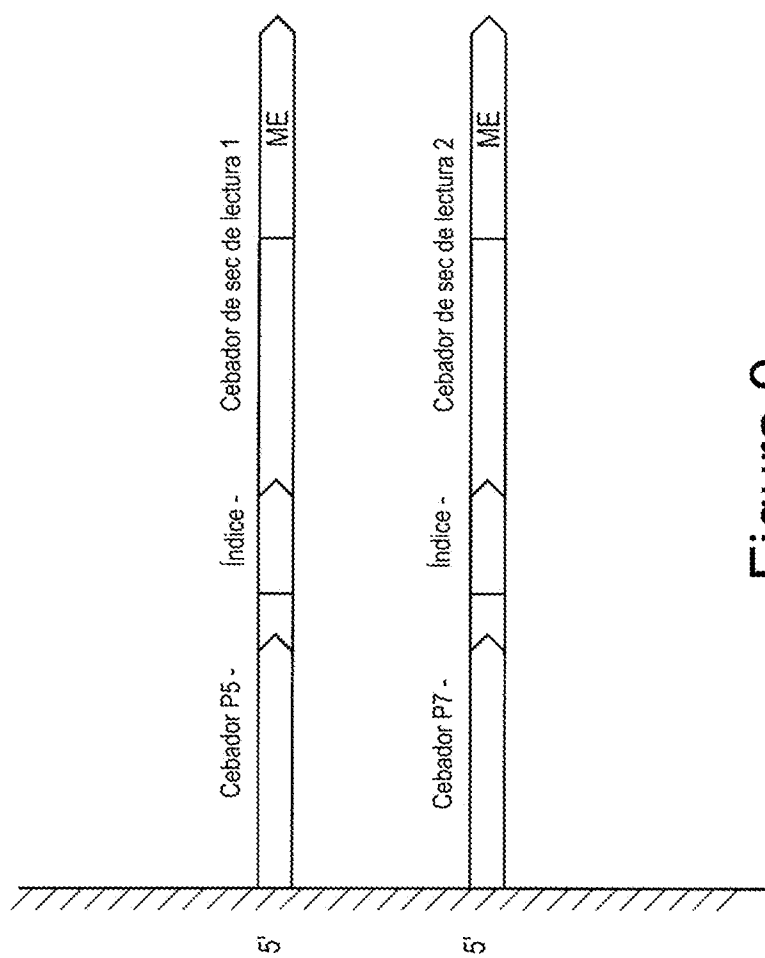


Figura 9

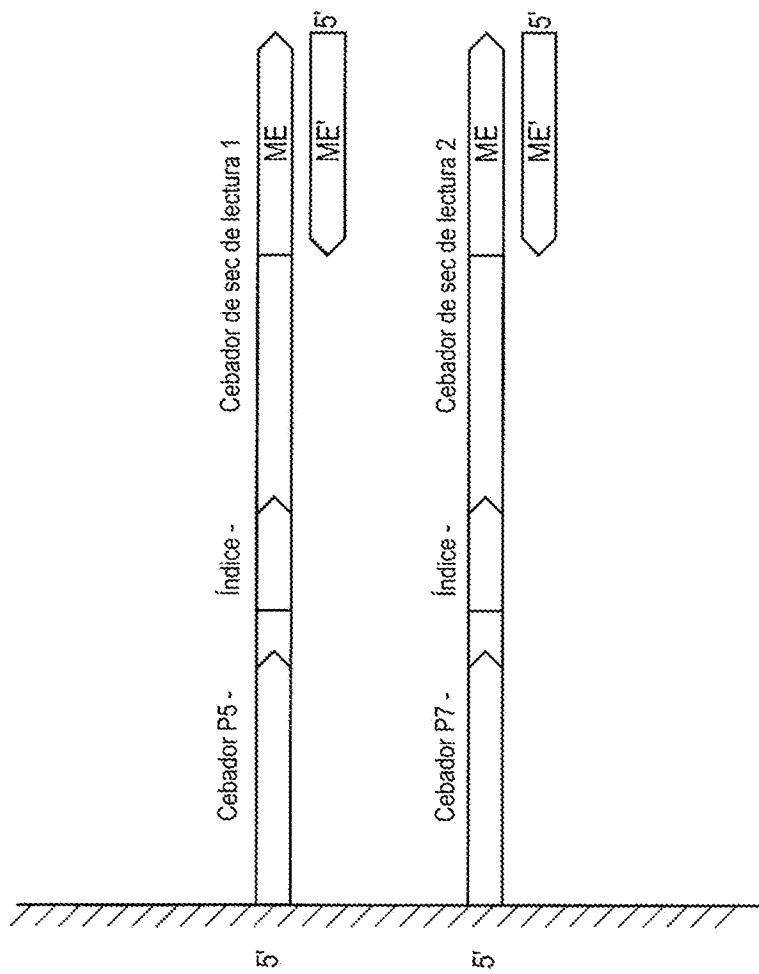


Figura 10

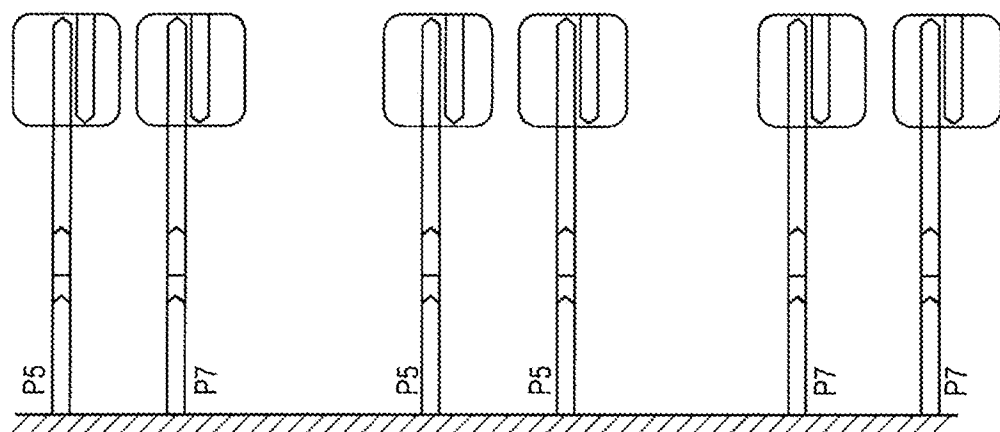


Figura 11



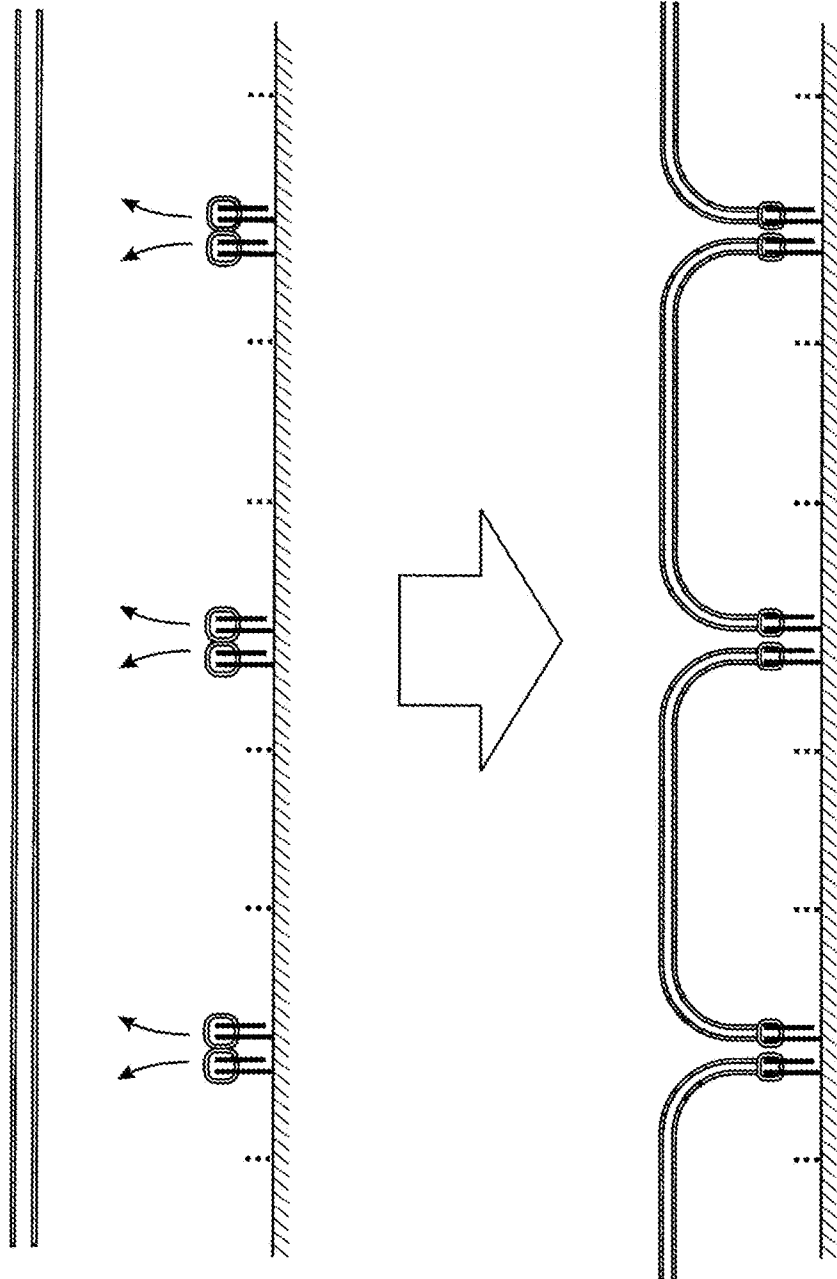


Figura 12

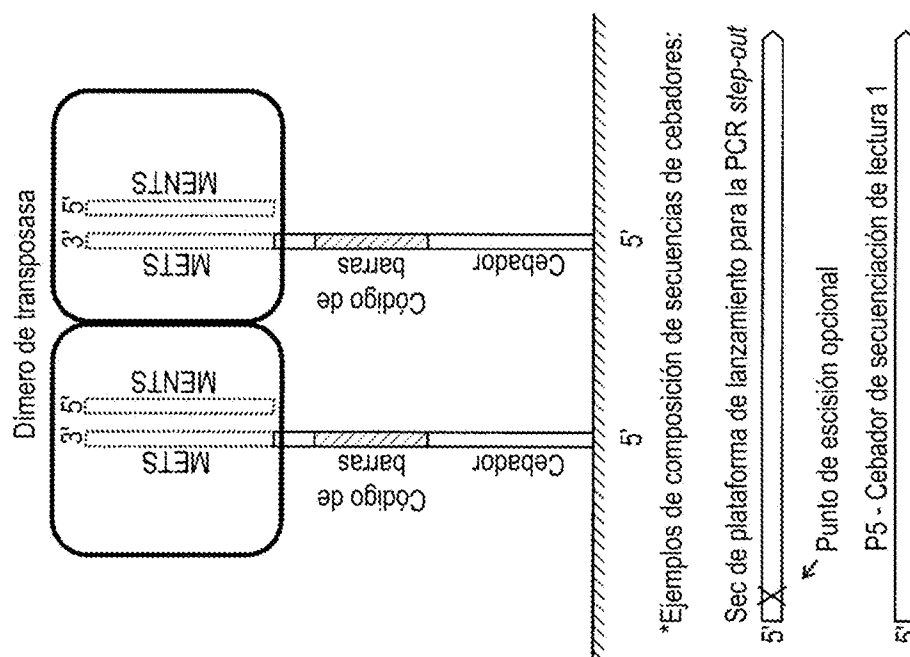


Figura 13

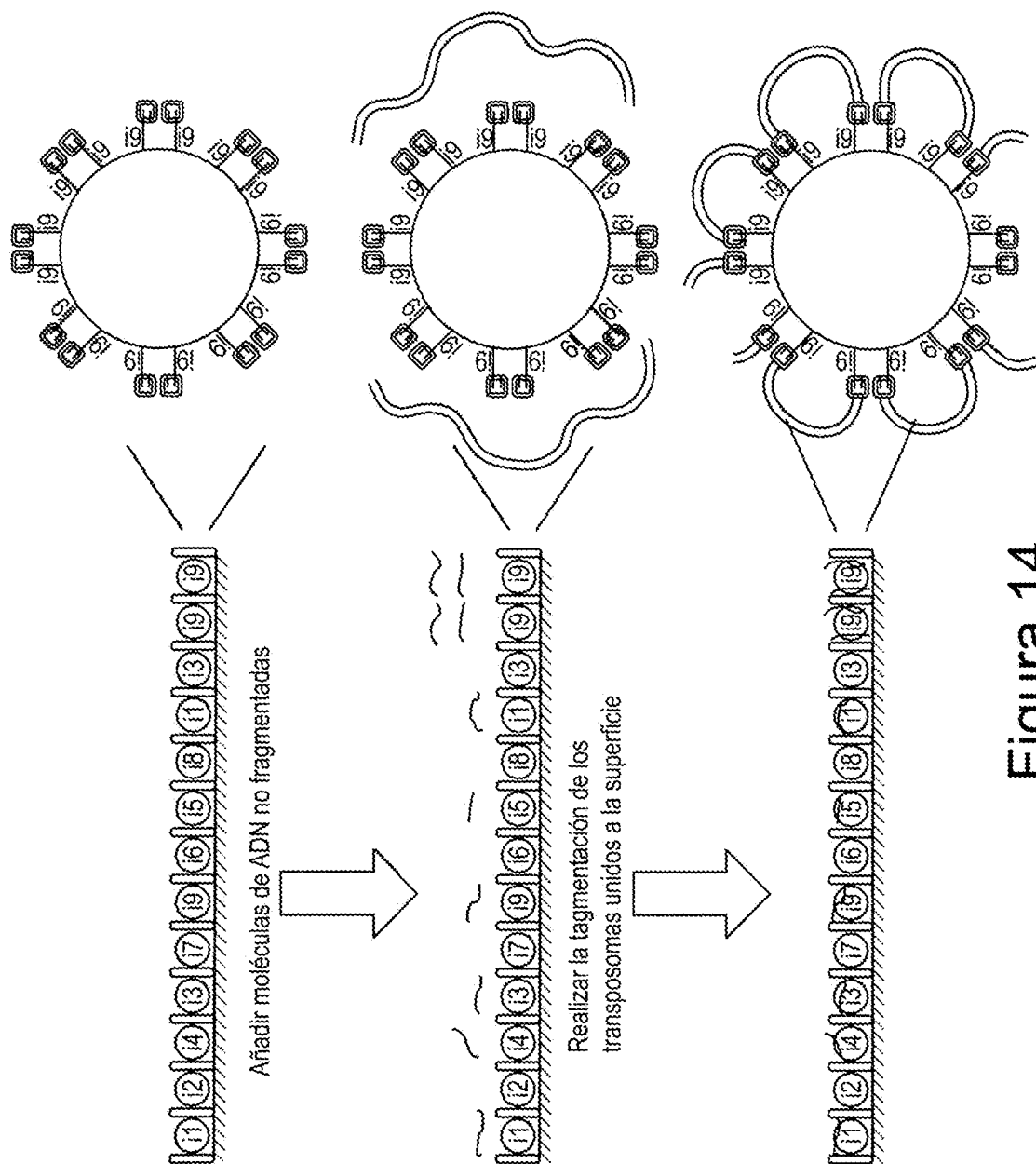


Figura 14

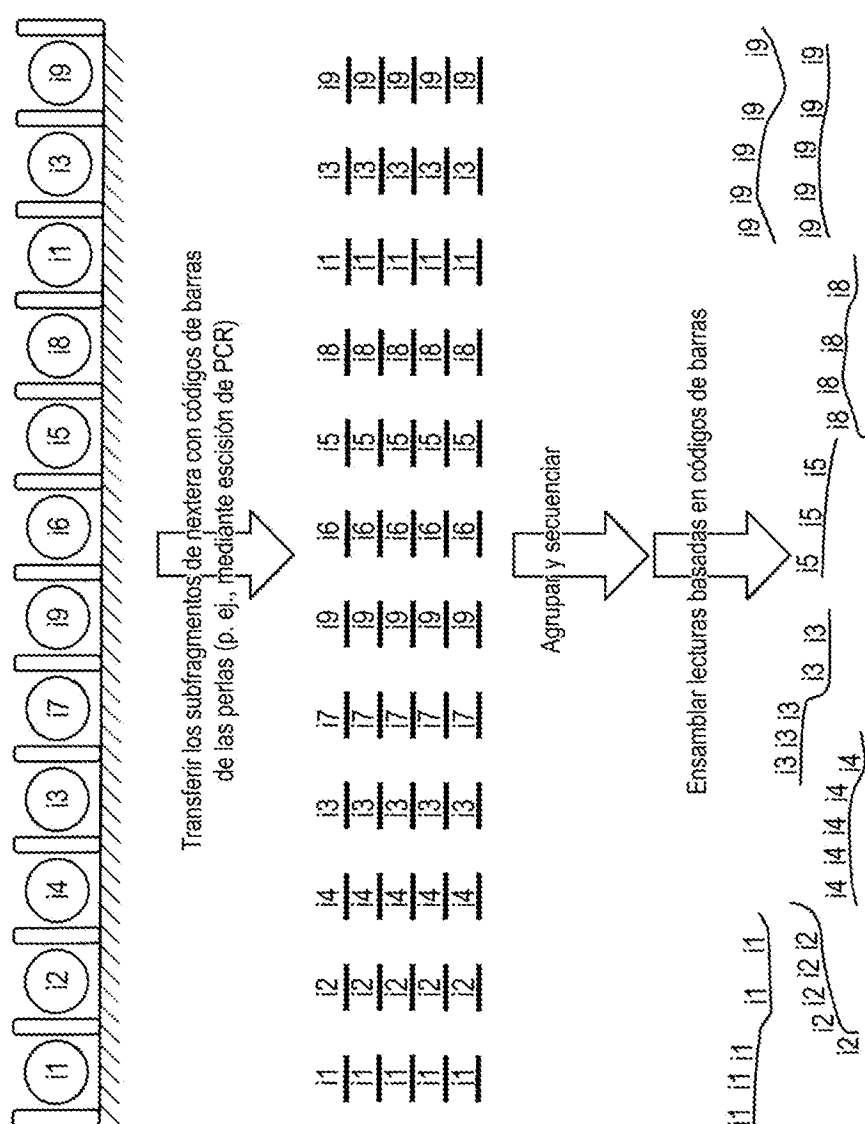


Figura 15

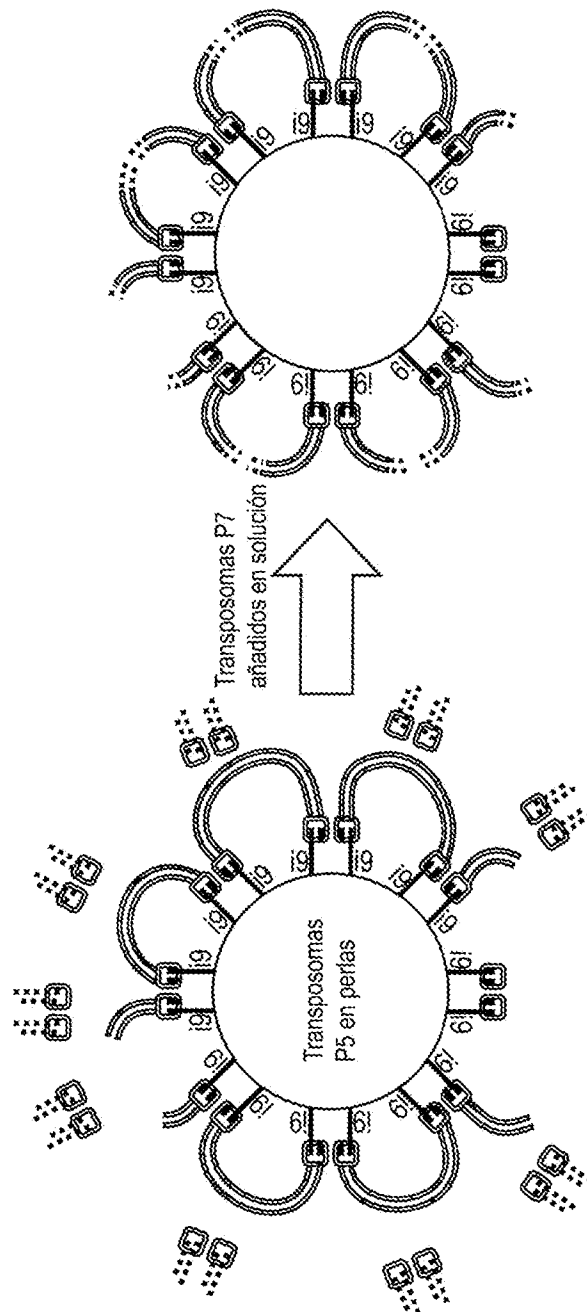


Figura 16

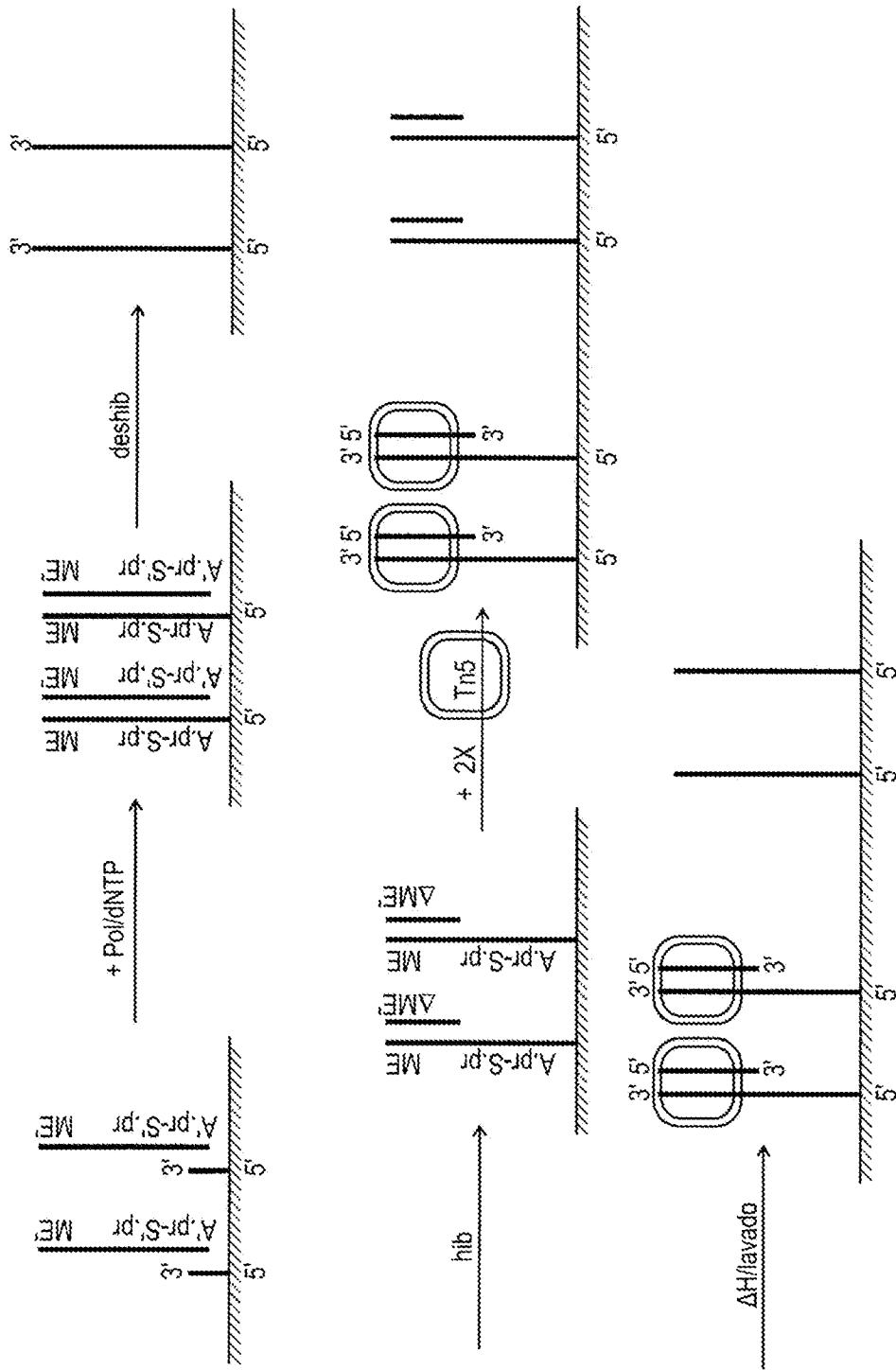


Figura 17

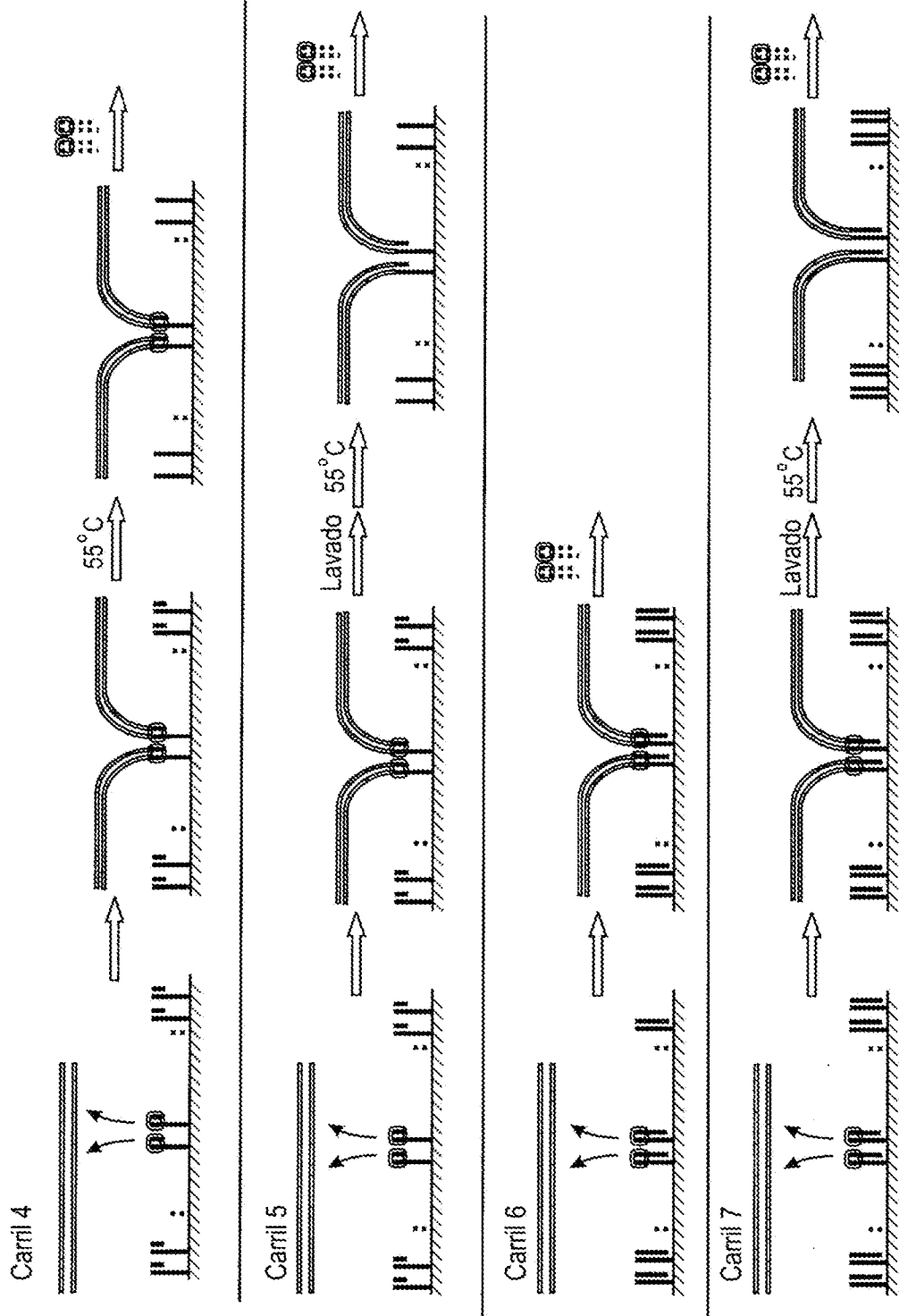


Figura 18a

	Rendimiento (Mb)	Grupos (sin procesar)	Grupos (PF)	Int de 1 <sup>er</sup> ciclo (PF)	% de intensidad después de 20 ciclos (PF)	% PF de los grupos	% de alineación (PF)	% tasa de desajuste (PF)	% >=0,30 bases (PF)	Puntuación media de calidad (PF)
carril1	204	6.172.822	5.821.910	226	84,42	94,32	98,24	0,04	98,75	39,16
carril2	178	5.676.854	5.096.648	170	84,02	89,78	93,81	0,08	97,78	38,74
carril3	261	8.915.997	7.450.306	168	82,73	83,56	89,9	0,13	95,42	37,7
carril4	89	2.963.760	2.542.838	120	85,56	85,8	96,47	0,11	95,99	38,14
carril5	96	3.223.900	2.728.667	103	83,66	84,64	96,41	0,12	95,22	37,82
carril6	109	3.801.294	3.103.977	105	85,96	81,66	22,59	0,11	95,21	37,69
carril7	123	4.239.062	3.513.343	112	85,31	82,88	15,52	0,17	95,64	37,82
carril8	0	2.230.575	0	0	0	0				

Lectura 1

Figura 18b