

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 973 302**

51 Int. Cl.:
C12Q 1/6869 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

- 86 Fecha de presentación y número de la solicitud internacional: **29.11.2019 PCT/EP2019/083175**
- 87 Fecha y número de publicación internacional: **18.06.2020 WO20120179**
- 96 Fecha de presentación y número de la solicitud europea: **29.11.2019 E 19816247 (1)**
- 97 Fecha y número de publicación de la concesión europea: **03.01.2024 EP 3894598**

54 Título: **Disminución del ajuste de fase con nucleótidos sin marcar durante la secuenciación**

30 Prioridad:
14.12.2018 US 201862779609 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
19.06.2024

73 Titular/es:
**ILLUMINA CAMBRIDGE LIMITED (100.0%)
19 Granta Park, Great Abington
Cambridge Cambridgeshire CB21 6DF, GB**

72 Inventor/es:
**GATTI-LAFRANCONI, PIETRO y
BALDING, PHILIP**

74 Agente/Representante:
DEL VALLE VALIENTE, Sonia

ES 2 973 302 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

Disminución del ajuste de fase con nucleótidos sin marcar durante la secuenciación

5 **Campo**

La presente descripción se refiere, entre otras cosas, a la secuenciación de polinucleótidos.

10 **Antecedentes**

10 La secuenciación de una cadena polinucleotídica molde puede producirse a través de múltiples ciclos de reacciones mediante los cuales se incorpora un nucleótido detectable por ciclo en una cadena copia. Los nucleótidos detectables de forma típica se bloquean para evitar la incorporación de más de un nucleótido detectable por ciclo. Después de un tiempo de incubación, de forma típica se realiza una etapa de lavado para retirar cualquier nucleótido detectable sin incorporar. Después puede realizarse una etapa de detección, en la que se determina la identidad del nucleótido detectable incorporado en la cadena copia. A continuación, se realiza una etapa de desbloqueo y una etapa de escisión o enmascaramiento en la que el agente bloqueador se retira del último nucleótido incorporado en la cadena copia y el resto detectable se escinde o enmascara en el último nucleótido incorporado en la cadena copia. En algunos casos, el resto detectable sirve como agente bloqueador, y la retirada del resto detectable puede retirar el agente bloqueador. Después se repite el ciclo introduciendo nucleótidos detectables en una etapa de incorporación.

25 En muchos casos, se secuencian simultáneamente grupos de cadenas polinucleotídicas molde que tienen la misma secuencia. Los grupos sirven para amplificar la señal producida por nucleótidos detectables incorporados en las cadenas copia. Debido a que los grupos contienen múltiples cadenas molde de la misma secuencia, el nucleótido incorporado en las cadenas copia correspondientes en cada ronda de adición de nucleótidos debe ser el mismo y la señal del nucleótido detectable debe potenciarse proporcionalmente al número de copias de la cadena molde en el grupo.

30 Un objetivo reciente de la secuenciación de polinucleótidos es disminuir el tiempo necesario para completar la secuenciación manteniendo al mismo tiempo una fidelidad elevada. Una forma de conseguir un menor tiempo de secuenciación es reducir el tiempo de ciclo acortando la duración de la etapa de incorporación. Se han desarrollado varias polimerasas más eficientes y nucleótidos modificados para proporcionar una incorporación más eficiente de nucleótidos en la cadena copia. Sin embargo, la etapa de incorporación todavía tiende a adolecer de una incorporación incompleta de nucleótidos en todas las cadenas molde que se secuencian.

35 Cuando, durante una etapa de incorporación, un nucleótido no se incorpora a una cadena copia en un grupo que contiene múltiples cadenas molde que tienen la misma secuencia, se dice que la cadena copia en la que no se incorpora el nucleótido está desfasada con aquellas cadenas copia en las que se incorpora el nucleótido durante la etapa de incorporación. A medida que el número de cadenas copia que están desfasadas aumenta con los ciclos adicionales, la señal de un grupo puede volverse demasiado heterogénea para determinar el nucleótido que se incorporó a las cadenas copia en fase. En GUO y col. (ACCOUNTS OF CHEMICAL RESEARCH, vol. 43, n.º 4, 20 de abril de 2010, páginas 551-563), el ajuste de fase se reduce incorporando nucleótidos sin marcar después de la detección de fluorescencia y antes del siguiente ciclo de extensión. En el documento US 2018/044715 A1, la etapa de sincronización puede producirse después o antes de la etapa de examen. En el documento US 2013/137091 A1, los nucleótidos marcados y sin marcar se incorporan al mismo tiempo y la lectura desfasada se corrige mediante desajuste de fase computacional.

40 **Resumen**

50 La presente descripción describe, entre otras cosas, métodos de secuenciación de polinucleótidos que permiten tiempos de ciclo cortos para la incorporación de nucleótidos marcados reduciendo al mismo tiempo el ajuste de fase. Los métodos incluyen incubar nucleótidos sin marcar con un grupo de cadenas polinucleotídicas molde que tienen la misma secuencia cuando se detecta la identidad del nucleótido marcado añadido anteriormente. La etapa de detección proporciona tiempo para que la adición de los nucleótidos sin marcar se incorpore en las cadenas copia en las que no se incorporó el nucleótido marcado añadido anteriormente. Por lo tanto, al final de la etapa de detección, todas o la mayor parte de las cadenas copia estarán en fase y listas para incorporar el nucleótido marcado adecuado en la etapa de incorporación posterior.

60 En algunas realizaciones, el método de secuenciación de polinucleótidos comprende introducir una enzima extensora de cadena y una mezcla de nucleótidos marcados y bloqueados en una célula de flujo que comprende un sitio en el que múltiples cadenas polinucleotídicas molde que tienen la misma secuencia de nucleótidos están unidas a una superficie de la célula de flujo. La enzima extensora de cadena está configurada para incorporar uno adecuado de los nucleótidos marcados y bloqueados en cadenas polinucleotídicas copia, basándose en la secuencia de las cadenas molde a las que corresponden las cadenas polinucleotídicas copia. El método comprende además retirar por lavado nucleótidos marcados y bloqueados sin incorporar de la célula de flujo e introducir una composición que comprende una mezcla de nucleótidos sin marcar y bloqueados en la célula de flujo durante o después de la retirada por lavado

de los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo. Un nucleótido sin marcar y bloqueado de la mezcla de los nucleótidos sin marcar y bloqueados está disponible para su incorporación en las cadenas polinucleotídicas copia, basándose en una secuencia de las cadenas molde a las que corresponden las cadenas polinucleotídicas copia, siempre que el nucleótido incorporado anteriormente en la cadena copia, si lo hay, no esté bloqueado. El método también comprende detectar la identidad de un nucleótido marcado y bloqueado incorporado en las cadenas copia, si lo hay, mientras la mezcla de nucleótidos sin marcar y bloqueados se incubaba con la célula de flujo.

Después, el marcador y el bloqueo pueden retirarse del nucleótido marcado y bloqueado incorporado en las cadenas copia y el bloque puede retirarse del nucleótido bloqueado y sin marcar, si lo hay, incorporado en las cadenas copia. Después, el proceso puede repetirse durante un número de ciclos predeterminado o hasta que se complete la secuenciación.

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.

Debe entenderse que tanto la descripción general anterior como la descripción detallada a continuación presentan realizaciones de la materia de la presente descripción y tienen previsto proporcionar una visión general o marco para comprender la naturaleza y el carácter de la materia de la presente descripción tal cual se reivindica. Las figuras adjuntas se incluyen para proporcionar una mayor comprensión de la materia de la presente descripción y se incorporan en y constituyen una parte de esta memoria descriptiva. Las figuras ilustran diversas realizaciones de la materia de la presente descripción y junto con la descripción sirven para explicar los principios y operaciones de la materia de la presente descripción. De forma adicional, las figuras y descripciones pretenden ser meramente ilustrativas y no pretenden limitar el alcance de las reivindicaciones de ninguna manera.

Descripción de los dibujos

La siguiente descripción detallada de las realizaciones específicas de la presente descripción puede entenderse mejor cuando se lee junto con los siguientes dibujos.

Las **FIG 1-2** son diagramas de flujo que ilustran realizaciones de los métodos de secuencias descritos en la presente memoria.

La **FIG. 3** es una vista en planta esquemática de una realización de una célula de flujo que puede emplearse según las enseñanzas presentadas en la presente memoria.

Las FIG. 4A-G son figuras esquemáticas que ilustran diversos ciclos de secuenciación y compensación para el ajuste de fase durante la etapa de barrido (detección).

La **FIG. 5** es una representación que ilustra la correlación entre el ajuste de fase (% de peso de ajuste de fase) y la tasa de error de PhiX (TE/%) usando un secuenciador Illumina MiniSeq™ con diferentes tiempos de incorporación (46 segundos, 23 segundos y 12 segundos).

La **FIG. 6** es una representación de tasas de error de PhiX acumuladas durante 80 ciclos de secuenciación en los que no se usaron nucleótidos sin marcar y bloqueados (mezcla de barrido convencional) durante una etapa de detección o se usaron nucleótidos sin marcar y bloqueados (mezcla ScanAndFill (BarrerYLLenar, en español)) durante la etapa de detección.

La **FIG. 7** es una representación de barras de tasas de ajuste de fase, preajuste de fase y % de Q30 durante 40 ciclos de secuenciación en los que se usaron nucleótidos sin marcar y bloqueados (ScanAndFill), con polimerasa adicional (Pol1671) o sin polimerasa adicional (Sin Pol), durante la etapa de detección.

La **FIG. 8** es una representación de la decadencia de la señal (P90Red) observada durante 100 ciclos de secuenciación en los que se usaron nucleótidos sin marcar y bloqueados (S&F) durante la etapa de detección en presencia y ausencia de ascorbato 3 mM, que protege el ADN del deterioro oxidativo provocado por el barrido óptico.

La **FIG. 9** es una representación de la tasa de error acumulada de PhiX (% de error, % de Q30, recuadro) observada durante 100 ciclos de secuenciación en los que se usaron nucleótidos sin marcar y bloqueados (S&F) durante la etapa de detección en presencia y ausencia de ascorbato 3 mM.

Los dibujos esquemáticos no son necesariamente a escala. Los mismos números que se utilizan en las figuras se refieren a los mismos componentes, pasos y similares. Sin embargo, se entenderá que el uso de un número para referirse a un componente en una figura dada no pretende estar limitado al componente en otra figura etiquetada con el mismo número. Además, el uso de números distintos para referirse a componentes no pretende indicar que los distintos componentes numerados no puedan ser iguales o similares a otros componentes numerados.

Descripción detallada

5 Ahora se hará referencia con mayor detalle a diversas realizaciones de la materia de la presente descripción, algunas realizaciones de las cuales se ilustran en los dibujos adjuntos.

10 Todos los términos científicos y técnicos utilizados en la presente memoria tienen los significados comúnmente utilizados en la técnica salvo que se indique lo contrario. Las definiciones proporcionadas en la presente memoria tienen por objeto facilitar la comprensión de determinados términos utilizados con frecuencia en la presente memoria y no pretenden limitar el ámbito de la presente descripción.

15 Como se utiliza en la presente memoria, las formas singulares “un”, “uno”, “una”, “el” y “la” incluyen referentes plurales salvo que el contexto indique claramente lo contrario. Por lo tanto, por ejemplo, la referencia a una “secuencia de polinucleótidos molde” incluye ejemplos que tienen dos o más de dichas “secuencias de polinucleótidos molde” salvo que el contexto indique claramente lo contrario.

20 Como se utiliza en esta memoria descriptiva y en las reivindicaciones adjuntas, el término “o” se emplea de forma general en su sentido que incluye “y/o” salvo que el contenido indique claramente lo contrario. El término “y/o” significa uno o todos los elementos enumerados o una combinación de cualquiera de dos o más de los elementos enumerados. El uso de “y/o” en algunos casos no implica que el uso de “o” en otros casos puede no significar “y/o”.

25 Como se utiliza en la presente memoria, “tienen”, “tiene”, “que tiene”, “incluyen”, “incluye”, “que incluye”, “comprenden”, “comprende”, “que comprende” o similares se utilizan en su sentido inclusivo abierto, y de forma general significan “incluyen, aunque no de forma limitativa”, “incluye, pero no se limita a”, o “que incluye, aunque no de forma limitativa”.

30 “Opcional” u “opcionalmente” significa que el evento, circunstancia o componente descrito posteriormente puede producirse o no, y que la descripción incluye instancias en las que el evento, circunstancia o componente se produce e instancias en las que no se produce.

35 Las palabras “preferido” y “preferiblemente” se refieren a realizaciones de la descripción que pueden proporcionar determinados beneficios, en determinadas circunstancias. Sin embargo, en las mismas circunstancias o en otras, también pueden preferirse otras realizaciones. Además, la mención de una o más realizaciones preferidas no implica que otras realizaciones no sean útiles y no pretende excluir otras realizaciones del ámbito de la tecnología de la invención.

40 Además, las menciones de intervalos numéricos en la presente memoria mediante extremos incluyen todos los números comprendidos dentro de ese intervalo (por ejemplo, 1 a 5 incluye 1, 1,5, 2, 2,75, 3, 3,80, 4, 5, etc.). Cuando un intervalo de valores es “mayor que”, “menor que”, etc. un valor particular, ese valor está incluido dentro del intervalo.

45 Salvo que se indique expresamente lo contrario, de ningún modo se pretende que las etapas de los métodos establecidos en la presente memoria deban llevarse a cabo en un orden específico. Por lo tanto, cuando una reivindicación de método no menciona realmente un orden a seguir para sus etapas o no se indica específicamente de otro modo en las reivindicaciones o descripciones que las etapas se deben limitar a un orden específico, de ningún modo se inferirá un orden particular. Sin embargo, se entenderá que un orden presentado es una realización de un orden mediante el cual puede realizarse el método. Cualquier característica o aspecto único o múltiple mencionado en cualquier reivindicación puede combinarse o permutarse con cualquier otra característica o aspecto mencionado en cualquier otra reivindicación o reivindicaciones.

50 Si bien pueden describirse diversas características, elementos o etapas de realizaciones particulares utilizando la expresión de transición “que comprende”, debe entenderse que están implícitas las realizaciones alternativas, incluidas las que pueden describirse utilizando las expresiones de transición “que consiste” o “que consiste esencialmente en”. Por lo tanto, por ejemplo, las realizaciones alternativas implícitas para un método que comprende una etapa de incorporación, una etapa de detección, una etapa de desprotección y una o más etapas de lavado incluyen realizaciones en las que el método consiste en etapas enumeradas y realizaciones en las que el método consiste esencialmente en las etapas enumeradas.

60 Como se utiliza en la presente memoria, “proporcionar” en el contexto de un compuesto, composición o artículo significa fabricar el compuesto, composición o artículo, adquirir el compuesto, composición o artículo u obtener de cualquier otra forma el compuesto, composición o artículo.

65 Como se utiliza en la presente memoria, la expresión “enzima extensora de cadena” es una enzima que produce una copia replicada de un polinucleótido usando el polinucleótido como cadena molde. Por ejemplo, la enzima extensora de cadena puede ser una enzima que tenga actividad polimerasa. De forma típica, las polimerasas de ADN se unen a la hebra molde y luego se mueven hacia abajo de la cadena molde añadiendo secuencialmente nucleótidos al grupo

hidroxilo libre del extremo 3' de una cadena en crecimiento de ácido nucleico. Las ADN polimerasas sintetizan de forma típica moléculas de ADN complementarias a partir de patrones de ADN y las ARN polimerasas sintetizan de forma típica moléculas de ARN a partir de patrones de ADN (transcripción). Las polimerasas pueden utilizar una cadena corta de ARN o ADN, llamada cebador, para iniciar el crecimiento de la cadena. Algunas polimerasas pueden desplazar la cadena corriente arriba del sitio donde están añadiendo bases a una cadena. Se dice que dichas polimerasas desplazan la cadena, lo que significa que tienen una actividad que elimina una cadena complementaria de una cadena molde que lee la polimerasa. Ejemplos de polimerasas que tienen actividad de desplazamiento de cadena incluyen, de modo no limitativo, el fragmento grande de polimerasa de Bst (*Bacillus stearothermophilus*), la polimerasa exo-Klenow o exo-polimerasa T7 de grado de secuenciación. Algunas polimerasas degradan la cadena delante de ellas, reemplazándola efectivamente por la cadena en crecimiento por detrás (actividad exonucleasa 5'). Algunas polimerasas tienen una actividad que degrada la cadena detrás de ellas (actividad exonucleasa 3'). Se han modificado algunas polimerasas útiles, ya sea por mutación o de otro modo, para reducir o eliminar la actividad exonucleasa 3' y/o 5'.

Como se utiliza en la presente memoria, el término "cebador" y sus derivados se refieren de forma general a cualquier polinucleótido que pueda hibridar con una secuencia diana de interés. De forma típica, el cebador funciona como un sustrato sobre el cual los nucleótidos una polimerasa polimeriza los nucleótidos; sin embargo, en algunas realizaciones, el cebador puede incorporarse a la cadena polinucleotídica sintetizada y proporcionar un sitio al que pueda hibridarse otro cebador para iniciar la síntesis de una nueva cadena que sea complementaria a la molécula de ácido nucleico sintetizada. El cebador puede estar compuesto por cualquier combinación de nucleótidos o análogos de los mismos. En algunos ejemplos, el cebador es un oligonucleótido o polinucleótido monocatenario.

Los términos "polinucleótido" y "oligonucleótido" se utilizan indistintamente en la presente memoria en referencia a una forma polimérica de nucleótidos de cualquier longitud, y pueden comprender ribonucleótidos, desoxirribonucleótidos, análogos de los mismos o mezclas de los mismos. Este término se refiere únicamente a la estructura primaria de la molécula. Por lo tanto, el término incluye ácido desoxirribonucleico ("ADN") tri-, bi- y monocatenario así como ácido ribonucleico ("ARN") tri-, bi- y mono-catenario. Como se utiliza en la presente memoria, "secuencias diana amplificadas" y sus derivados se refiere de forma general a una secuencia de polinucleótidos producida mediante la amplificación de las secuencias diana utilizando cebadores específicos de diana y los métodos proporcionados en la presente memoria. Las secuencias diana amplificadas pueden ser del mismo sentido (es decir, la cadena positiva) o antisentido (es decir, la cadena negativa) con respecto a las secuencias diana.

Los nucleótidos adecuados para su uso en los métodos proporcionados incluyen, aunque no de forma limitativa, trifosfatos de desoxinucleótidos, trifosfato de desoxiadenosina (dATP), trifosfato de desoxitimidina (dTTP), trifosfato de desoxicitidina (dCTP) y trifosfato de desoxiguanosina (dGTP). Opcionalmente, los nucleótidos utilizados en los métodos proporcionados, ya sean marcados o sin marcar, pueden incluir un resto bloqueador tal como un resto terminador reversible que inhibe la extensión de la cadena. Los marcadores adecuados para su uso en los nucleótidos marcados incluyen, aunque no de forma limitativa, haptenos, radionucleótidos, enzimas, marcadores fluorescentes, marcadores quimioluminiscentes y agentes cromógenos.

Un polinucleótido generalmente contendrá enlaces fosfodiéster, aunque en algunos casos, los análogos de ácido nucleico pueden tener cadenas principales alternativas, que comprenden, por ejemplo, fosforamida (Beaucage y col., *Tetrahedron*, 49(10): 1925 (1993) y las referencias en el mismo; Letsinger, *J. Org. Chem.* 35:3800 (1970); Sprinzl y col., *Eur. J. Biochem.* 81:579 (1977); Letsinger y col., *Nucl. Acids Res.* 14:3487 (1986); Sawai y col., *Chem. Lett.* 805 (1984), Letsinger y col., *J. Am. Chem. Soc.* 110:4470 (1988); y Pauwels y col., *Chemica Scripta*, 26:141 (1986)), fosforotioato (Mag y col., *Nucleic Acids Res.* 19:1437 (1991); y la Patente US 5.644.048), fosforoditioato (Briu y col., *J. Am. Chem. Soc.* 111:2321 (1989), enlaces O-metilfosforoamidita (véanse Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press) y cadenas principales de ácidos nucleicos peptídicos y enlaces (véase Egholm, *J. Am. Chem. Soc.* 114:1895 (1992); Meier y col., *Chem. Int. Ed. Otros ácidos nucleicos análogos incluyen aquellos con cadenas principales positivas (Denpcy y col., Proc. Natl. Acad. Sci. USA 92:6097 (1995); cadenas principales no iónicas (Patentes US 5.386.023, US 5.637.684, US 5.602.240, US 5.216.141 y US 4.469.863; Kiedrowski y col., *Angew. Chem. Intl. Ed. English* 30:423 (1991); Letsinger y col., *J. Am. Chem. Soc.* 110:4470 (1988); Letsinger y col., *Nucleoside & Nucleotide* 13:1597 (1994); Capítulos 2 y 3, *ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research"*, Ed. Y.S. Sanghui y P. Dan Cook; Mesmaeker y col., *Bioorganic & Medicinal Chem. Lett.* 4:395 (1994); Jeffs y col., *J. Biomolecular NMR*, 34:17 (1994); *Tetrahedron Lett.* 37:743 (1996)) y cadenas principales sin ribosa, incluyendo aquellas descritas en las Patentes US 5.235.033 y US 5.034.506, y los Capítulos 6 y 7, *ASC Symposium Series 580, "Carbohydrate Modifications in Antisense Research"*, Ed. Y.S. Sanghui y P. Dan Cook. Los polinucleótidos que contienen uno o más azúcares carbocíclicos también se incluyen dentro de la definición de polinucleótidos (véase Jenkins y col., *Chem. Soc. Rev.* (1995) págs.169-176). Se describen varios análogos de polinucleótidos en Rawls, *C & E News*, 2 de junio de 1997, página 35. Estas modificaciones de la cadena principal de ribosa-fosfato pueden realizarse para facilitar la adición de marcadores o para aumentar la estabilidad y la semivida de dichas moléculas en entornos fisiológicos.*

Un polinucleótido generalmente contendrá una secuencia específica de cuatro bases nucleotídicas: adenina (A); citosina (C); guanina (G); y timina (T). También puede haber presente uracilo (U), por ejemplo, como un reemplazo natural de timina cuando el ácido nucleico es ARN. En el ADN también puede utilizarse uracilo. Un polinucleótido

también puede incluir bases nativas o no nativas. En este sentido, un polinucleótido de ácido desoxirribonucleico nativo puede tener una o más bases seleccionadas del grupo que consiste en adenina, timina, citosina o guanina y un ácido ribonucleico puede tener una o más bases seleccionadas del grupo que consiste en uracilo, adenina, citosina o guanina. Se entenderá que los polinucleótidos de ácido desoxirribonucleico utilizados en los métodos o composiciones establecidos en la presente memoria pueden incluir, por ejemplo, bases de uracilo y un ácido ribonucleico puede incluir, por ejemplo, una base de timina. Las bases no nativas ilustrativas que pueden incluirse en un ácido nucleico, tanto si tiene una cadena principal natural como una estructura análoga, incluyen, sin limitarse a, inosina, xantina, hipoxantina, isocitocina, isoguanina, 2-aminopurina, 5-metilcitosina, 5-hidroximetil citosina, 2-aminoadenina, 6-metil adenina, 6-metil guanina, 2-propil guanina, 2-propil adenina, 2-tio-uracilo, 2-tiotimina, 2-tiocitosina, 15-halouracilo, 15-halocitosina, 5-propinil uracilo, 5-propinil citosina, 6-azo uracilo, 6-azo citosina, 6-azo timina, 5-uracilo, 4-tiouracilo, 8-halo adenina o guanina, 8-amino adenina o guanina, 8-tiol adenina o guanina, 8-tioalquil adenina o guanina, 8-hidroxil adenina o guanina, uracilo o citosina sustituido con 5-halo, 7-metilguanina, 7-metiladenina, 8-aza guanina, 8-azaadenina, 7-deazaguanina, 7-deazaadenina, 3-deazaguanina, 3-deazaadenina o similares. Opcionalmente, pueden incluirse isocitosina e isoguanina en un ácido nucleico para reducir la hibridación no específica, como se describe generalmente en la Patente US 5.681.702.

Una base no nativa utilizada en un polinucleótido puede tener actividad de emparejamiento de bases universal de manera que sea capaz de emparejarse con cualquier otra base de origen natural. Las bases ilustrativas que tienen actividad de emparejamiento de bases universal incluyen 3-nitropirrol y 5-nitroindol. Otras bases que pueden usarse incluyen aquellas que tienen actividad de emparejamiento de bases con un subconjunto de bases naturales tales como inosina, cuyas bases se emparejan con citosina, adenina o uracilo.

La incorporación de un nucleótido a una cadena polinucleotídica se refiere a la unión del nucleótido a un grupo hidroxilo 3' libre de la cadena polinucleotídica a través de la formación de un enlace fosfodiéster con el grupo fosfato 5' del nucleótido. El molde de polinucleótido que ha de secuenciarse puede ser ADN o ARN, o incluso una molécula híbrida que incluya tanto desoxinucleótidos como ribonucleótidos. El polinucleótido puede incluir nucleótidos de origen natural y/o no natural y uniones de cadena principal de origen natural o no natural.

La presente descripción describe, entre otras cosas, métodos de secuenciación de polinucleótidos que permiten tiempos de ciclo cortos para la incorporación de nucleótidos marcados reduciendo al mismo tiempo el ajuste de fase. Los métodos incluyen incubar nucleótidos sin marcar con un grupo de cadenas polinucleotídicas molde que tienen la misma secuencia cuando se detecta la identidad del nucleótido marcado añadido anteriormente. La etapa de detección proporciona tiempo para que la adición de los nucleótidos sin marcar se incorpore en las cadenas copia en las que no se incorporó el nucleótido marcado anteriormente de la etapa de incorporación anterior. Por lo tanto, al final de la etapa de detección, todas o la mayor parte de las cadenas copia estarán en fase y listas para incorporar el nucleótido marcado adecuado en la etapa de incorporación de la subsecuencia.

La etapa de detección tiende a llevar sustancialmente más tiempo que la etapa de incorporación, particularmente si la etapa de detección incluye la formación de imágenes. En consecuencia, la etapa de detección puede proporcionar sustancialmente más tiempo para la incorporación de nucleótidos que la etapa de incorporación, lo que puede reducir sustancialmente el ajuste de fase debido al retraso. Debido a que los nucleótidos incorporados en las cadenas copia retrasadas no están marcados, no producen una señal durante la detección y, por lo tanto, no interfieren con la detección.

Como se analiza en toda la presente memoria, se proporcionan métodos mejorados para secuenciar polinucleótidos. Se describen métodos de secuenciación ilustrativos, 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. Un método útil para una secuenciación rápida o de alto rendimiento es la secuenciación por síntesis (SBS, por sus siglas en inglés). Las técnicas de SBS incluyen, aunque no de forma limitativa, los sistemas Genome Analyzer (Illumina Inc., San Diego, CA) y los sistemas True Single Molecule Sequencing (tSMS)TM (Helicos BioSciences Corporation, Cambridge, MA). Brevemente, se utilizan varias reacciones de secuenciación por síntesis para dilucidar la identidad de una pluralidad de bases en posiciones diana dentro de una secuencia diana. Todas estas reacciones se basan en el uso de una secuencia de ácido nucleico diana que tiene al menos dos dominios; un primer dominio con el que se hibridará un cebador de secuenciación, y un segundo dominio adyacente, para el que se desea información de secuencia. Tras la formación de un complejo de ensayo, se utilizan enzimas de extensión para añadir trifosfatos de desoxinucleótidos (dNTP) a un cebador de secuenciación que se hibrida con el primer dominio, y cada adición de dNTP se lee para determinar la identidad del dNTP añadido. Esto puede transcurrir durante muchos ciclos. Las técnicas de SBS tales como los sistemas Genome Analyzer (Illumina Inc., San Diego, CA) y los sistemas True Single Molecule Sequencing (tSMS)TM

(Helicos BioSciences Corporation, Cambridge, MA), utilizan nucleótidos marcados para determinar la secuencia de una molécula de ácido nucleico diana. Una molécula de ácido nucleico diana puede hibridarse con un cebador e incubarse en presencia de una polimerasa y un nucleótido marcado que contenga un grupo bloqueador. El cebador se extiende de manera que se incorpore el nucleótido. La presencia del grupo bloqueador permite solamente una ronda de incorporación, es decir, la incorporación de un único nucleótido. La presencia del marcador permite la identificación del nucleótido incorporado. Puede añadirse una pluralidad de bases nucleotídicas individuales

homogéneas durante cada ciclo, tales como las que se usan en los sistemas True Single Molecule Sequencing (tSMS)[™] (Helicos BioSciences Corporation, Cambridge, MA) o, alternativamente, pueden añadirse las cuatro bases nucleotídicas durante cada ciclo simultáneamente, tal como se utiliza en los sistemas Genome Analyzer (Illumina Inc., San Diego, CA), particularmente cuando cada base se asocia a un marcador distinguible. Después de identificar el nucleótido incorporado mediante su correspondiente marcador, tanto el marcador como el grupo bloqueador pueden retirarse, permitiendo de este modo una ronda posterior de incorporación e identificación. La determinación de la identidad de la base nucleotídica añadida incluye, en algunas realizaciones, la exposición repetida de las bases marcadas recién añadidas a una fuente de luz que puede inducir una emisión detectable debido a la adición de una base nucleotídica específica, es decir, dATP, dCTP, dGTP o dTTP. Los métodos y composiciones descritos en la presente memoria son particularmente útiles para dichas técnicas de SBS. Además, los métodos y composiciones descritos en la presente memoria pueden ser particularmente útiles para la secuenciación de una matriz de ácidos nucleicos, donde pueden leerse múltiples secuencias simultáneamente desde múltiples posiciones en la matriz, puesto que cada nucleótido en cada posición puede identificarse basándose en su marcador identificable. Se describen métodos ilustrativos en los documentos US 2009/0088327; US 2010/0028885; y US 2009/0325172.

Refiriéndose ahora a las **FIG. 1-2**, se muestran visiones generales de realizaciones de métodos de secuenciación. Los métodos incluyen introducir una primera enzima extensora de cadena y una mezcla de nucleótidos marcados y bloqueados en un sitio que comprende múltiples cadenas polinucleotídicas molde que tienen la misma secuencia (**100**) de nucleótidos. La primera enzima extensora de cadena está configurada para incorporar uno adecuado de los nucleótidos marcados y bloqueados en cadenas polinucleotídicas copia, basándose en la secuencia de las cadenas molde a las que corresponden las cadenas polinucleotídicas copia. Durante o antes de la primera ronda de incorporación de nucleótidos, un cebador polinucleotídico puede hibridarse con la cadena molde para facilitar la secuenciación desde una posición de partida predeterminada.

Los nucleótidos marcados y bloqueados se incuban con las cadenas molde y la enzima extensora de cadena durante un período de tiempo en una “etapa de incorporación”. La etapa de incorporación puede llevar cualquier cantidad de tiempo adecuada. A medida que aumenta el tiempo de la etapa de incorporación, el ajuste de fase tiende a disminuir porque se proporciona tiempo suficiente para la incorporación de un nucleótido marcado y bloqueado en una cadena copia. Sin embargo, las etapas de incorporación más largas impiden el objetivo común de una secuenciación más rápida.

Aunque muchos métodos de secuenciación actuales proporcionan un tiempo de etapa de incorporación suficientemente largo para reducir el ajuste de fase a un nivel suficientemente bajo para proporcionar una tasa de error aceptable durante el número de ciclos en un ciclo, los presentes métodos pueden, en algunas realizaciones, acomodar un nivel más alto de ajuste de fase durante la etapa de incorporación porque la corrección del ajuste de fase puede producirse durante la etapa de detección.

Una ventaja de los métodos descritos en la presente memoria es que el tiempo de la etapa de incorporación puede reducirse a un nivel que dé como resultado un nivel de ajuste de fase anteriormente inaceptable porque las cadenas copia retrasadas pueden volver a ponerse en fase durante una etapa de detección. Al reducir el tiempo de la etapa de incorporación, puede reducirse el tiempo del proceso de secuenciación global.

En algunas realizaciones, el tiempo que una composición que comprende una mezcla de nucleótidos marcados y bloqueados se incuba con los polinucleótidos molde es de entre 1 segundo y 60 segundos, tal como de 5 segundos a 45 secciones, o de 10 segundos a 30 segundos. En algunas realizaciones, el tiempo que una composición que comprende una mezcla de nucleótidos marcados y bloqueados con los polinucleótidos molde es de 15 segundos o menos, tal como 10 segundos o menos, o 7,5 segundos o menos.

Preferiblemente, la etapa de incorporación es suficientemente larga para permitir la incorporación de nucleótidos marcados y bloqueados en un número suficiente de cadenas copia para producir una señal detectable durante una etapa de detección.

Todavía con referencia a las **FIG. 1-2**, después de incubar los nucleótidos marcados y bloqueados con las cadenas molde y la enzima extensora de cadena en la etapa de incorporación, los nucleótidos marcados y bloqueados sin incorporar pueden retirarse por lavado del sitio (**110**), y puede introducirse una composición que comprenda una mezcla de nucleótidos sin marcar y bloqueados en el sitio que comprende las múltiples cadenas (**120**) polinucleotídicas molde. En algunas realizaciones, la introducción de la composición que comprende la mezcla de nucleótidos sin marcar y bloqueados (etapa **120**) sirve para retirar por lavado los nucleótidos marcados y bloqueados sin incorporar (etapa **110**). La presencia de una etapa de lavado separada (**110**) puede ser preferible, en algunas realizaciones.

La identidad de los nucleótidos bloqueados y marcados incorporados en la cadena copia durante la etapa **100** después se detecta en presencia de la mezcla de nucleótidos (**130**) sin marcar y bloqueados en una “etapa de detección”. Los nucleótidos desbloqueados y marcados pueden incorporarse en cadenas copia que están retrasadas y desfasadas. Es decir, si un nucleótido marcado y bloqueado no se incorporó en la cadena copia durante la etapa **100**, puede incorporarse un nucleótido bloqueado y sin marcar durante la etapa de detección **130** para volver a poner en fase las cadenas copia retrasadas.

Los componentes de la composición que comprende los nucleótidos sin marcar y bloqueados no deberían interferir sustancialmente con la detección de la identidad del nucleótido marcado incorporado. Por ejemplo, si los nucleótidos marcados comprenden marcadores fluorescentes, los componentes de la composición preferiblemente no producen una señal fluorescente en condiciones que den como resultado una señal fluorescente a partir de los nucleótidos marcados, y los componentes de la composición preferiblemente no interfieren con la señal fluorescente producida por el nucleótido marcado.

El tiempo que la composición que comprende los nucleótidos marcados y bloqueados se incuba con las cadenas molde durante la etapa de detección es preferiblemente suficiente para reducir el ajuste de fase. Preferiblemente, la duración de la etapa de detección no aumenta para permitir tiempo adicional para la incorporación de los nucleótidos marcados y bloqueados. Es decir, la duración de las etapas de detección es preferiblemente una duración suficiente para detectar la identidad del nucleótido marcado y bloqueado incorporado en la etapa de incorporación.

Aunque los métodos de secuenciación actuales pueden proporcionar etapas de detección prolongadas, es probable que etapas de detección mucho más cortas aún permitan una corrección suficiente del ajuste de fase (es decir, la incorporación de nucleótidos sin marcar y bloqueados) incluso en condiciones menos que óptimas. Por ejemplo, frecuentemente se utilizan etapas de detección de 90 segundos con los secuenciadores Illumina MiniSeq™. Sin embargo, se cree que tiempos de detección mucho más cortos pueden ser suficientes para permitir la incorporación de nucleótidos en cadenas retrasadas durante la etapa de detección.

En algunas realizaciones, el tiempo que la composición que comprende los nucleótidos marcados y bloqueados se incuba con las cadenas molde durante una etapa de detección es de entre 5 segundos y 120 segundos, tal como entre 10 segundos y 90 segundos. En algunas realizaciones, el tiempo que la composición que comprende los nucleótidos marcados y bloqueados se incuba con las cadenas molde durante una etapa de detección es de 120 segundos o menos, tal como 60 segundos o menos, 30 segundos o menos, 10 segundos o menos, o 5 segundos o menos. En algunas realizaciones, el tiempo que la composición que comprende los nucleótidos marcados y bloqueados se incuba con las cadenas molde durante una etapa de detección es de 1 segundo o más, tal como 5 segundos o más, o 10 segundos o más.

Todavía con referencia a las FIG. 1-2, en la etapa 140, el método puede incluir además desbloquear y desmarcar cualquier nucleótido bloqueado y marcado incorporado en la cadena copia en la etapa 100 y desbloquear cualquier nucleótido bloqueado y sin marcar incorporado en la cadena copia en la etapa 130. La etapa de desbloqueo implica retirar un resto bloqueador del último nucleótido incorporado en la cadena copia y el resto detectable se escinde o enmascara en el último nucleótido incorporado en la cadena copia.

Puede realizarse una etapa de lavado (no mostrada) entre las etapas 130 y 140 o después de la etapa 140, o ambos.

Después del desbloqueo y el desmarcaje (140), el proceso puede repetirse hasta que se haya ejecutado un número predeterminado de ciclos o se haya completado la secuenciación. Alternativamente, puede introducirse una mezcla de nucleótidos marcados y bloqueados, sin una enzima que extienda la cadena, en el sitio que comprende las múltiples cadenas polinucleotídicas molde (150) y el proceso puede repetirse, comenzando en la etapa 110. La enzima extensora de cadena puede permanecer en el sitio. Por ejemplo, la enzima extensora de cadena puede unirse a la cadena molde, a la cadena copia o a ambas, y puede permanecer activa durante múltiples ciclos de secuenciación. Sin embargo, en algunas realizaciones se prefiere introducir una enzima extensora de cadena adicional junto con la mezcla de nucleótidos (100) marcados y bloqueados ya que la cantidad de enzima extensora de cadena funcional tiende a disminuir después de la etapa de desbloqueo y desmarcado (140).

Como se muestra en la FIG. 2 en la etapa 125, una segunda enzima extensora de cadena, que puede ser la misma enzima o una enzima diferente de la primera enzima extensora de cadena introducida en la etapa 100, puede incluirse con la composición que comprende la mezcla de nucleótidos sin marcar y bloqueados e introducirse en el sitio que comprende las múltiples cadenas molde. La etapa 125 en la FIG. 2 puede reemplazar la etapa 120 en uno o más ciclos del proceso de la FIG. 1. Sorprendentemente, se ha descubierto que puede producirse menos ajuste de fase si no se añade una segunda enzima extensora de cadena (por ejemplo, como se indica en la etapa 120 en la FIG. 1) que si se añade una segunda enzima extensora de cadena en la etapa 125.

En los métodos proporcionados, las etapas de incorporación y detección de nucleótidos pueden repetirse una o más veces. Por ejemplo, las etapas pueden repetirse al menos 25 veces en algunas realizaciones, al menos 75 veces en otras realizaciones y al menos 100 veces en otras realizaciones más. Por lo tanto, los métodos proporcionados incluyen, aunque no de forma limitativa, repetir las etapas de incorporación y detección durante un número de ciclos en un intervalo de entre aproximadamente 100 ciclos y aproximadamente 1000 ciclos, en algunas realizaciones, de entre aproximadamente 100 ciclos y aproximadamente 500 ciclos, en otras realizaciones, y entre aproximadamente 100 ciclos y aproximadamente 300 ciclos, en otras realizaciones más.

Los métodos de secuenciación descritos en la presente memoria pueden realizarse de cualquier manera adecuada, utilizando cualquier equipo adecuado. En algunas realizaciones, los métodos de secuenciación emplean un soporte

sólido sobre el cual se inmovilizan múltiples cadenas polinucleotídicas molde. El término inmovilizado tal como se utiliza en la presente memoria pretende abarcar la unión directa o indirecta a un soporte sólido a través de uno o más enlaces covalentes o no covalentes. En realizaciones particulares, todo lo que se requiere es que los polinucleótidos permanezcan inmovilizados o unidos a un soporte en las condiciones en las que se pretende usar el soporte, por ejemplo, en aplicaciones que requieran amplificación y/o secuenciación de ácidos nucleicos. Por ejemplo, pueden inmovilizarse oligonucleótidos o cebadores de manera que un extremo 3' esté disponible para la extensión enzimática y/o al menos una porción de la secuencia sea capaz de hibridarse con una secuencia complementaria. La inmovilización puede producirse a través de hibridación con un cebador unido a la superficie, en cuyo caso el cebador u oligonucleótido inmovilizado puede estar en la orientación 3'-5'. Alternativamente, la inmovilización puede producirse mediante hibridación sin emparejamiento de bases, tal como la unión covalente.

A manera de ejemplo, los polinucleótidos pueden unirse a la superficie mediante hibridación o recocido con uno o más cebadores en un parche de cebadores. La hibridación puede lograrse, por ejemplo, ligando un adaptador a los extremos de los polinucleótidos molde. La secuencia de ácido nucleico del adaptador puede ser complementaria a la secuencia de ácido nucleico del cebador, permitiendo por lo tanto que el adaptador se una o hibride con el cebador en la superficie. Opcionalmente, los polinucleótidos pueden ser monocatenarios o bicatenarios y pueden añadirse adaptadores a los extremos 5' y/o 3' de los polinucleótidos. Opcionalmente, los polinucleótidos pueden ser bicatenarios y pueden ligarse adaptadores en los extremos 3' del polinucleótido bicatenario. Opcionalmente, pueden usarse polinucleótidos sin ningún adaptador. En algunas realizaciones, los polinucleótidos molde pueden unirse a una superficie mediante interacciones distintas de la hibridación con un cebador complementario. Por ejemplo, un polinucleótido puede unirse covalentemente a una superficie usando un enlace químico tal como los que son resultado de la química de clic o una interacción receptor-ligando tal como la unión estreptavidina-biotina.

Oligonucleótidos cebadores, cebadores oligonucleotídicos y cebadores se usan indistintamente y son secuencias de polinucleótidos que son capaces de hibridarse específicamente con uno o más moldes de polinucleótidos que han de amplificarse o secuenciarse. Generalmente, los oligonucleótidos cebadores son monocatenarios o parcialmente monocatenarios. Los cebadores también pueden contener una mezcla de bases no naturales, modificaciones químicas no nucleotídicas o enlaces de cadena principal no naturales siempre que las entidades no naturales no interfieran con la función del cebador. Opcionalmente, un parche de cebadores en una superficie de un soporte sólido puede comprender una o más pluralidades diferentes de moléculas de cebador. A manera de ejemplo, un parche puede comprender una primera, segunda, tercera, cuarta o más pluralidades de moléculas cebadoras, teniendo cada pluralidad una secuencia diferente. Se entenderá que para realizaciones que tienen diferentes pluralidades de cebadores en un único parche, las diferentes pluralidades de cebadores pueden compartir una secuencia común siempre que exista una diferencia de secuencia entre al menos una porción de las diferentes pluralidades. Por ejemplo, una primera pluralidad de cebadores puede compartir una secuencia con una segunda pluralidad de cebadores siempre que los cebadores de una pluralidad tengan una secuencia diferente que no se encuentre en los cebadores de la otra pluralidad.

Los polinucleótidos molde pueden amplificarse en la superficie del soporte sólido. La amplificación de polinucleótidos incluye el proceso de amplificación o aumento del número de un molde de polinucleótido y/o de un complemento del mismo que están presentes mediante la producción de una o más copias del molde y/o su complemento. La amplificación puede realizarse mediante una diversidad de métodos conocidos en condiciones que incluyen, aunque no de forma limitativa, amplificación por termociclado o amplificación isotérmica. Por ejemplo, se describen métodos para realizar la amplificación en la Publicación US 2009/0226975; el documento WO 98/44151; el documento WO 00/18957; el documento WO 02/46456; el documento WO 06/064199; y el documento WO 07/010251. Brevemente, en los métodos proporcionados, la amplificación puede producirse en la superficie a la que están unidas las moléculas de polinucleótidos. Este tipo de amplificación puede denominarse amplificación en fase sólida, que cuando se utiliza en referencia a polinucleótidos, se refiere a cualquier reacción de amplificación de polinucleótidos realizada sobre o en asociación con una superficie (por ejemplo, un soporte sólido). De forma típica, la totalidad o una porción de los productos amplificados se sintetizan mediante la extensión de un cebador inmovilizado. Las reacciones de amplificación en fase sólida son análogas a las amplificaciones en fase de solución convencionales excepto por que al menos uno de los cebadores de amplificación se inmoviliza en una superficie (por ejemplo, un soporte sólido).

Las condiciones adecuadas incluyen proporcionar tampones/soluciones adecuadas para amplificar polinucleótidos. Dichas soluciones incluyen, por ejemplo, una enzima con actividad polimerasa, trifosfatos de nucleótidos y, opcionalmente, aditivos tales como DMSO o betaína. Opcionalmente, la amplificación se realiza en presencia de un agente recombinasa como se describe en la Patente US 7.485.428, que permite la amplificación sin fusión térmica. Brevemente, los agentes recombinasa tales como la proteína RecA de *E. coli* (o un pariente de RecA de otros filos), en presencia de, por ejemplo, ATP, dATP, ddATP, UTP o ATP γ S, formará un filamento nucleoproteico alrededor del ADN monocatenario (por ejemplo, un cebador). Cuando este complejo entra en contacto con secuencias homólogas, el agente recombinasa catalizará una reacción de invasión de cadena y el emparejamiento del cebador con la cadena homóloga del ADN diana. La cadena de emparejamiento original es desplazada por la invasión de cadena, dejando una burbuja de ADN monocatenario en la región, que sirve como molde para la amplificación.

La amplificación en fase sólida puede comprender una reacción de amplificación de polinucleótidos que comprende solo una especie de cebador oligonucleotídico inmovilizado en una superficie. Alternativamente, la superficie puede

comprender una pluralidad de primera y segunda especies de cebadores de oligonucleótidos inmovilizados diferentes. Las reacciones de amplificación de ácidos nucleicos en fase sólida generalmente comprenden al menos uno de dos tipos diferentes de amplificación de ácidos nucleicos, la amplificación interfacial y la de superficie (o puente). Por ejemplo, en la amplificación interfacial, el soporte sólido comprende un polinucleótido molde que se inmoviliza indirectamente en el soporte sólido mediante hibridación con un cebador oligonucleotídico inmovilizado, el cebador inmovilizado puede extenderse en el curso de una reacción de elongación dirigida por molde catalizada por polimerasa (por ejemplo, extensión del cebador) para generar un polinucleótido inmovilizado que permanece unido al soporte sólido. Después de la fase de extensión, los polinucleótidos (por ejemplo, el molde y su producto complementario) se desnaturalizan de manera que el polinucleótido molde se libere en solución y esté disponible para la hibridación con otro cebador oligonucleotídico inmovilizado. El polinucleótido molde puede estar disponible en 1, 2, 3, 4, 5 o más rondas de extensión del cebador o puede retirarse por lavado de la reacción después de 1, 2, 3, 4, 5 o más rondas de extensión del cebador.

En la amplificación de superficie (o puente), un polinucleótido inmovilizado se hibrida con un cebador oligonucleotídico inmovilizado. El extremo 3' del polinucleótido inmovilizado proporciona el molde para una reacción de elongación dirigida por molde catalizada por polimerasa (por ejemplo, extensión del cebador) que se extiende desde el cebador oligonucleotídico inmovilizado. El producto bicatenario resultante "une por puente" los dos cebadores y ambas cadenas se unen covalentemente al soporte. En el siguiente ciclo, después de la desnaturalización que produce un par de cadenas individuales (el molde inmovilizado y el producto del cebador extendido) inmovilizadas en el soporte sólido, ambas cadenas inmovilizadas pueden servir como moldes para una nueva extensión del cebador.

Puede usarse amplificación para producir colonias de polinucleótidos inmovilizados. Por ejemplo, los métodos pueden producir matrices agrupadas de colonias de polinucleótidos, análogas a las descritas en la Patente US 7.115.400; y la Publicación US 2005/0100900; el documento WO 00/18957; y el documento WO 98/44151. "Grupos" y "colonias" se usan indistintamente y se refieren a una pluralidad de copias de un polinucleótido que tiene la misma secuencia y/o complementos de la misma unidos a una superficie. De forma típica, el grupo comprende una pluralidad de copias de un polinucleótido que tiene la misma secuencia y/o complementos de la misma, unidas a través de sus extremos 5' a la superficie. Las copias de polinucleótidos que componen los grupos pueden estar en forma monocatenaria o bicatenaria.

Por lo tanto, la pluralidad de polinucleótidos molde puede estar en un grupo, conteniendo cada grupo polinucleótidos molde de la misma secuencia. Puede secuenciarse una pluralidad de grupos, comprendiendo cada grupo polinucleótidos de la misma secuencia. Opcionalmente, la secuencia de los polinucleótidos en un primer grupo es diferente de la secuencia de las moléculas de ácido nucleico de un segundo grupo. Opcionalmente, el grupo se forma hibridando con un cebador sobre una superficie sólida un polinucleótido molde y amplificando el polinucleótido molde en condiciones para formar el grupo que comprende la pluralidad de polinucleótidos molde de la misma secuencia. La amplificación puede ser térmica o isotérmica.

Cada colonia puede comprender polinucleótidos de las mismas secuencias. En realizaciones particulares, la secuencia de los polinucleótidos de una colonia es diferente de la secuencia de los polinucleótidos de otra colonia. Por lo tanto, cada colonia comprende polinucleótidos que tienen diferentes secuencias de ácidos nucleicos. Todos los polinucleótidos inmovilizados en una colonia se producen de forma típica mediante amplificación del mismo polinucleótido. En algunas realizaciones, es posible que una colonia de polinucleótidos inmovilizados contenga uno o más cebadores sin un polinucleótido inmovilizado al que pueda unirse otro polinucleótido de secuencia diferente tras la aplicación adicional de soluciones que contienen polinucleótidos libres o sin unir. Sin embargo, debido a la falta de un número suficiente de cebadores libres en una colonia, este polinucleótido segundo o polinucleótido invasor puede no amplificarse a números significativos. El segundo polinucleótido o polinucleótido invasor de forma típica es menos del 1, 0,5, 0,25, 0,1, 0,001 o 0,0001 % de la población total de polinucleótidos en una única colonia. Por lo tanto, el segundo polinucleótido o polinucleótido invasor puede no detectarse ópticamente o la detección del segundo polinucleótido o polinucleótido invasor se considera ruido de fondo o no interfiere con la detección de los polinucleótidos originales inmovilizados en la colonia. En dichas realizaciones, la colonia será aparentemente homogénea o uniforme según la resolución de los métodos o aparatos utilizados para detectar la colonia.

Los grupos pueden tener diferentes formas, tamaños y densidades dependiendo de las condiciones utilizadas. Por ejemplo, los grupos pueden tener una forma que sea sustancialmente redonda, de múltiples lados, en forma de rosquilla o en forma de anillo. El diámetro o sección transversal máxima de un grupo puede ser de aproximadamente 0,2 μm a aproximadamente 6 μm , de aproximadamente 0,3 μm a aproximadamente 4 μm , de aproximadamente 0,4 μm a aproximadamente 3 μm , de aproximadamente 0,5 μm a aproximadamente 2 μm , de aproximadamente 0,75 μm a aproximadamente 1,5 μm , o cualquier diámetro intermedio. Opcionalmente, el diámetro o sección transversal máxima de un grupo puede ser de al menos aproximadamente 0,5 μm , al menos aproximadamente 1 μm , al menos aproximadamente 1,5 μm , al menos aproximadamente 2 μm , al menos aproximadamente 2,5 μm , al menos aproximadamente 3 μm , en al menos aproximadamente 4 μm , al menos aproximadamente 5 μm o al menos aproximadamente 6 μm . El diámetro de un grupo puede verse influenciado por una serie de parámetros que incluyen, aunque no de forma limitativa, el número de ciclos de amplificación realizados para producir el grupo, la longitud del molde de polinucleótido, el contenido de GC del molde de polinucleótido, la forma de un parche al que se unen los cebadores, o la densidad de los cebadores unidos a la superficie sobre la que se forman los grupos. Sin embargo,

como se ha analizado anteriormente, en todos los casos, el diámetro de un grupo no puede ser mayor que el parche sobre el que se forma el grupo. Por ejemplo, si un parche es una perla, el tamaño del grupo no será superior al área de superficie de la perla. La densidad de los grupos puede estar en el intervalo de al menos aproximadamente $0,1/\text{mm}^2$, al menos aproximadamente $1/\text{mm}^2$, al menos aproximadamente $10/\text{mm}^2$, al menos aproximadamente $100/\text{mm}^2$, al menos aproximadamente $1.000/\text{mm}^2$, al menos aproximadamente $10.000/\text{mm}^2$ a al menos aproximadamente $100.000/\text{mm}^2$. Opcionalmente, los grupos tienen una densidad de, por ejemplo, $100.000/\text{mm}^2$ a $1.000.000/\text{mm}^2$ o $1.000.000/\text{mm}^2$ a $10.000.000/\text{mm}^2$. Los métodos proporcionados en el presente documento pueden producir colonias de aproximadamente el mismo tamaño. Esto se produce independientemente de las diferencias en las eficiencias de amplificación de los polinucleótidos de diferente secuencia.

Los grupos pueden detectarse, por ejemplo, usando un medio de formación de imágenes adecuado, tal como un dispositivo de formación de imágenes confocal o un dispositivo de carga acoplada (CCD, por sus siglas en inglés) o una cámara CMOS. Los dispositivos de formación de imágenes ilustrativos incluyen, aunque no de forma limitativa, los descritos en las Patentes US 7.329.860; US 5.754.291; y US 5.981.956; y el documento WO 2007/123744. El aparato de formación de imágenes puede usarse para determinar una posición de referencia en un grupo o en una pluralidad de grupos en la superficie, tal como la ubicación, límite, diámetro, área, forma, superposición y/o centro de uno o una pluralidad de grupos (y/o de una señal detectable que se origina a partir de los mismos). Una posición de referencia de este tipo puede registrarse, documentarse, anotarse, convertirse en una señal interpretable o similar, para proporcionar información significativa.

Como se utiliza en la presente memoria, el término “soporte” se refiere a un sustrato para unir polinucleótidos. Un soporte es un material que tiene una superficie rígida o semirrígida a la que puede unirse un polinucleótido o sobre la que pueden sintetizarse y/o modificarse ácidos nucleicos. Los soportes pueden incluir cualquier resina, gel, perla, pocillo, columna, chip, célula de flujo, membrana, matriz, placa, filtro, vidrio, vidrio de poro controlado (CPG, por sus siglas en inglés), soporte de polímero, membrana, papel, plástico, tubo o tableta de plástico, perla de plástico, perla de vidrio, portaobjetos, cerámica, chip de silicio, placa multipocillos, membrana de nailon, fibra óptica y membrana de PVDF.

Un soporte puede incluir cualquier sustrato plano similar a una oblea y sustratos planos que tengan pocillos, tales como una placa de microtitulación, incluyendo placas de 96 pocillos. Los sustratos planos ilustrativos incluyen chips, portaobjetos, sustratos grabados, placas de microtitulación y reactores de célula de flujo, incluyendo reactores de célula de flujo de múltiples calles que tienen múltiples canales microfluídicos, tales como la célula de flujo de ocho canales utilizada en la estación de trabajo de secuenciación cBot (Illumina, Inc., San Diego, CA). En el documento WO 2007/123744 se describen células de flujo ilustrativas. Opcionalmente, la célula de flujo es una célula de flujo con patrón. Las células de flujo con patrón adecuadas incluyen, aunque no de forma limitativa, células de flujo descritas en el documento WO 2008/157640.

Un soporte también puede incluir perlas, incluyendo perlas magnéticas, perlas huecas y perlas sólidas. Pueden usarse perlas junto con soportes planos, conteniendo también dichos soportes planos opcionalmente pocillos. Las perlas, o alternativamente microesferas, se refieren generalmente a un pequeño cuerpo hecho de un material rígido o semirrígido. El cuerpo puede tener una forma caracterizada, por ejemplo, como una esfera, un óvalo, una microesfera u otra forma de partícula reconocida que tenga dimensiones regulares o irregulares. Los tamaños de las perlas, en particular, incluyen, sin limitarse a, aproximadamente $1\ \mu\text{m}$, aproximadamente $2\ \mu\text{m}$, aproximadamente $3\ \mu\text{m}$, aproximadamente $5\ \mu\text{m}$, aproximadamente $10\ \mu\text{m}$, aproximadamente $20\ \mu\text{m}$, aproximadamente $30\ \mu\text{m}$, aproximadamente $40\ \mu\text{m}$, aproximadamente $60\ \mu\text{m}$, aproximadamente $100\ \mu\text{m}$, aproximadamente $150\ \mu\text{m}$ o aproximadamente $200\ \mu\text{m}$ de diámetro. Pueden usarse otras partículas de manera similar a las descritas en la presente memoria para perlas y microesferas.

La composición de un soporte puede variar dependiendo, por ejemplo, del formato, la química y/o el método de unión y/o del método de síntesis de ácido nucleico. Los materiales de soporte que pueden usarse según la presente descripción incluyen, aunque no de forma limitativa, polipropileno, polietileno, polibutileno, poliuretanos, nailon, metales y otros materiales adecuados. Las composiciones ilustrativas incluyen soportes y funcionalidades químicas transmitidas a las mismas, utilizadas en la síntesis de polipéptidos, polinucleótidos y/o restos orgánicos. Dichas composiciones incluyen, por ejemplo, plásticos, cerámica, vidrio, poliestireno, melamina, metilestireno, polímeros acrílicos, materiales paramagnéticos, sol de óxido de torio, grafito de carbono, dióxido de titanio, látex o dextranos reticulados tales como Sepharose™, celulosa, nailon, micelas reticuladas y Teflon™, así como cualquier otro material que pueda encontrarse descrito, por ejemplo, en “*Microsphere Detection Guide*” de Bangs Laboratories, Fishers IN. Una partícula de soporte puede estar hecha de almidón reticulado, dextranos, celulosa, proteínas, polímeros orgánicos que incluyen polímeros de estireno que incluyen poliestireno y metilestireno, así como otros copolímeros de estireno, plásticos, vidrio, cerámica, polímeros acrílicos, materiales de respuesta al magnetismo, coloides, sol de óxido de torio, grafito de carbono, dióxido de titanio, nailon, látex o TEFLON®. La “*Microsphere Detection Guide*” de Bangs Laboratories, Fishers Inc. es una guía útil. Otros soportes ilustrativos dentro del alcance de la presente descripción incluyen, por ejemplo, los descritos en la Publicación de solicitud US 02/0102578 y la Patente US 6.429.027.

Por ejemplo, y en referencia a la FIG. 3, se muestra una realización de un soporte sólido **200**, tal como una célula de flujo. El soporte sólido **200** tiene una superficie **210** a la que se unen grupos **300** que contienen múltiples cadenas

polinucleotídicas molde que tienen la misma secuencia de nucleótidos. La superficie **210** del soporte sólido **200** puede ser plana.

5 Las composiciones fluidas que contienen reactivos, tampones de lavado y similares pueden fluir sobre la superficie **210** del soporte sólido **200** para interactuar con los polinucleótidos molde en los grupos **300**. El flujo de las composiciones puede producirse en cualquier dirección, tal como la dirección indicada por las flechas en la **FIG. 3**.

10 El aparato de secuenciación con el que puede utilizarse la célula de flujo **300** puede configurarse para hacer fluir reactivos y composiciones a través de la superficie **210** para interactuar con las cadenas molde en los grupos **300**. Por ejemplo, el aparato puede provocar que enzimas extensoras de cadena, cebadores de secuenciación, nucleótidos, composiciones de lavado, reactivos de desbloqueo, reactivos de desmarcaje y similares fluyan a través de la superficie **210** del soporte sólido **200**, tal como una célula de flujo, para interactuar con los polinucleótidos molde en los grupos **300** en los momentos adecuados para realizar la secuenciación de las cadenas molde.

15 Cada grupo **300** puede contener los mismos polinucleótidos molde o polinucleótidos diferentes que otro grupo **300**.

20 Los polinucleótidos molde que han de secuenciarse pueden obtenerse de cualquier muestra biológica usando métodos rutinarios conocidos. Las muestras biológicas adecuadas incluyen, aunque no de forma limitativa, una muestra de sangre, una muestra de ensayo de biopsia, un explante de tejido, un cultivo de órgano, un fluido biológico o cualquier otra preparación de tejido o célula, o fracción o derivado de los mismos o aislado de los mismos. La muestra biológica puede ser un cultivo celular primario o una estirpe celular adaptada al cultivo incluyendo, aunque no de forma limitativa, estirpes celulares modificadas genéticamente que pueden contener secuencias de ácido nucleico recombinantes episómicas o integradas cromosómicamente, estirpes celulares inmortalizadas o inmortalizables, estirpes celulares híbridas de células somáticas, diferenciadas o diferenciables, estirpes celulares transformadas, células madre, células germinales (por ejemplo, espermatozoides, ovocitos), estirpes celulares transformadas y similares. Por ejemplo, las moléculas de polinucleótidos pueden obtenerse a partir de células primarias, estirpes celulares, células o tejidos recién aislados, células o tejidos congelados, células o tejidos incluidos en parafina, células o tejidos fijados y/o células o tejidos disecados con láser. Las muestras biológicas pueden obtenerse de cualquier sujeto o fuente biológica incluyendo, por ejemplo, animales humanos o no humanos, incluyendo mamíferos y no mamíferos, vertebrados e invertebrados, y también puede ser cualquier organismo multicelular u organismo unicelular tal como un organismo eucariota (incluyendo plantas y algas) o procariota, arquea, microorganismos (por ejemplo, bacterias, arqueas, hongos, protistas, virus) y plancton acuático.

35 Una vez que se obtienen los polinucleótidos, puede prepararse una pluralidad de moléculas de polinucleótidos de diferente secuencia para su uso en los métodos proporcionados usando una diversidad de técnicas convencionales disponibles y conocidas. Los métodos ilustrativos de preparación de moléculas de polinucleótidos incluyen, aunque no de forma limitativa, aquellos descritos en Bentley y col., Nature 456:49-51 (2008); la Patente US 7.115.400; las Publicaciones de solicitud de patente US 2007/0128624; US 2009/0226975; US 2005/0100900; US 2005/0059048; US 2007/0110638; y US 2007/0128624. Los polinucleótidos molde pueden contener una diversidad de secuencias incluyendo, aunque no de forma limitativa, secuencias universales y secuencias conocidas o desconocidas. Por ejemplo, el polinucleótido puede comprender una o más regiones de secuencia conocida (por ejemplo, un adaptador) ubicadas en los extremos 5' y/o 3'. Dichos polinucleótidos molde pueden formarse uniendo adaptadores a los extremos de un polinucleótido de secuencia desconocida. Cuando los polinucleótidos comprenden secuencias conocidas en los extremos 5' y 3', las secuencias conocidas pueden ser secuencias iguales o diferentes. Opcionalmente, una secuencia conocida ubicada en los extremos 5' y/o 3' de los polinucleótidos es capaz de hibridarse con uno o más cebadores inmovilizados en la superficie. Por ejemplo, un polinucleótido que comprende una secuencia conocida en 5' puede hibridarse con una primera pluralidad de cebadores mientras que la secuencia conocida en 3' puede hibridarse con una segunda pluralidad de cebadores. Opcionalmente, los polinucleótidos comprenden uno o más marcadores detectables. El uno o más marcadores detectables pueden unirse al molde de polinucleótido en el extremo 5', en el extremo 3' y/o en cualquier posición de nucleótido dentro de la molécula de polinucleótido. Los polinucleótidos para su uso en los métodos proporcionados pueden comprender el polinucleótido que se ha de amplificarse y/o secuenciarse y, opcionalmente, secuencias cortas de ácido nucleico en los extremos 5' y/o 3'.

55 Una secuencia corta de ácido nucleico que se añade al extremo 5' y/o 3' de un polinucleótido puede ser una secuencia universal. Una secuencia universal es una región de secuencia de nucleótidos que es común, es decir, compartida por, dos o más polinucleótidos, donde los dos o más polinucleótidos también tienen regiones de diferencias de secuencia. Una secuencia universal que puede estar presente en diferentes miembros de una pluralidad de polinucleótidos puede permitir la replicación o la amplificación de múltiples secuencias diferentes utilizando un solo cebador universal que es complementario a la secuencia universal. Similarmente, al menos una, dos (por ejemplo, un par) o más secuencias universales que pueden estar presentes en diferentes miembros de una colección de polinucleótidos pueden permitir la replicación o amplificación de múltiples secuencias diferentes usando al menos uno, dos (por ejemplo, un par) o más cebadores universales únicos que son complementarios a las secuencias universales. Por lo tanto, un cebador universal incluye una secuencia que puede hibridarse específicamente con una secuencia universal de este tipo. El polinucleótido puede modificarse para unirse a adaptadores universales (por ejemplo, secuencias de ácido nucleico no diana) a uno o ambos extremos de las diferentes secuencias diana, proporcionando los adaptadores sitios para la hibridación de cebadores universales. Este enfoque tiene la ventaja de que no es

necesario diseñar un par específico de cebadores para cada polinucleótido que se ha de generar, amplificar, secuenciar y/o analizar de cualquier otra manera; puede utilizarse un único par de cebadores para la amplificación de polinucleótidos diferentes siempre que cada polinucleótido se modifique por la adición de las mismas secuencias de unión a cebador universal en sus extremos 5' y 3'.

Los polinucleótidos también pueden modificarse para incluir cualquier secuencia de ácido nucleico deseable usando métodos conocidos y convencionales. Dichas secuencias adicionales pueden incluir, por ejemplo, sitios de enzimas de restricción o marcadores de indexación para permitir la identificación de productos de amplificación de una secuencia de ácido nucleico dada.

Como se utiliza en la presente memoria, el término diferente cuando se usa en referencia a dos o más polinucleótidos significa que los dos o más polinucleótidos tienen secuencias de nucleótidos que no son iguales. Por ejemplo, dos polinucleótidos pueden diferir en el contenido y el orden de los nucleótidos en la secuencia de un polinucleótido en comparación con el otro polinucleótido. El término puede usarse para describir polinucleótidos, ya sea que se los denomine copias, amplicones, moldes, dianas, cebadores, oligonucleótidos o similares.

Haciendo referencia ahora a las FIG. 4A-G, se muestran dibujos esquemáticos que ilustran una realización de un método de secuenciación. La realización representada ilustra un problema con el ajuste de fase y cómo el método aborda el problema. En las FIG. 4A-G, se muestra una pluralidad de polinucleótidos **400** que tienen la misma secuencia. Los polinucleótidos **400** pueden estar unidos a una superficie de un soporte sólido en sus extremos 3' o 5'. Por ejemplo, los polinucleótidos pueden estar unidos a la superficie **210** en un grupo **300** en una célula **200** de flujo como se muestra en la FIG. 3.

La FIG. 4A muestra los polinucleótidos molde **400** antes del inicio de la secuenciación. La FIG. 4B muestra un único nucleótido **510** marcado y bloqueado incorporado en una cadena copia **590** durante una primera ronda de secuenciación, tal como después de la etapa de lavado **110** de las FIG. 1-2. Por comodidad de uso y claridad, la enzima extensora de cadena, que puede permanecer en los polinucleótidos molde **400** después de una etapa de lavado, no se muestra. En algunas realizaciones, puede usarse un polinucleótido cebador (no mostrado), complementario a una porción de la secuencia del molde, para facilitar la incorporación de los nucleótidos **510**. En la realización representada en la FIG. 4B, un nucleótido **510** marcado y bloqueado se incorpora en las cadenas copia **590** para cuatro de los cinco polinucleótidos molde **400**. Sin más, la cadena en la que no se incorporó ningún nucleótido **510** marcado y bloqueado estaría desfasada.

Sin embargo, como se muestra en la FIG. 4C, si se incubasen nucleótidos sin marcar y bloqueados (por ejemplo, etapa **120** o **125** en la FIG. 1 o FIG. 2) con los polinucleótidos molde **400** durante la detección de la identidad de los nucleótidos **510** marcados y bloqueados (por ejemplo, etapa **130** de las FIG. 1-2), un nucleótido **515** sin marcar y bloqueado podría incorporarse en la cadena retrasada para poner la cadena retrasada en fase antes del siguiente ciclo de secuenciación. Como se ha indicado anteriormente, el nucleótido **515** sin marcar y bloqueado puede incorporarse mediante una enzima extensora de cadena (no mostrada) que permanece después de una etapa de lavado (por ejemplo, etapa **110** de las FIG. 1-2) o que se introduce con la mezcla de nucleótidos sin marcar y desbloqueados (por ejemplo, etapa **125** de la FIG. 2). Los nucleótidos incorporados **510**, **515** después pueden desbloquearse y desmarcarse, en el caso de los nucleótidos **510** (por ejemplo, etapa **140** de las FIG. 1-2), y puede realizarse otro ciclo de secuenciación.

La FIG. 4D muestra la incorporación de segundos nucleótidos **520** marcados y bloqueados en la cadena copia **590** después de una etapa de incorporación (tal como la etapa **100** de las FIG. 1-2) y una etapa de lavado (tal como la etapa **110** de las FIG. 1-2) en un segundo ciclo de secuenciación. En la realización representada en la FIG. 4D, un nucleótido **520** marcado y bloqueado se incorpora en las cadenas copia **590** para cuatro de los cinco polinucleótidos molde **400**. Sin más, la cadena copia **590** en la que no se incorporó ningún nucleótido **510** marcado y bloqueado estaría desfasada.

Sin embargo, como se muestra en la FIG. 4E, si se incubasen nucleótidos sin marcar y bloqueados (por ejemplo, etapa **120** o **125** en la FIG. 1 o FIG. 2) con los polinucleótidos molde **400** durante la detección de la identidad de los nucleótidos **520** marcados y bloqueados (por ejemplo, etapa **130** de las FIG. 1-2), un nucleótido **525** sin marcar y bloqueado podría incorporarse en la cadena retrasada para poner la cadena retrasada en fase antes del siguiente ciclo de secuenciación. Los nucleótidos incorporados **520**, **525** después pueden desbloquearse y desmarcarse, en el caso de los nucleótidos **520** (por ejemplo, etapa **140** de las FIG. 1-2) y puede realizarse otro ciclo de secuenciación.

La FIG. 4F ilustra otro ciclo de secuenciación en el que se incorporan nucleótidos **530** marcados y bloqueados en cuatro de cinco cadenas copia **590** después de una etapa de incorporación. La FIG. 4G ilustra la incorporación de un nucleótido **535** sin marcar y bloqueado en la cadena retrasada para poner la cadena retrasada en fase antes del siguiente ciclo de secuenciación. Después de tres ciclos de secuenciación como se ilustra en la FIG. 4G, todas las cadenas copia **590** están en fase. Sin embargo, sin incubar los nucleótidos sin marcar y bloqueados durante la etapa de detección, tres de las cinco cadenas habrían estado desfasadas en la realización ilustrada. Con el 60 % de las cadenas copia desfasadas, la señal producida durante la etapa de detección probablemente sería demasiado

heterogénea, debido a la incorporación de un nucleótido marcado diferente, para identificar adecuadamente el nucleótido adecuado.

Aunque las figuras presentadas en las FIG. 4A-G son dibujos esquemáticos, ilustran algunos puntos clave. Por ejemplo, una tasa pequeña de ajuste de fase puede combinarse a lo largo de múltiples ciclos de secuenciación. Sin embargo, poner en fase las cadenas copia retrasadas durante la detección a través de la incorporación de nucleótidos sin marcar no requiere una gran actividad de extensión de cadena debido al número relativamente bajo de cadenas retrasadas después de una ronda de incorporación. Por lo tanto, pueden tolerarse condiciones inferiores a las óptimas para la incorporación de los nucleótidos sin marcar para la incorporación de los nucleótidos sin marcar. Por lo tanto, pueden tolerarse condiciones que no pretenden optimizar la eficiencia y precisión de la enzima extensora de cadena.

Puede usarse cualquier composición adecuada que comprenda la mezcla de nucleótidos sin marcar y bloqueados. En algunas realizaciones, la composición comprende una solución tamponada que tiene un pH de aproximadamente 8 a aproximadamente 10, tal como de aproximadamente 9 a aproximadamente 9,5. Puede usarse cualquier tampón adecuado, tal como un tampón de Tris. Los nucleótidos sin marcar y bloqueados pueden estar presentes en cualquier concentración adecuada tal como de aproximadamente 0,001 a aproximadamente 15 μ M, tal como de aproximadamente 0,5 μ M a aproximadamente 5 μ M, tal como aproximadamente 2 μ M. La composición puede comprender cualquier cantidad adecuada de un quelante, tal como de aproximadamente 0 mM a aproximadamente 5 mM de ácido etilendiaminotetraacético (EDTA), o de aproximadamente 0 mM a aproximadamente 2 mM de EDTA, o de aproximadamente 0 mM a aproximadamente 1 mM de EDTA. La composición puede comprender cualquier cantidad adecuada de sulfato de magnesio, tal como $MgSO_4$ de aproximadamente 1 mM a aproximadamente 10 mM, $MgSO_4$ de aproximadamente 2 mM a aproximadamente 6 mM o $MgSO_4$ aproximadamente 4 mM. La composición puede comprender cualquier cantidad adecuada de un detergente. Por ejemplo, la composición puede comprender de aproximadamente el 0,05 % a aproximadamente el 1 % de detergente en peso, tal como de aproximadamente el 0,1 % a aproximadamente el 0,5 % de detergente en peso, o aproximadamente el 0,2 % de detergente en peso. Puede usarse cualquier detergente adecuado. Por ejemplo, puede usarse monolaurato de polioxietilensorbitano (también denominado monolaurato de polioxietilensorbitano o Tween) o detergentes de hidrato de 3-[(3-colamidopropil)dimetilamonio]-1-propanosulfonato (CHAPS). La composición puede comprender cualquier cantidad adecuada de un antioxidante. Por ejemplo, el detergente puede comprender uno o más antioxidantes en una concentración de antioxidante total combinada de aproximadamente 1 mM a aproximadamente 50 mM, tal como de aproximadamente 2 mM a aproximadamente 40 mM, de aproximadamente 3 mM a aproximadamente 20 mM o de aproximadamente 15 mM a aproximadamente 25 mM. Los antioxidantes adecuados incluyen ascorbato, acetovainillona y Trolox. Preferiblemente, la composición está exenta de nucleótidos marcados.

En algunas realizaciones, la composición que comprende una mezcla de nucleótidos sin marcar y bloqueados comprende un tampón de Tris, tiene un pH de aproximadamente 9 a aproximadamente 9,5, tiene un detergente, $MgSO_4$, EDTA, ascorbato y acetovainillona. Por ejemplo, la composición puede incluir componentes como se describe en la **Tabla 1** a continuación:

REACTIVO	CONCENTRACIÓN
Tampón de Tris	200 mM
pH	de 9 a 9,5
Nucleótidos sin marcar y bloqueados	2 μ M cada uno
CHAPS	0,2 % en peso
$MgSO_4$	4 mM
EDTA	de 0 a 1 mM
Ascorbato	de 3 a 20 mM
Acetovainillona	de 10 a 15 mM

Se describen materiales, composiciones y componentes que pueden usarse, pueden usarse junto con, pueden usarse en preparación para o son productos de los métodos y composiciones descritos. En la presente memoria se describen estos y otros materiales, y se entiende que cuando se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales, aunque no se describa explícitamente la referencia específica de cada una de las diversas combinaciones y permutaciones individuales y colectivas, cada uno se contempla y describe específicamente en la presente memoria. Por ejemplo, si se describe y analiza un método y se analiza una serie de modificaciones que pueden realizarse en las etapas del método, todas y cada una de las combinaciones y permutaciones de las etapas del método, y las modificaciones que son posibles se contemplan específicamente a menos que se indique específicamente lo contrario. Análogamente, también se contempla y se describe específicamente cualquier subconjunto o combinación de éstos. Este concepto se aplica a todos los aspectos de esta descripción. Por lo tanto, si existe una diversidad de etapas adicionales que pueden realizarse, se entiende que cada una de estas etapas

adicionales puede realizarse con cualquier etapa de método específica o combinación de etapas de método de los métodos descritos, y que cada combinación o subconjunto de combinaciones se contempla específicamente y debe considerarse descrita.

5 **Ejemplos**

Ejemplo 1. El aumento del ajuste de fase se correlaciona con un menor tiempo de incorporación

10 Para comprender el efecto de la reducción del tiempo de incorporación sobre el ajuste de fase y la tasa de error, se secuenció la secuenciación de polipéptidos de una secuencia conocida utilizando un secuenciador MiniSeq™ de Illumina utilizando tiempos de incorporación diferentes. Específicamente, los tiempos de incorporación fueron de 46 segundos, 23 segundos y 12 segundos. La correlación entre el ajuste de fase y la tasa de error se muestra en la FIG. 5. Como se muestra, la tasa de error y el ajuste de fase se mantienen bajos en tiempos de incorporación largos (46 segundos). Sin embargo, a medida que disminuye el tiempo de incorporación (por ejemplo, 12 segundos), el ajuste de fase y la tasa de error aumentan. En consecuencia, la disminución del tiempo de incorporación tiende a dar como resultado un aumento del ajuste de fase y de la tasa de error.

20 Ejemplo 2. Disminución del ajuste de fase a tiempos de incorporación reducidos con nucleótidos sin marcar y bloqueados durante la detección

25 Para evaluar si el ajuste de fase y la tasa de error se ven afectados por la incubación con polimerasa y nucleótidos sin marcar y bloqueados durante una etapa de detección, se secuenció un polinucleótido de una secuencia conocida utilizando un secuenciador MiniSeq™ de Illumina, durante la etapa de detección, (i) la mezcla de barrido del fabricante (sin nucleótidos sin marcar y bloqueados y sin polimerasa, los “reactivos patrón”) y (ii) una mezcla de barrido modificada que incluía nucleótidos sin marcar y bloqueados y una polimerasa [“ScanAndFill”]. La mezcla ScanAndFill incluía polimerasa Pol812. Se ejecutaron 80 ciclos de secuenciación utilizando tiempos de incorporación de 25 segundos y 7,5 segundos con la mezcla de barrido patrón y un tiempo de incorporación de 7,5 segundos con la mezcla ScanAndFill y la mezcla ScanAndFill. Se determinó la tasa de error, el ajuste de fase y el preajuste de fase (donde la cadena copia incorpora un nucleótido adicional más allá de las cadenas copia en fase) en cada ciclo. Los resultados se presentan en la FIG. 6 y la Tabla 2 a continuación.

Tabla 2. Tasa de error, ajuste de fase y preajuste de fase

Mezcla de barrido	Tiempo de inc.	Densidad	% de PF	% de Alin.	Ajuste de fase	Preajuste de fase	% de Error
Reactivos patrón	2 x 25s	122	95,6	99,2	0,07	0,08	0,13
Reactivos patrón	2 x 7,5s	122	95,6	99,2	0,65	0,04	0,86
ScanAndFill	2 x 7,5s	133	92,5	97,9	0,10	0,09	0,35

45 Como se muestra en las FIG. 6 y la Tabla 2, el uso de la mezcla ScanAndFill produjo métricas algo comparables con las basales (reactivos patrón a 2x25 segundos), reduciendo al mismo tiempo el tiempo total de química por ciclo en aproximadamente un 32 % (de 112 segundos a 77 segundos). Como se indica en la Tabla 2, la reducción de la tasa de error parece deberse a una reducción del ajuste de fase.

50 Ejemplo 3. Disminución del ajuste de fase a tiempos de incorporación reducidos con nucleótidos sin marcar y bloqueados durante la detección sin polimerasa adicional

55 Para evaluar el efecto de la polimerasa añadida en la mezcla ScanAndFill sobre el ajuste de fase, el preajuste de fase y la tasa de error, se secuenció un polinucleótido de una secuencia conocida utilizando un secuenciador MiniSeq™ de Illumina, durante la etapa de detección, una mezcla ScanAndFill con y sin polimerasa. Las mezclas ScanAndFill eran las mismas, excepto por que la mezcla con la polimerasa incluía Pol1671 60 µg/ml, que se describe en la Solicitud de patente provisional US 62/753.558, titulada “Polimerasas, composiciones y métodos de uso”, presentada el 31 de octubre de 2018. Los otros componentes de las mezclas ScanAndFill fueron tampón de EA; pH 9,85; MgSO4; EDTA; trifosfatos de nucleótidos A, C y T modificados para incluir un colorante fluorescente con un enlazador LN3 escindible como se describe, por ejemplo, en las Solicitudes de patente publicadas US 2013/0079232 y US 2016/0040225 (A-LN3, C-LN3 y T-LN3); Dark G (sin marcar) y CHAPS. Se usó un tiempo de incorporación de 7,5 segundos. Se calcularon los valores de Q30, ajuste de fase y preajuste de fase en cada ciclo. “Q30” se refiere al porcentaje de lecturas que pasan el filtro de calidad Q30, es decir, una tasa de error inferior o igual a 1 en 1000, o el 0,1 %. Los resultados se presentan en la FIG. 7.

65 Como se muestra, el ajuste de fase se mantuvo bajo en presencia y ausencia de polimerasa añadida, lo que sugiere que la polimerasa de la etapa de incorporación permaneció activa y presente durante la etapa de detección.

Sorprendentemente, se observó una menor tasa de ajuste de fase en ausencia de polimerasa adicional. La tasa de error global se redujo, como se refleja en el valor de Q30 superior, al igual que la tasa de ajuste de fase y preajuste de fase.

5 Ejemplo 4. La adición de antioxidante reduce adicionalmente la tasa de error y la decadencia de la señal

10 Para evaluar el efecto de un antioxidante añadido en la mezcla ScanAndFill (sin polimerasa) sobre la decadencia de la señal de la etapa de detección y la tasa de error de secuenciación, se secuenció un polinucleótido de una secuencia conocida utilizando un secuenciador MiniSeq™ de Illumina, durante la etapa de detección, una mezcla ScanAndFill sin polimerasa y con o sin ascorbato 3 mM. Las mezclas ScanAndFill incluían tampón de EA, pH 9,85, MgSO₄, EDTA, A-LN3, C-LN3, T-LN3, Dark G y CHAPS. Se ejecutaron 100 ciclos de secuenciación utilizando tiempos de incorporación de 7,5 segundos y se compararon con un valor basal utilizando la mezcla de barrido patrón con un tiempo de incorporación de 25 segundos por ciclo. Los resultados se muestran en las FIG. 8-9.

15 La adición de ascorbato dio como resultado una reducción de la decadencia de la señal sustancialmente mejorada (FIG. 8) y una reducción en la tasa de error (FIG. 9) con respecto a ScanAndFill sin ascorbato. Se permiten longitudes de lectura superiores a 75 ciclos con la adición del antioxidante. Como se muestra en el recuadro de la FIG. 9, el ascorbato también mejoró el % de Q30. Además, el ascorbato mejoró el % de Alin PhiX.

20 Se observaron resultados similares con otras concentraciones de ascorbato (hasta 20 mM), con concentraciones variables de acetovainillona (hasta 20 mM) y combinaciones de ascorbato y acetovainillona (datos no mostrados).

Además, puede variarse el pH y pueden obtenerse resultados similares.

25 Se ha descrito un número de realizaciones. No obstante, se entenderá que se podrán realizar diversas modificaciones.

REIVINDICACIONES

1. Un método de secuenciación de polinucleótidos que comprende:
 - 5 (a) introducir una primera enzima extensora de cadena y una mezcla de nucleótidos marcados y bloqueados en una célula de flujo que comprende un sitio en el que múltiples cadenas polinucleotídicas molde que tienen la misma secuencia de nucleótidos están unidas a una superficie de la célula de flujo, en donde la primera enzima extensora de cadena se configura para incorporar uno adecuado de los nucleótidos marcados y bloqueados en cadenas polinucleotídicas copia, basándose en la secuencia de las cadenas molde a las que corresponden las cadenas polinucleotídicas copia;
 - 10 (b) retirar por lavado los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo;
 - (c) introducir una composición que comprende una mezcla de nucleótidos sin marcar y bloqueados en la célula de flujo durante o después de la retirada por lavado de los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo, en donde un nucleótido sin marcar y bloqueado de la mezcla de los nucleótidos sin marcar y bloqueados está disponible para su incorporación en las cadenas polinucleotídicas copia, basándose en una secuencia de las cadenas molde a las que corresponden las cadenas polinucleotídicas copia, siempre que el nucleótido incorporado anteriormente en la cadena copia, si lo hay, no esté bloqueado; y
 - 15 (d) detectar la identidad de un nucleótido marcado y bloqueado incorporado en las cadenas copia, si lo hay, mientras la mezcla de nucleótidos sin marcar y bloqueados se incubaba con la célula de flujo.
2. El método según la reivindicación 1, en donde la composición que comprende la mezcla de nucleótidos sin marcar y bloqueados se introduce en la célula de flujo después de retirar por lavado los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo.
- 25 3. Un método según la reivindicación 1 o 2, en donde retirar por lavado los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo comprende introducir la composición que comprende la mezcla de nucleótidos sin marcar y bloqueados en la célula de flujo.
- 30 4. El método de una cualquiera de las reivindicaciones 1 a 3, en donde el nucleótido sin marcar y bloqueado de la mezcla de los nucleótidos sin marcar y bloqueados está disponible para su incorporación en las cadenas polinucleotídicas copia cuando se detecta la identidad del primer nucleótido marcado y bloqueado incorporado en las cadenas copia.
- 35 5. El método según una cualquiera de las reivindicaciones 1 a 4, en donde detectar la identidad del primer nucleótido marcado y bloqueado incorporado en las cadenas copia comprende formar imágenes de la célula de flujo.
- 40 6. El método según una cualquiera de las reivindicaciones 1 a 4, en donde detectar la identidad del primer nucleótido marcado y bloqueado incorporado en las cadenas copia se produce durante un período de tiempo de 120 segundos o menos.
- 45 7. El método según una cualquiera de las reivindicaciones 1 a 4, en donde detectar la identidad del primer nucleótido marcado y bloqueado incorporado en las cadenas copia se produce durante un período de tiempo de 10 segundos o más.
8. El método según una cualquiera de las reivindicaciones 1 a 7, que comprende además:
 - 50 (e) retirar un marcador y resto de bloqueo de cualquier primer nucleótido bloqueado y marcado incorporado en las cadenas copia, y retirar un resto de bloqueo de cualquier segundo nucleótido sin marcar y bloqueado incorporado en las cadenas copia.
9. El método según la reivindicación 8, que comprende además repetir las etapas (a)-(d).
- 55 10. El método según cualquiera de las reivindicaciones 1 a 9, en donde la primera enzima extensora de cadena y la mezcla de nucleótidos marcados y bloqueados se incuban con la célula de flujo durante 15 segundos o menos antes de retirar por lavado los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo.
- 60 11. Un método según cualquiera de las reivindicaciones 1 a 9, en donde la primera enzima extensora de cadena y la mezcla de nucleótidos marcados y bloqueados se incuban con la célula de flujo durante 10 segundos o menos antes de retirar por lavado los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo.
- 65 12. Un método según cualquiera de las reivindicaciones 1 a 9, en donde la primera enzima extensora de cadena y la mezcla de nucleótidos marcados y bloqueados se incuban con la célula de flujo durante 7,5 segundos o menos antes de retirar por lavado los nucleótidos marcados y bloqueados sin incorporar de la célula de flujo.

13. El método según una cualquiera de las reivindicaciones 1 a 12, en donde la composición que comprende la mezcla de nucleótidos sin marcar y bloqueados comprende una segunda enzima extensora de cadena configurada para incorporar uno adecuado de los nucleótidos sin marcar y bloqueados en las cadenas polinucleotídicas copia, basándose en una secuencia de las cadenas molde a las que corresponden las cadenas polinucleotídicas copia, siempre que el nucleótido incorporado anteriormente en la cadena copia, si lo hay, no esté bloqueado.
14. El método según la reivindicación 13, en donde la primera enzima extensora de cadena y la segunda enzima extensora de cadena son las mismas enzimas o enzimas diferentes.
15. El método según una cualquiera de las reivindicaciones 1-14, en donde la composición que comprende la mezcla de nucleótidos sin marcar y bloqueados tiene un pH de 9 a 9,5 y comprende un tampón, un detergente, sulfato de magnesio, EDTA y un antioxidante.

5

10

15

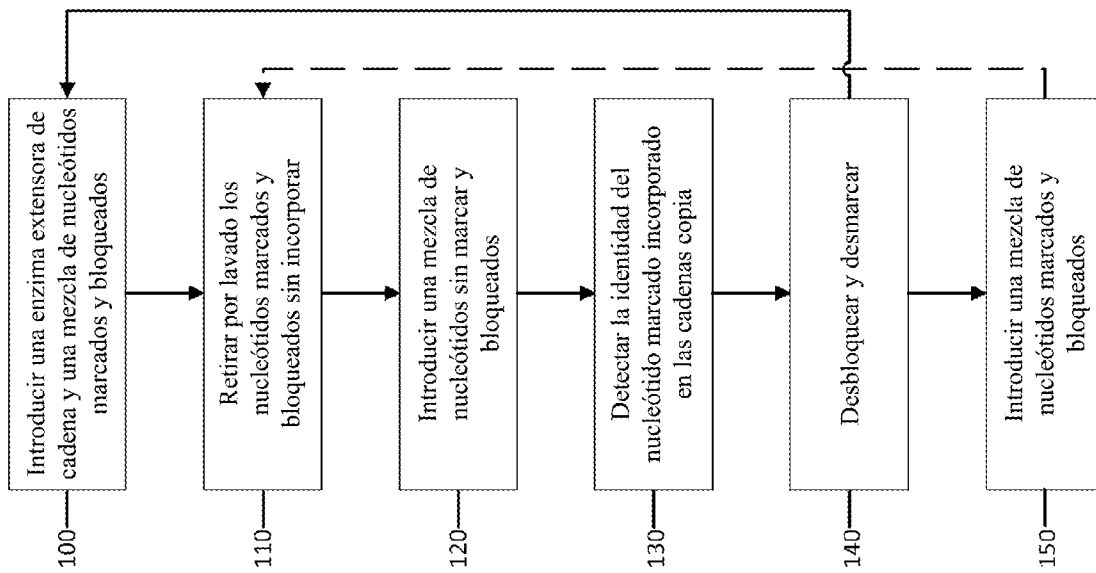


Figura 1

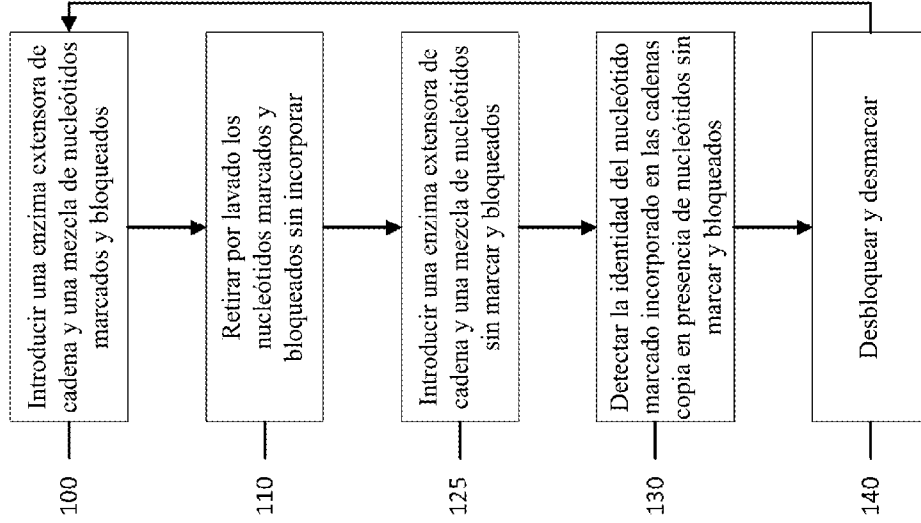


Figura 2

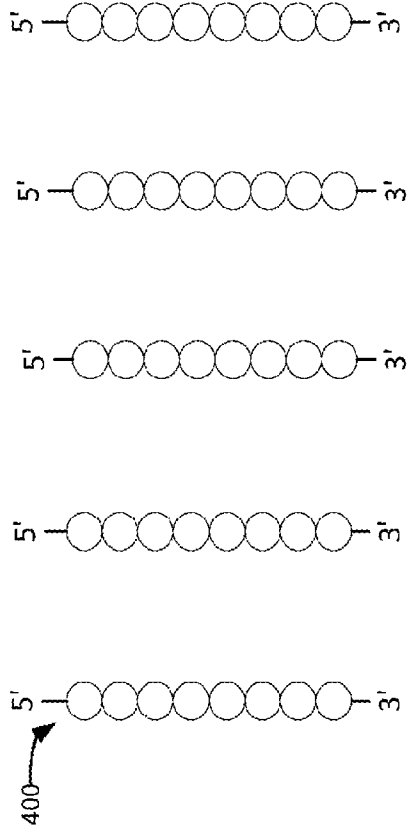


Figura 4A

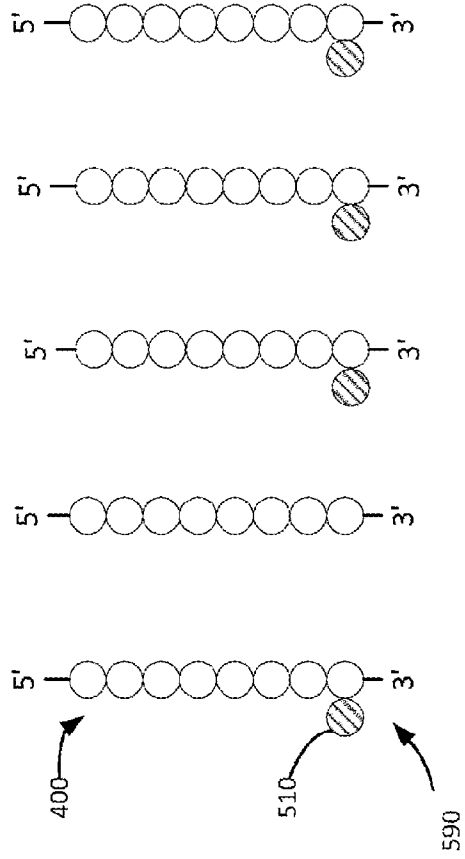


Figura 4B

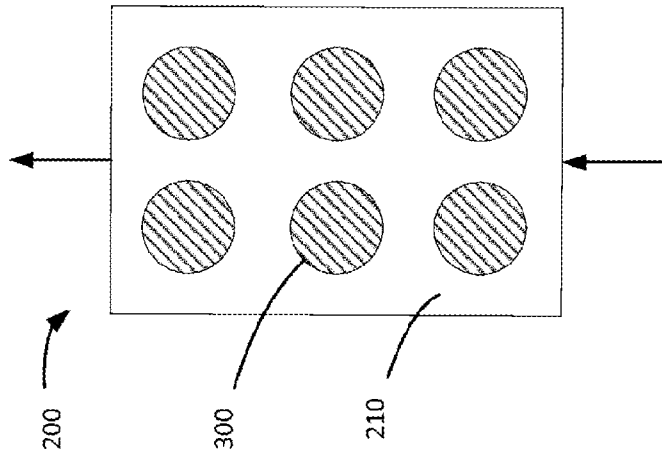


Figura 3

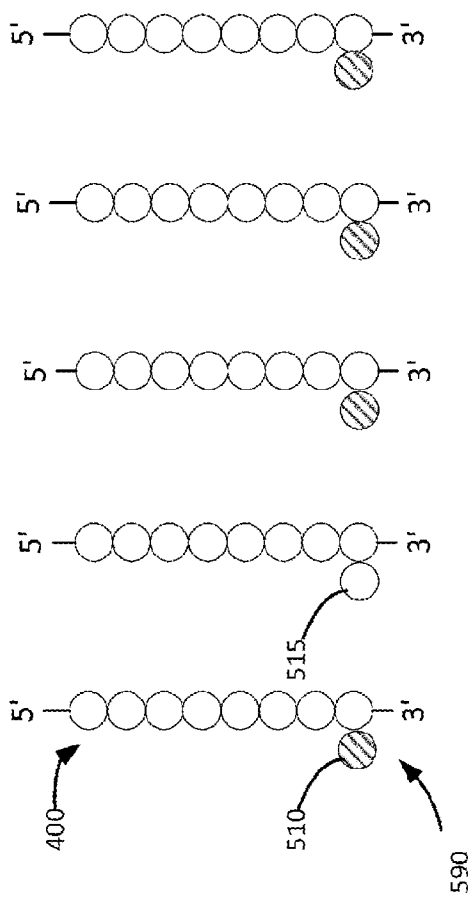


Figure 4C

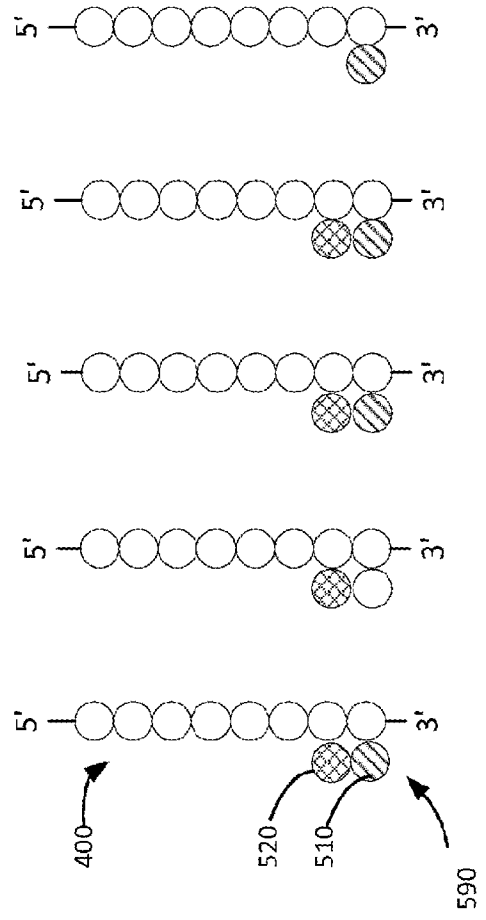


Figure 4D

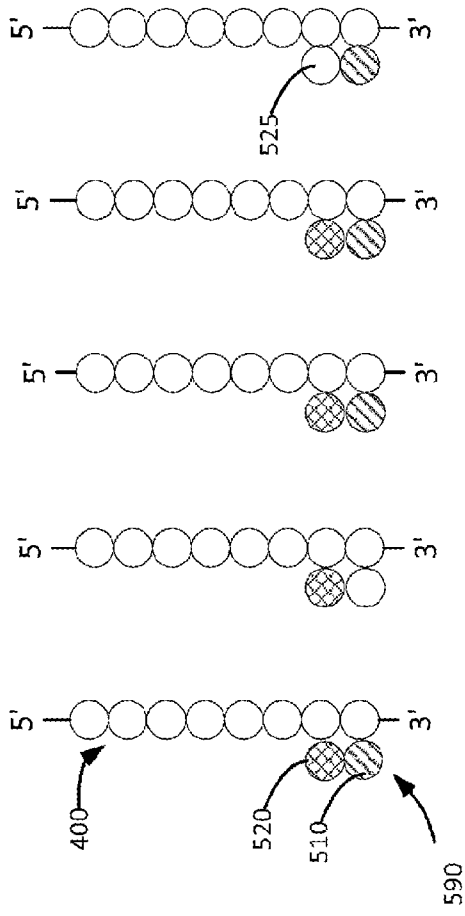


Figura 4E

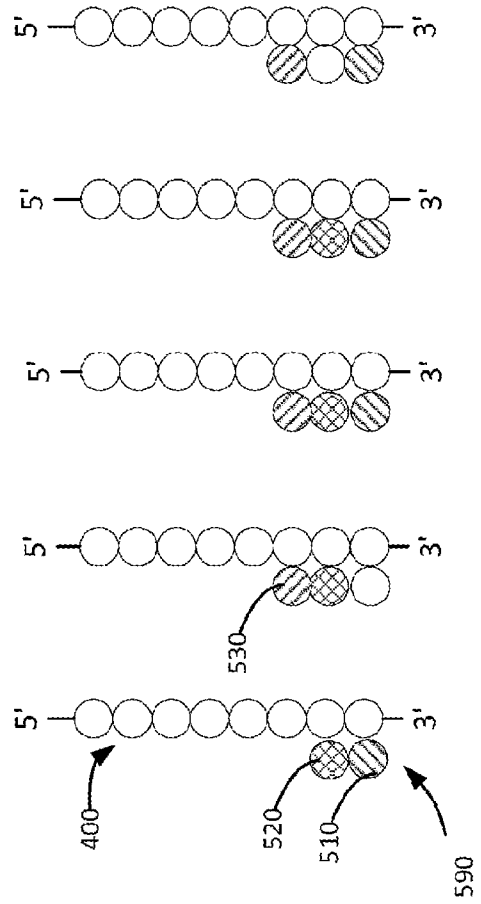


Figura 4F

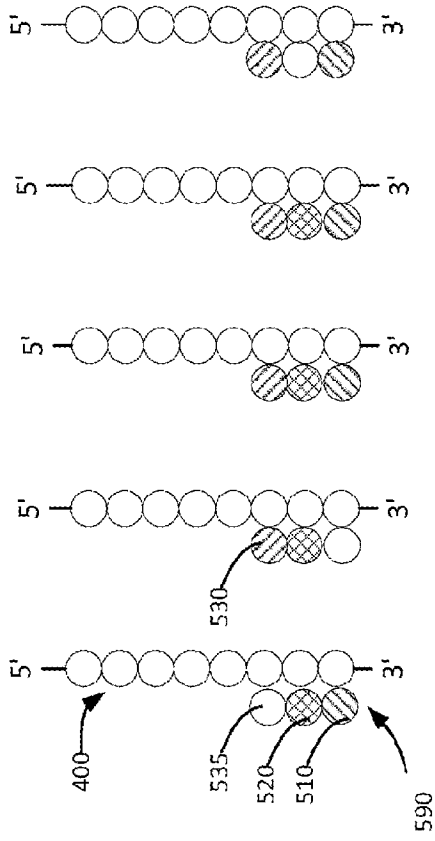


Figura 4G

Tasa de error de Phix frente a ajuste de fase

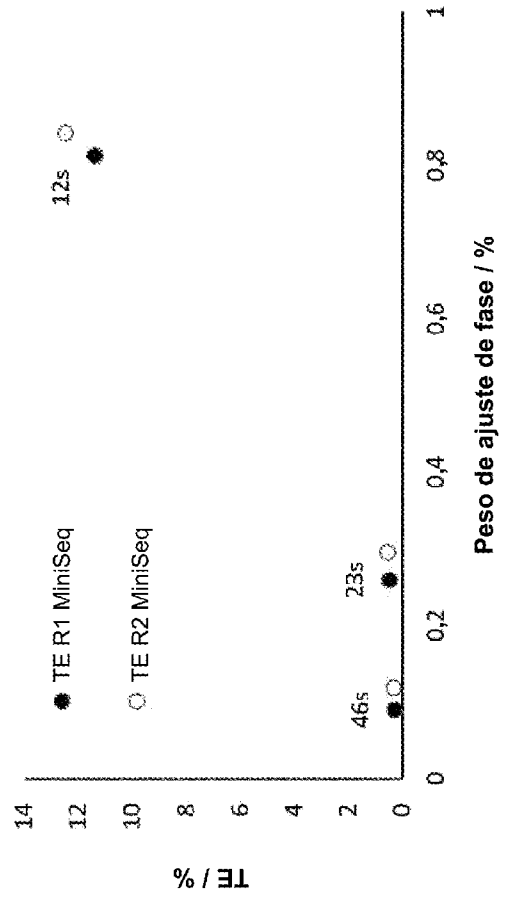


Figura 5

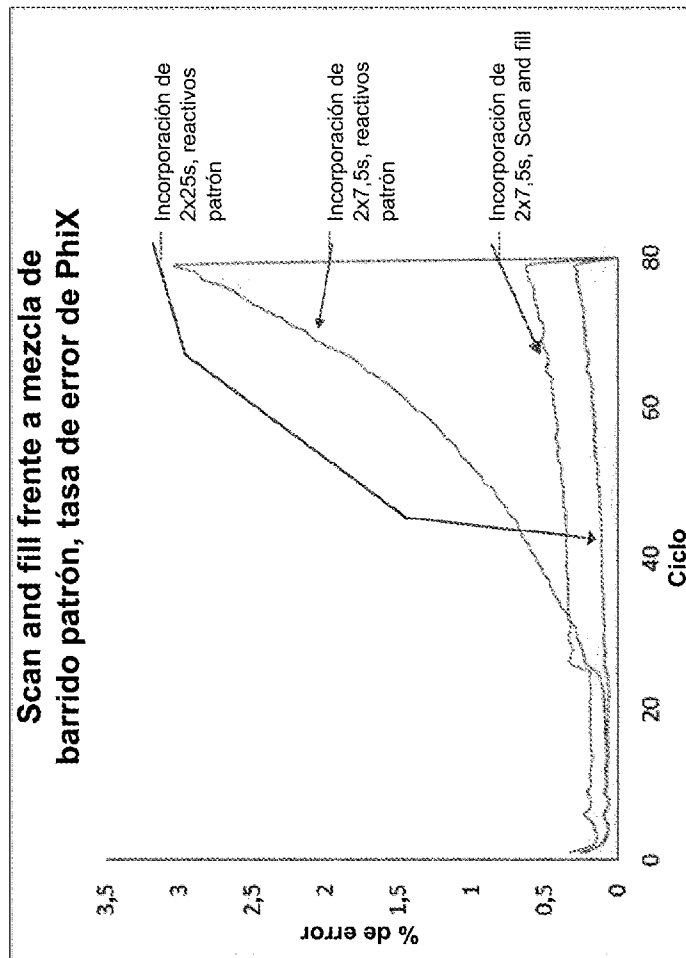


Figura 6

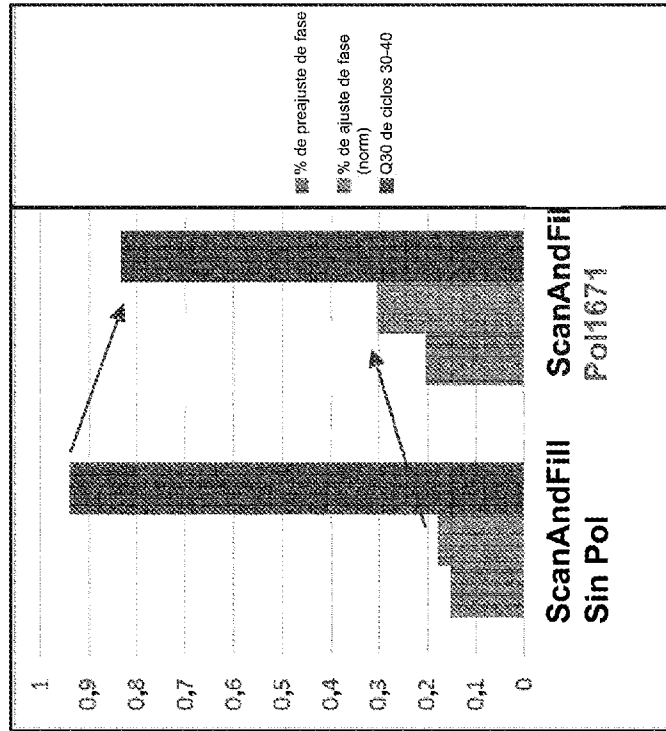


Figura 7

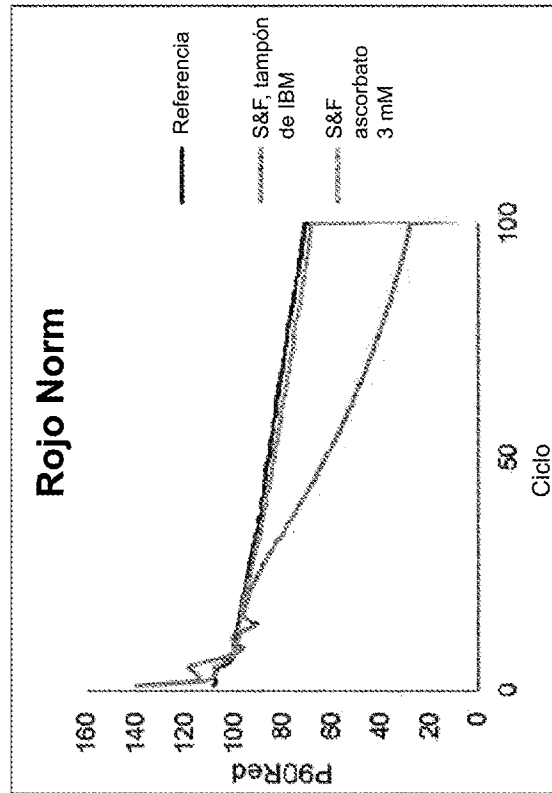


Figura 8

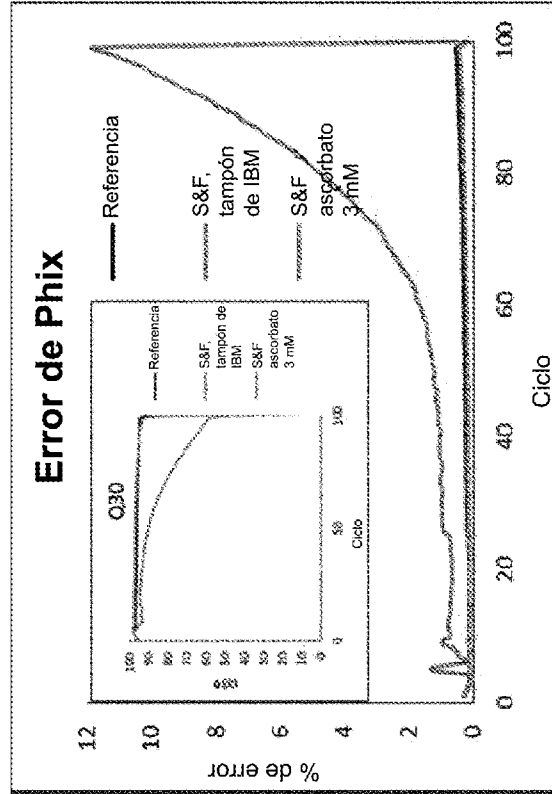


Figura 9