

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 817 425**

51 Int. Cl.:

**C12Q 1/6869** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **22.05.2014** **E 17207528 (5)**

97 Fecha y número de publicación de la concesión europea: **15.07.2020** **EP 3351645**

54 Título: **Secuenciado pirofosforolítico usando nanoporos**

30 Prioridad:

**24.05.2013 US 201361827175 P**

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

**07.04.2021**

73 Titular/es:

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

72 Inventor/es:

**MEULEMAN, WOUTER**

74 Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 817 425 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

Secuenciado pirofosforolítico usando nanoporos

**Antecedentes**

5 Esta descripción se refiere generalmente al análisis de ácidos nucleicos, y más específicamente a la síntesis de ácidos nucleicos usando nanoporos.

10 Las plataformas comerciales actualmente disponibles para secuenciar ADN son relativamente costosas. Estas plataformas utilizan un método de 'secuenciación por síntesis', llamado así porque los polímeros de ADN se sintetizan al detectar la adición de cada monómero (es decir, nucleótido) a la estructura del polímero en crecimiento. Debido a que una cadena de ADN plantilla dirige estrictamente la síntesis de un nuevo polímero de ADN, se puede deducir la secuencia del ADN plantilla de la serie de monómeros de nucleótidos que se añadieron a la cadena de crecimiento durante la síntesis. La capacidad para detectar adiciones de monómeros se ve facilitada por variantes especialmente modificadas genéticamente de los componentes bioquímicos que normalmente llevan a cabo la síntesis de ADN en sistemas biológicos. Estos componentes modificados genéticamente son costosos de construir y se consumen en cantidades relativamente grandes durante la secuenciación por síntesis. Además, el seguimiento de la reacción utiliza equipo informático relativamente costoso, tales como láseres, ópticas de detección y sistemas complejos de suministro de fluidos. Las plataformas comerciales más exitosas hasta la fecha también requieren reactivos costosos y equipo informático para amplificar las plantillas de ADN antes de que pueda comenzar incluso la secuenciación por síntesis. Los

20 Se han considerado otros métodos de secuenciación para reducir el costo, aumentar el rendimiento y/o simplificar el proceso. Uno de estos métodos se basa en formar una única cadena de ADN a través de un nanoporo e identificar su secuencia a partir de la variación en la corriente iónica que circula a través del poro a medida que se forma la cadena. Una alternativa a este método de secuenciación 'nanoporo-cadena' es la secuenciación 'nanoporo-exonucleasa', que implica la eliminación de monofosfatos de nucleótidos catalizada por exonucleasa, uno de una vez, a partir de una cadena de ADN y pasando sucesivamente los monofosfatos de nucleótidos liberados a través de un nanoporo. Sin embargo, las variaciones resultantes en la corriente iónica que circula a través de los nanoporos son bastante pequeñas y es difícil distinguir un nucleótido de otro. Se han realizado intentos para modificar el ADN antes de la digestión o para modificar los monofosfatos de nucleótidos una vez que se han liberado. En los documentos WO 03/080861 y US 6,268,146 se describen por ejemplo métodos de secuenciación de ácidos nucleicos diana.

30 Sin embargo, a pesar de estos esfuerzos, la secuenciación de nanoporo-exonucleasa aún no se ha demostrado a un nivel comercialmente viable hasta la fecha.

Por lo tanto, existe una necesidad de plataformas más rentables, rápidas y convenientes que proporcionen una alternativa a las actualmente disponibles para la secuenciación de ácidos nucleicos. La presente descripción aborda esta necesidad y proporciona otras ventajas también.

**Breve compendio**

35 La presente invención proporciona un método para determinar la secuencia de un ácido nucleico diana. El método incluye las etapas siguientes: (a) proporcionar un ácido nucleico diana; (b) poner en contacto el ácido nucleico diana con una polimerasa para eliminar sucesivamente trifosfatos de nucleótidos del ácido nucleico diana, en donde los trifosfatos de nucleótidos que se eliminan tienen una variedad de diferentes restos de bases; y (c) distinguir los diferentes restos de bases para los trifosfatos de nucleótidos que se eliminan, determinando de este modo la secuencia del ácido nucleico diana.

En algunas realizaciones se puede llevar a cabo un método para determinar la secuencia de un ácido nucleico diana en el que el ácido nucleico diana tiene dos cadenas;

45 La presente descripción también proporciona un aparato que incluye (a) una barrera impermeable a los fluidos que separa un primer depósito de fluido de un segundo depósito de fluido; (b) un nanoporo situado en la barrera impermeable a los fluidos para formar un paso a través del cual puede pasar un trifosfato de nucleótido desde el primer depósito de fluido al segundo depósito de fluido; y (c) una mezcla de reacción en el primer depósito de fluido, comprendiendo la mezcla de reacción una polimerasa, ácido nucleico diana bicatenario y una concentración pirofosforolítica de pirofosfato.

**Breve descripción de los dibujos**

La FIG. 1 presenta un diagrama de una reacción de secuenciación pirofosforolítica que utiliza una polimerasa unida a un nanoporo.

La FIG. 2 presenta un diagrama de una reacción de secuenciación pirofosforolítica que utiliza una polimerasa unida a un nanoporo y un ácido nucleico plantilla unido a una barrera impermeable a fluidos.

La FIG. 3. presenta secuenciación pirofosforolítica con siembra en membrana de múltiples plantillas de ácido nucleico.

### Descripción detallada

5 La presente descripción proporciona un método para secuenciar ácidos nucleicos en modo inverso en comparación con las técnicas de secuenciación habituales por síntesis (SBS). El método de la presente descripción aprovecha una función catalítica de la polimerasa conocida como pirofosforólisis que generalmente se considera injustamente como culpable de artificios no deseados en técnicas de SBS. La pirofosforólisis da como resultado la eliminación de trifosfatos de nucleótidos de una cadena de ácido nucleico por una polimerasa, y como tal es la inversa de la reacción de polimerización que impulsa las técnicas habituales de SBS.

10 La pirofosforólisis puede distinguirse de la actividad de exonucleasa (que está presente en algunas polimerasas), por ejemplo, en función del diferente mecanismo catalítico para las dos reacciones, diferentes zonas activas en la estructura de la polimerasa donde se producen las dos reacciones, y los diferentes productos para las reacciones. Con respecto al mecanismo catalítico y las diferencias de la zona activa, se sabe que la actividad de exonucleasa puede eliminarse de muchas especies de polimerasas por eliminación de determinados dominios, mientras que la actividad de pirofosforólisis se cree que está catalizada por el mismo dominio que cataliza la polimerización.

15 Además, el ciclo de reacción catalizado por exonucleasa es la conversión del ADN<sub>n</sub> (ADN de longitud n) en ADN<sub>n-1</sub> (ADN que es un nucleótido más corto que el ADN de longitud n) y un monofosfato de nucleótido. Por el contrario, un ciclo de pirofosforólisis produce ADN<sub>n-1</sub> y un trifosfato de nucleótido a partir del ADN<sub>n</sub> y pirofosfato. Principalmente, el pirofosfato se consume en una reacción de pirofosforólisis pero no se consume en una reacción de exonucleasa.

20 Los métodos de secuenciación pirofosforolítica utilizan la detección de nanoporos. Los nanoporos pueden usarse para detectar sucesivamente los trifosfatos de nucleótidos que se liberan de un ácido nucleico para determinar la secuencia del ácido nucleico. Los métodos proporcionan una combinación de ventajas que generalmente se satisfacen solo parcialmente por secuenciación de nanoporo-exonucleasa o secuenciación nanoporo-cadena. Específicamente, los métodos de secuenciación pirofosforolíticos de la presente descripción abordan algunos de los retos incurridos en la secuenciación nanoporo-exonucleasa, es decir, bajas probabilidades de captura y

25 detección, conservando su principal ventaja sobre la secuenciación de cadenas, es decir, la resolución de una sola base. Esta ventaja procede del hecho de que la afinidad de los nanoporos por los monofosfatos de nucleótidos es más bien débil (del orden de la afinidad micromolar), incluso en presencia de un adaptador de am6-ciclodextrina que se ha utilizado para mejorar la señal en algunos sistemas de nanoporo. Véase Clarke *et al.*, *Nat. Nanotechnol. Apr.*; 4 (4): 265-70 (2009). Para la distinción satisfactoria de diferentes tipos de nucleótidos en un contexto de secuenciación, se desea una afinidad de intervalo nanomolar. El ATP tiene una afinidad que está en un buen intervalo, incluso sin el empleo de un adaptador en el poro. Véase Cheley *et al.*, *Chem. Biol.* 9:829-838 (2002). Sin pretender imponer ninguna teoría, se supone que la afinidad mejorada del ATP sobre el monofosfato de nucleótido se debe a la triple carga negativa transportada por el ATP, que puede hacer que se una más fuertemente dentro del nanoporo. Además, la triple carga negativa puede ayudar a la captura de la molécula en presencia de un campo eléctrico, especialmente cuando el campo es muy débil, como es el caso fuera del nanoporo, donde los nucleótidos se liberan realmente.

Además de la captura mejorada y la detección de trifosfatos de nucleótidos, hay varias otras numerosas ventajas que pueden ser proporcionadas por los métodos expuestos en la presente memoria, tales como el uso de una polimerasa frente a una exonucleasa. Se ha demostrado que las polimerasas forman un buen "ajuste" con nanoporos con el fin de secuenciar nanoporo-cadena (Cherf *et al.*, *Nat. Biotech.*, 30:344-348 (2012); Manrao *et al.*, *Nat. Biotech.* 30: 349-353 (2012)). Un ajuste igualmente bueno aún no se ha demostrado con exonucleasas. Además, el sustrato para polimerasas es ADN bicatenario que generalmente no entra en el nanoporo. Por el contrario, el ADN monocatenario, el sustrato para la mayoría de las exonucleasas, puede plantear el problema de

45 entrar y bloquear el nanoporo. Por último, al contrario que en el secuenciado a base de exonucleasas o en el secuenciado de cadenas a base de polimerasa, la velocidad de una reacción de secuenciación pirofosforolítica puede controlarse ajustando la concentración de pirofosfatos.

Los términos utilizados en la presente memoria toman su significado ordinario a menos que se especifique lo contrario. Se exponen a continuación ejemplos de varios términos utilizados en este documento y sus definiciones.

50 Tal como se emplea en la presente memoria, el término "unido" quiere decir conectado por fuerzas que evitan la separación por difusión. El término puede incluir enlaces que son de naturaleza covalente o no covalente. Por ejemplo, dos proteínas se pueden unir covalentemente por su secuencia primaria (p. ej., un enlace peptídico o fusión de proteínas) o entre sus secuencias primarias (p. ej., un enlace disulfuro o reticulación química mediante grupos R de aminoácido). Se pueden unir dos proteínas de forma no covalente, por ejemplo, mediante uno o más enlaces de

55 hidrógeno, enlaces iónicos, fuerzas de Van der Waals, enlaces hidrófobos o similares.

Tal como se emplea en la presente memoria, la expresión "cada uno", cuando se emplea en referencia a una colección de elementos, está destinado a identificar un solo elemento en la colección, pero no necesariamente se refiere a cada elemento de la colección. Puede haber excepciones si la descripción explícita o el contexto imponen inequívocamente lo contrario.

Como se emplea en la presente memoria, la expresión "actividad de exonucleasa" quiere decir hidrólisis del enlace fosfodiéster que une un nucleótido al extremo de un ácido nucleico de longitud n para producir un monofosfato de nucleótido y un ácido nucleico de longitud n-1. La hidrólisis puede tener lugar en el enlace que une el nucleótido 5' al ácido nucleico (es decir, actividad exonucleasa 5' a 3') o en el enlace que une el nucleótido 3' al ácido nucleico (es decir, actividad exonucleasa 3' a 5').

Como se emplea en la presente memoria, la expresión "barrera impermeable a los fluidos" quiere decir cualquier cosa que impida el paso de un fluido. Por ejemplo, una barrera impermeable a los fluidos puede estar presente entre dos depósitos de manera que un fluido en el primer depósito está separado del fluido en el segundo depósito, y los fluidos no se mezclen. Una barrera impermeable a los fluidos puede incluir un poro o abertura que permita el paso de un fluido que de otra manera estaría obstruido por el resto de la barrera. En determinadas realizaciones, el fluido puede ser un fluido acuoso y la barrera puede ser impermeable a fluidos acuosos.

Como se emplea en la presente memoria, la expresión "carece de actividad de exonucleasa" quiere decir que no tiene actividad de exonucleasa medible. Por ejemplo, una polimerasa u otro agente que carece de actividad de exonucleasa 3' a 5' no tendrá actividad de exonucleasa 3' a 5' medible. De forma similar, una polimerasa u otro agente que carece de actividad de exonucleasa 5' a 3' no tendrá actividad de exonucleasa 5' a 3' mensurable. Las polimerasas conocidas en la técnica como "exo-negativas" (o "exo-") ya sean naturales o genéticamente modificadas son ejemplos no restrictivos de polimerasas que carecen de actividad de exonucleasa. Las variantes conocidas incluyen aquellas que son 5' a 3' exo negativas y/o 3' a 5' exo negativas.

Como se emplea en la presente memoria, el término "nanoporo" quiere decir un pequeño orificio que permite el paso de trifosfatos de nucleótidos a través de una barrera impermeable. La barrera es generalmente una capa eléctricamente aislante y el nanoporo generalmente permite que los iones circulen de un lado de la barrera al otro, impulsados por un potencial aplicado. El nanoporo preferiblemente permite que los nucleótidos circulen a través de la luz del nanoporo por el potencial aplicado. El nanoporo también puede permitir que un ácido nucleico, tal como ADN o ARN, sea empujado o arrastrado a través de la luz del nanoporo. Sin embargo, en determinadas realizaciones, el nanoporo no necesita permitir el paso de un ácido nucleico bicatenario o monocatenario. Un nanoporo utilizado en una determinada realización puede tener un diámetro de luz mínimo de no más de 10 nm, 5 nm, 4 nm, 3 nm, 2 nm, 1 nm, 0,5 nm o menos. Un tipo de nanoporo es una "nanoporo de proteína" que es un polipéptido o una colección de polipéptidos que forma el orificio pequeño para permitir el paso de trifosfatos de nucleótidos a través de una barrera tal como una bicapa lipídica. Los ejemplos de nanoporos proteicos incluyen nanoporo de alfa hemolisina, porin A de *mycobacterium smegmatis* (MspA) y sus variantes. Otro tipo de nanoporo es un "nanoporo de estado sólido", que es un pequeño orificio fabricado a través de un material sólido. El material sólido puede ser, por ejemplo, grafeno o silicio.

Como se emplea en la presente memoria, el término "nucleótido" pretende incluir ribonucleótidos, desoxinucleótidos o análogos de los mismos. Por ejemplo, el término se emplea en la presente memoria para referirse en general a un resto de nucleósido (ya sea ribosa, desoxirribosa o uno de sus análogos) incluido un resto de base y opcionalmente unido a uno o más restos de fosfato. Los nucleótidos ejemplares incluyen, pero no se limitan a, monofosfato de ribonucleótido (a veces denominado monofosfato de ribonucleósido), difosfato de ribonucleótido (a veces denominado difosfato de ribonucleósido), trifosfato de ribonucleótido (a veces denominado trifosfato de ribonucleósido), monofosfato de desoxinucleótido (a veces denominado monofosfato de desoxinucleósido), difosfato de desoxinucleótido (a veces denominado difosfato de desoxinucleósido) y trifosfato de desoxinucleótido (a veces denominado trifosfato de desoxinucleósido). Para mayor claridad cuando se desea distinguir los componentes de ARN de los componentes de ADN, el término "ribonucleótido" puede emplearse para especificar nucleótidos de ARN, tales como trifosfato de ribouridina, trifosfato de riboguanidina, trifosfato de ribocitidina o trifosfato de riboadenosina; y el término "desoxinucleótido" puede emplearse para especificar nucleótidos de ADN, tales como trifosfato de desoxitimidina, trifosfato de desoxiguanidina, trifosfato de desoxicitidina y trifosfato de desoxiadenosina. En determinadas realizaciones, los nucleótidos son "extensibles", por ejemplo, que carecen de un resto de bloqueo de extensión en el hidroxilo 3' o en cualquier otra posición en el nucleótido.

Como se emplea en la presente memoria, el término "pirofosforólisis" quiere decir la reacción entre el pirofosfato y la unidad 3'-nucleótido de un ácido nucleico para liberar el nucleótido en forma del correspondiente trifosfato de nucleótido. Otro producto de la reacción es el ácido nucleico que carece del nucleótido correspondiente (es decir, la reacción convierte un ácido nucleico de longitud n en un ácido nucleico de longitud n-1). La reacción típicamente está catalizada por una polimerasa y es la inversa de la reacción de polimerización llevada a cabo por la polimerasa en condiciones biológicas habituales.

Como se emplea en la presente memoria, la expresión "concentración pirofosforolítica", cuando se usa en referencia a pirofosfato, quiere decir una concentración de pirofosfato que da lugar a que se produzca una reacción de pirofosforólisis a un nivel sustancial. Por ejemplo, una concentración pirofosforolítica de pirofosfato puede dar como resultado una polimerasa que presente una velocidad de pirofosforólisis más alta que de polimerización. Por lo tanto, una concentración pirofosforolítica de pirofosfato puede dar como resultado una reversión sustancial de la actividad de polimerización que, si no, sería catalizada por una polimerasa.

Como se emplea en la presente memoria, el término "depósito" quiere decir un receptáculo o cámara para contener

o restringir la circulación de fluido. Un depósito puede estar completamente cerrado, al menos temporalmente. Alternativamente, un depósito puede estar parcialmente cerrado, por ejemplo, teniendo una o más aberturas (p. ej., una o más entradas o salidas). Un fluido puede circular a través de un depósito, por ejemplo, en realizaciones en las que el depósito está en una celda de flujo.

5 Como se emplea en la presente memoria, el término "especie" se emplea para identificar moléculas que compartan la misma estructura química. Por ejemplo, una mezcla de nucleótidos puede incluir varias moléculas de dCTP. Se entenderá que las moléculas de dCTP son de la misma especie que las demás. Asimismo, cada una de las moléculas de ADN que tienen la misma secuencia de nucleótidos son la misma especie.

10 Las realizaciones expuestas a continuación y mencionadas en las reivindicaciones se pueden entender a la vista de las definiciones anteriores.

La presente invención proporciona un método para determinar la secuencia de un ácido nucleico diana definido en las reivindicaciones.

15 Una realización ejemplar se muestra en la FIG. 1. Como se muestra, una polimerasa se une a una molécula de ADN bicatenario y, en presencia de pirofosfato en exceso, cataliza una reacción de pirofosforólisis para liberar trifosfatos de nucleótidos del extremo 3' de una de las cadenas. En este ejemplo, se ha producido un trifosfato de desoxiguanidina seguido de un trifosfato de desoxitimidina. La polimerasa está acoplada a un nanoporo y el trifosfato de desoxiguanidina está en la luz del nanoporo, mientras que el trifosfato de desoxitimidina está en proceso de ser liberado en la luz del nanoporo. Como tales, los desoxitriposfatos de nucleótidos entran en el nanoporo sucesivamente, en el mismo orden en que se liberaron de la cadena de ADN mediante la acción pirofosforolítica de la polimerasa. Los trifosfatos de desoxinucleótidos tienen una carga neta negativa debido a los trifosfatos y se conducen a través del nanoporo por un potencial a través de la membrana (como indica el signo positivo en el lado de la membrana donde ocurre la pirofosforólisis y un signo negativo en el lado opuesto de la membrana). Generalmente, los trifosfatos de nucleótidos reactivos no están presentes en una reacción de pirofosforólisis. En algunos casos, pueden estar presentes cantidades en trazas de trifosfatos de nucleótidos, pero en cantidades que no catalizan sustancialmente una reacción de extensión del cebador directo por la polimerasa. Por lo tanto, en la mayoría de las realizaciones, los únicos trifosfatos de nucleótidos que están sustancialmente presentes en una reacción de pirofosforólisis son los producidos por la acción catalítica de la polimerasa en el ácido nucleico.

20 Una reacción similar se ejemplifica en la FIG. 2 excepto que la cadena plantilla (es decir, la cadena que no está siendo escindida por pirofosforólisis) está unida a la membrana. En este caso, la cadena plantilla tiene un resto de esterol unido por enlace covalente (p. ej., colesterol) que se une al interior hidrófobo de la bicapa lipídica de la membrana.

25 Un método de la presente descripción se puede usar con cualquiera de una variedad de ácidos nucleicos diana. El ácido nucleico diana puede tener una estructura natural como se encuentra, por ejemplo, en ADN o ARN. El ADN contiene naturalmente monómeros que tienen bases de timina, guanina, citosina o adenina. Una cadena de ADN que se somete a pirofosforólisis puede producir trifosfato de desoxitimidina, trifosfato de desoxiguanidina, trifosfato de desoxicitidina y trifosfato de desoxiadenosina, respectivamente. El ADN también puede contener algunas variantes de las cuatro bases nucleotídicas tales como 5-metil citosina, 5-hidroximetilcitosina y otras bases metiladas. Los trifosfatos de desoxinucleótidos que tienen estas bases variantes pueden producirse y/o detectarse usando un método o aparato expuesto en la presente memoria. Por ejemplo, puede distinguirse la presencia o ausencia de metilación en la citosina para facilitar los análisis epigenéticos. El ARN en la naturaleza contiene monómeros que tienen bases de uracilo, guanina, citosina o adenina. Una cadena de ARN que se somete a pirofosforólisis puede producir trifosfato de ribouridina, trifosfato de riboguanidina, trifosfato de ribocitidina o trifosfato de riboadenosina, respectivamente.

30 Un ácido nucleico puede incluir modificaciones no naturales tales como bases no naturales, modificaciones en los restos de fosfato o modificaciones en los restos de azúcar. Las bases no naturales ejemplares que se pueden incluir en un ácido nucleico, ya sea que tengan un eje principal natural o estructura análoga, incluyen, sin limitación, inosina, xatanina, hipoxatanina, isocitosina, isoguanina, 2-aminoadenina, 6-metiladenina, 6-metilguanina, 2-propilguanina, 2-propiladenina, 2-tiouracilo, 2-tiotimina, 2-tiocitosina, 15-haloacilo, 15-halocitosina, 5-propiniluracilo, 5-propinilcitosina, 6-azouracilo, 6-azocitosina, 6-azotimina, 5-uracilo, 4-tiouracilo, 8-haloadenina o guanina, 8-aminoadenina o guanina, 8-tioladenina o guanina, 8-tioalquiladenina o guanina, 8-hidroxiadenina o guanina, 5-halouracilo o citosina sustituido, 7-metilguanina, 7-metiladenina, 8-azaguanina, 8-azaadenina, 7-deazaguanina, 7-deazaadenina, 3-deazaguanina, 3-deazaadenina o similares.

35 Las bases no naturales pueden seleccionarse, por ejemplo, para impartir un tamaño mayor o menor, o para impartir una carga aumentada o disminuida, para influir en la capacidad de los análogos de trifosfato de nucleótidos resultantes para distinguirse por un nanoporo u otro componente de detección. Asimismo, dichas bases pueden ser beneficiosas si se seleccionan para impartir una velocidad deseada de pirofosforólisis. Pueden incorporarse bases no naturales a un ácido nucleico usando métodos conocidos tales como la amplificación o la replicación de un ácido nucleico plantilla en presencia de los análogos de nucleótidos. Una o más de las copias de ácido nucleico resultantes se pueden usar luego como ácido(s) nucleico(s) diana en el aparato o método de secuenciación

expuesto en la presente memoria.

En determinadas realizaciones, un ácido nucleico que se usa en un método o aparato en la presente memoria carecerá de una o más de las bases no naturales u otros restos no naturales expuestos en la presente memoria. Por ejemplo, en determinadas realizaciones, los métodos no se usan para eliminar un nucleótido terminador del extremo 3' de una cadena de ácido nucleico. Por consiguiente, en algunas realizaciones, un aparato o método puede no incluir ningún ácido nucleico que tenga un nucleótido terminador en su extremo 3'. Alternativamente o además, en algunas realizaciones, un aparato o método puede no incluir ningún nucleótido terminador.

Como se ejemplifica en la FIG. 1 y en otra parte en la presente memoria, un ácido nucleico diana puede ser ADN bicatenario, por ejemplo, cuando se usa una ADN polimerasa para pirofosforólisis. También se puede usar un heterodúplex, formado entre una cadena de ADN y una cadena de ARN. Por ejemplo, una ARN polimerasa puede usarse para catalizar la pirofosforólisis en el extremo 3' de una cadena de ARN que se hibrida con una cadena de ADN plantilla, produciendo de este modo trifosfatos de ribonucleótidos.

Un ácido nucleico que se utiliza en un método o aparato en la presente memoria puede aislarse de una fuente biológica y usarse directamente o procesarse para producir un producto amplificado o modificado. Alternativamente, se puede utilizar un ácido nucleico sintético, otra vez, directamente o después del tratamiento. El tratamiento puede incluir, por ejemplo, uno o más de aislamiento de componentes naturales, escisión para formar fragmentos, amplificación (p. ej., por PCR, Rolling Circle u otras técnicas conocidas), ligadura de secuencias adaptadoras o secuencias de etiquetas, tagmentación usando un transposón o métodos de "preparación de muestras" usados antes de las técnicas de secuenciación de ácidos nucleicos. Las técnicas de tratamiento útiles son conocidas en la técnica como se expone, por ejemplo, en Sambrook *et al.*, *Molecular Cloning: A Laboratory Manual*, 3ª edición, Cold Spring Harbor Laboratory, Nueva York (2001) o en Ausubel *et al.*, *Current Protocols in Molecular Biology*, John Wiley and Sons, Baltimore, Md. (1998).

Los ejemplos de métodos de preparación de muestras que pueden utilizarse para tratar ácidos nucleicos antes de la secuenciación basada en pirofosforólisis incluyen métodos que se han desarrollado para la amplificación del genoma completo o la amplificación completa del exoma en combinación con técnicas de secuenciación masivamente paralelas. Por ejemplo, se pueden utilizar técnicas de preparación de muestras disponibles comercialmente de Illumina, Inc. (San Diego, CA), Life Technologies (Carlsbad, CA), 454 Life Sciences (una filial de Roche, Basilea, Suiza) o New England Biolabs (Ipswich, MA). Otros métodos útiles de preparación de muestras que se pueden utilizar se describen, por ejemplo, en las patentes de EE.UU. nº 6.107.023 o nº 7.741.463; o la solicitud de patente de EE.UU. Pub. nº 2010/0120098 A1. También se pueden utilizar métodos de preparación de muestras dirigidos para aislar un subconjunto del contenido de la secuencia de una muestra compleja para secuenciación posterior. Los métodos comerciales ejemplares que pueden usarse para la captura dirigida de un subconjunto de fragmentos de ácido nucleico incluyen, pero no están limitados a SureSelect™ kits (Agilent, Santa Clara, CA), TruSeq® Enrichment kits (Illumina, Inc., San Diego, CA) o Nextera® Enrichment kits (Illumina, Inc., San Diego, CA).

Los ácidos nucleicos se pueden aislar de cualquiera de una variedad de fuentes incluidas, sin limitación, un mamífero tal como un roedor, ratón, rata, conejo, conejillo de Indias, ungulado, caballo, oveja, cerdo, cabra, vaca, gato, perro, primate, primate humano o no humano; una planta tal como *Arabidopsis thaliana*, maíz (*Zea mays*), sorgo, avena (*oryza sativa*), trigo, arroz, canola o soja; una alga tal como *Chlamydomonas reinhardtii*; un nematodo tal como *Caenorhabditis elegans*; un insecto tal como *Drosophila melanogaster*, mosquito, mosca de la fruta, abeja melífera o araña; un pez tal como el pez cebra (*Danio rerio*); un reptil; un anfibio tal como una rana o *Xenopus laevis*; un *dictyostelium discoideum*; un hongo tal como *Pneumocystis carinii*, *Takifugu rubripes*, levadura, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*; o un *Plasmodium falciparum*. Los ácidos nucleicos también pueden proceder de genomas más pequeños, tales como los de un procarionte, tal como una bacteria, *Escherichia coli*, estafilococos o *mycoplasma pneumoniae*; una arquea; un virus tal como el virus de la Hepatitis C o el virus de la inmunodeficiencia humana; o un viroide.

Puede usarse cualquiera de una variedad de polimerasas en un método o aparato expuesto en la presente memoria incluidas, por ejemplo, enzimas a base de proteínas aisladas de sistemas biológicos y sus variantes funcionales. Se entenderá que la referencia a una determinada polimerasa, tales como las ejemplificadas a continuación, incluyen sus variantes funcionales a menos que se indique lo contrario. Los ejemplos de polimerasas útiles incluyen ADN polimerasas y ARN polimerasas. Las ADN polimerasas ejemplares incluyen las que han sido clasificadas por homología estructural en familias identificadas como A, B, C, D, X, Y y RT. Las ADN polimerasas en la familia A incluyen, por ejemplo, la T7 ADN polimerasa, la ADN polimerasa y mitocondrial eucariota, el ADN Pol I de *E. coli*, Pol I de *Thermus aquaticus*, y Pol I de *Bacillus stearothermophilus*. Las ADN polimerasas en la Familia B incluyen, por ejemplo, ADN polimerasas  $\alpha$ ,  $\delta$ , y  $\epsilon$  eucariotas; ADN polimerasa  $\zeta$ ; T4 ADN polimerasa, *Phi29* DNA polimerasa y DNA polimerasa de bacteriófago RB69. La familia C incluye, por ejemplo, la subunidad ADN polimerasa III alfa de *E. coli*. La familia D incluye, por ejemplo, polimerasas derivadas del subdominio Euryarchaeota de Archaea. Las ADN polimerasas en la familia X incluyen, por ejemplo, polimerasas eucarióticas Pol  $\beta$ , pol  $\sigma$ , Pol  $\lambda$  y Pol  $\mu$ , y Pol4 de *S. cerevisiae*. Las ADN polimerasas en la familia Y incluyen, por ejemplo, Pol  $\eta$ , Pol iota, Pol kappa, Pol IV de *E. coli* (DINB) y Pol V de *E. coli* (UmuD'2C). La familia RT (transcriptasa inversa) de ADN polimerasas incluye, por ejemplo, transcriptasas inversas de retrovirus y telomerasas eucarióticas. Las ARN polimerasas ejemplares incluyen, pero no se limitan a, ARN polimerasas víricas tales como T7 ARN polimerasa; ARN polimerasas eucariotas tales como ARN

polimerasa I, ARN polimerasa II, ARN polimerasa III, ARN polimerasa IV y ARN polimerasa V; y RNA polimerasa de Archaea.

Las clasificaciones anteriores se proporcionan con fines ilustrativos. Se entenderá que son posibles variaciones en el sistema de clasificación. Por ejemplo, en al menos un sistema de clasificación, las polimerasas de la familia C se han clasificado como una subcategoría de la familia X. Además, las polimerasas se pueden clasificar según otras características, funcionales o estructurales, que pueden solaparse o no con las características estructurales ejemplificadas anteriormente. Algunas características ejemplares se exponen con mayor detalle a continuación.

Muchas polimerasas tienen una actividad de exonucleasa de corrección 3' a 5' intrínseca que puede ser útil para algunas realizaciones. Se prefieren las polimerasas que sustancialmente carecen de actividad de exonucleasa de corrección 3' a 5'. La ausencia de actividad de exonucleasa puede ser una característica natural o una característica impartida por una variante o polimerasa genéticamente modificada. Por ejemplo, el fragmento exo negativo de Klenow es una versión mutada del fragmento de Klenow que carece de actividad de exonucleasa de corrección 3' a 5'. El fragmento Klenow y su variante exo negativa pueden ser útiles en un método o aparato expuesto epm. Las polimerasas que catalizan la pirofosforólisis, la inversión directa de la polimerización en la misma zona activa, son particularmente útiles.

Dependiendo de la realización que se vaya a utilizar, una polimerasa puede ser termófila o inactivable por calor. Las polimerasas termófilas son generalmente útiles para condiciones de alta temperatura o en condiciones termocíclicas tales como las empleadas para técnicas de reacción en cadena de la polimerasa (PCR). Los ejemplos de polimerasas termófilas incluyen, pero no se limitan a, 9°N ADN polimerasa, Taq ADN polimerasa, Phusion® ADN polimerasa, Pfu ADN polimerasa, RB69 ADN polimerasa, KOD ADN polimerasa, y VentR® ADN polimerasa. La mayoría de las polimerasas aisladas de organismos no termófilos son inactivables por calor. La inactivación por calor puede ser útil para detener una reacción de secuenciación expuesta en la presente memoria. La reacción puede reiniciarse opcionalmente añadiendo un nuevo suministro de polimerasa a la reacción a la temperatura permisiva apropiada. Ejemplos de polimerasas inactivables por calor son las de fagos. Debe entenderse que las polimerasas de cualquiera de una variedad de fuentes pueden modificarse para aumentar o disminuir su tolerancia a condiciones de alta temperatura.

Puede utilizarse una variante modificada genéticamente de una polimerasa que tiene actividad aumentada de pirofosforólisis. Las variantes ejemplares son las que se han creado para su uso en técnicas de PCR incluidas, pero no limitadas a, las variantes descritas en la patente de EE.UU. nº 8,071,536.

Se puede inducir una polimerasa para eliminar sucesivamente nucleótidos de una de dos cadenas de ácido nucleico por pirofosforólisis en un método expuesto en la presente memoria. La polimerasa puede colocarse en condiciones para que se produzca la pirofosforólisis. Por ejemplo, la polimerasa puede ponerse en contacto con un ácido nucleico bicatenario y una concentración pirofosforolítica de pirofosfato. Cualquier concentración de pirofosfato puede usarse que provoca que se produzca una reacción de pirofosforólisis a un nivel sustancial incluida, pero no limitada a, al menos aproximadamente 100  $\mu\text{M}$  de pirofosfato. Se pueden emplear concentraciones más altas de pirofosfato, por ejemplo, para conducir la pirofosforólisis a una velocidad más rápida. Por consiguiente, se puede usar una concentración de al menos aproximadamente 250  $\mu\text{M}$  de pirofosfato, al menos aproximadamente 500  $\mu\text{M}$  de pirofosfato, al menos aproximadamente 750  $\mu\text{M}$  de pirofosfato, al menos aproximadamente 1 mM de pirofosfato, al menos aproximadamente 2 mM de pirofosfato, al menos aproximadamente 5 mM de pirofosfato, al menos aproximadamente 10 mM de pirofosfato, al menos aproximadamente 20 mM de pirofosfato o superior.

La capacidad de alterar la velocidad de pirofosforólisis es una ventaja para ajustar la velocidad de la reacción de secuenciación, por ejemplo, para acomodar una velocidad óptima o deseada de detección de trifosfato de nucleótidos para un dispositivo de detección dado. Por ejemplo, la velocidad de pirofosforólisis puede reducirse utilizando una concentración menor de pirofosfato que las expuestas anteriormente. Por lo tanto, como una alternativa o adición a las concentraciones umbral expuestas anteriormente, una concentración máxima de pirofosfato presente en un aparato o método expuesto en la presente memoria puede ser como máximo aproximadamente pirofosfato 20 mM, como máximo aproximadamente pirofosfato 10 mM, como máximo aproximadamente pirofosfato 5 mM, como máximo aproximadamente pirofosfato 2 mM, como máximo aproximadamente pirofosfato 1 mM, como máximo aproximadamente pirofosfato 750  $\mu\text{M}$ , como máximo aproximadamente pirofosfato 500  $\mu\text{M}$ , como máximo aproximadamente pirofosfato 250  $\mu\text{M}$ , como máximo aproximadamente pirofosfato 100  $\mu\text{M}$ , o menos.

Los trifosfatos de nucleótidos reactivos están generalmente ausentes en condiciones de pirofosforólisis. Por lo tanto, en la mayoría de las realizaciones, los únicos trifosfatos de nucleótidos que están sustancialmente presentes en una reacción de pirofosforólisis son los producidos por la acción catalítica de la polimerasa en el ácido nucleico. Si los trifosfatos de nucleótidos están presentes en condiciones de pirofosforólisis, los trifosfatos de nucleótidos estarán presentes en lo que puede considerarse una concentración no catalítica. Una concentración no catalítica de trifosfato de nucleótido es una concentración que es insuficiente para permitir que se produzca actividad de extensión de la polimerasa sustancial o detectable. Por ejemplo, una concentración no catalítica de trifosfato de nucleótido es una concentración que está sustancialmente por debajo de la constante de unión de asociación para la unión de los trifosfatos de nucleótidos a la polimerasa. Por consiguiente, la pirofosforólisis puede llevarse a cabo poniendo en

contacto una polimerasa con un ácido nucleico bicatenario en presencia de una concentración pirofosforolítica de pirofosfato y no más de una concentración no catalítica de trifosfato de nucleótido.

Las condiciones ilustrativas para provocar pirofosforólisis que se pueden utilizarse en la presente memoria se exponen en Patel *et al. Biochem.* 30:511-525 (1991), o las patentes de EE. UU. nº 7,090,975; nº 7,914,995 o nº 8,071,536. En varios casos, estas referencias describen reacciones que también incluyen componentes usados para la extensión de un ácido nucleico. Debe entenderse que la extensión no se utiliza en determinadas realizaciones de la presente descripción. Uno o más de los componentes que se describen en Patel *et al. Biochem.* 30:511-525 (1991) o las patentes de EE. UU. nº 7,090,975; nº 7,914,995 o nº 8,071,536, incluidos, por ejemplo, los componentes usados para la reacción de extensión de la polimerasa pueden excluirse de un método o aparato expuesto en la presente memoria.

Se pueden incluir amortiguadores, sales, iones metálicos, glicerol, DMSO y otros componentes generalmente incluidos en el almacenamiento de la polimerasa o mezclas de reacción en un aparato o método de la presente descripción, según se desee. Las cantidades de los componentes a incluir serán evidentes para los expertos en la técnica o determinables, por ejemplo, mediante técnicas de valoración rutinarias.

Un aspecto beneficioso de algunas realizaciones es la capacidad de detener o interrumpir el método de secuenciación alterando las condiciones de reacción para inhibir la pirofosforólisis. El método de secuenciación puede reiniciarse opcionalmente alterando las condiciones de reacción para permitir que se reanude la pirofosforólisis. Cualquiera de una variedad de las condiciones de reacción expuestas en la presente memoria, o en las referencias citadas en la presente memoria, puede alterarse entre la pirofosforólisis interrumpida y la pirofosforólisis reanudada, permitiendo de ese modo un cambio eficaz entre la secuenciación interrumpida y la activa, respectivamente.

En determinadas realizaciones, un método de la presente descripción puede incluir una etapa de interrupción de la producción sucesiva de trifosfatos de nucleótidos eliminando el pirofosfato del contacto con una polimerasa que está catalizando la pirofosforólisis, seguido de una etapa de reanudación de la producción sucesiva de los trifosfatos de nucleótidos poniendo en contacto la polimerasa con pirofosfato. El pirofosfato se puede administrar a una reacción usando técnicas apropiadas para el sistema de fluido que se utiliza incluido, por ejemplo, pipeteo de alícuotas de líquido, movimiento de bolos de fluido bajo presión positiva o negativa (p. ej. mediante bombas o gravedad), electroforesis, isotacoforesis, manipulación de gotitas (p. ej. electrohumección) o similar. Se pueden usar técnicas fluidicas similares para eliminar el pirofosfato, por ejemplo, por desplazamiento y/o reemplazo con una solución de lavado. Por supuesto, dichas técnicas fluidicas pueden utilizarse para añadir o eliminar otros componentes utilizados en los métodos y aparatos expuestos en la presente memoria.

Alternativamente o además de los métodos fluidicos expuestos anteriormente, el pirofosfato puede eliminarse de la reacción por secuestro, degradación o inactivación. Por ejemplo, se pueden utilizar manipulaciones físicas tales como adsorción a un agente secuestrante o degradación por calor, luz o electricidad. Pueden usarse métodos químicos para modificar la estructura o actividad del pirofosfato o para degradar la molécula. También pueden utilizarse métodos enzimáticos tales como la degradación por pirofosfatasa, como se muestra para las reacciones de PCR en los documentos US 4,800,159 y US 5,498,523 y para las reacciones de secuenciación en gel en el documento US 4,962,020.

Alternativamente o además de las técnicas expuestas anteriormente, la pirofosforólisis puede detenerse o interrumpirse eliminando otros componentes de la reacción. Por ejemplo, la polimerasa puede eliminarse de una reacción y opcionalmente reemplazarse o volver a un estado activo. Por ejemplo, la polimerasa puede eliminarse por métodos fluidicos, de secuestro, degradación o inactivación tales como los ejemplificados anteriormente para pirofosfato. En determinadas realizaciones, puede usarse una polimerasa sensible al calor (no termófila) en una reacción de pirofosforólisis y a continuación inactivarse por calor. De forma similar, una polimerasa puede degradarse por modificación química o degradación enzimática (p. ej., mediante una proteasa). Ya sea degradada por técnicas físicas, químicas o enzimáticas, la polimerasa usada puede eliminarse por lavado y a continuación la pirofosforólisis puede reanudarse mediante la adición de una nueva polimerasa al ácido nucleico que se está secuenciando.

La actividad de la polimerasa también puede cambiarse mediante la adición y eliminación de inhibidores, alternando entre temperaturas permisivas y no permisivas para polimerasas termoestables, o la presencia y ausencia de un agente secuestrante o sustrato competitivo. La pirofosforólisis también puede detenerse e iniciarse por desnaturalización y renaturalización, respectivamente, del ácido nucleico que se está secuenciando.

Aunque en la presente memoria se han ejemplificado métodos y aparatos para las realizaciones que usan pirofosfato para conducir la pirofosforólisis, debe entenderse que en su lugar se pueden usar análogos de pirofosfato. Un ejemplo de análogo es pirovanadato, que puede utilizarse, por ejemplo, como se describe en Akabayov *et al. J. Biol. Chem.* 286: 29146-29157 (2011). Como ejemplos adicionales, pueden usarse análogos de pirofosfato que tienen restos adicionales. Generalmente se seleccionan análogos de pirofosfato que no inhiben por completo la pirofosforólisis o el paso de los trifosfatos de nucleótidos resultantes, o uno de sus análogos, a través de un nanoporo. Sin embargo, los análogos de pirofosfato pueden alterar las características de la pirofosforólisis y/o la

detección de nanoporos. Por ejemplo, puede ser beneficioso usar un análogo de pirofosfato para ralentizar o acelerar la pirofosforólisis para proporcionar una velocidad de detección deseada u óptima. Asimismo, los análogos de los trifosfatos de nucleótidos que se producen cuando se utiliza un análogo de pirofosfato en una reacción de pirofosforólisis también pueden impartir características deseadas para la detección de nanoporos. Por ejemplo, restos que alteran carga o tamaño, en comparación con difosfato solo, pueden aumentar o disminuir la velocidad de paso de análogos de trifosfato de nucleótido a través de un nanoporo, o alterar las interacciones de los análogos de trifosfato de nucleótido con el nanoporo, para proporcionar mejores resultados de secuenciación.

Sin embargo, en algunas realizaciones, un método o aparato de la presente descripción excluirá pirofosfato que tenga cualquier resto añadido. Por ejemplo, se puede utilizar pirofosfato que carece de un resto ópticamente detectable, tal como un resto fluorescente.

En realizaciones particulares, los trifosfatos de nucleótidos se detectan utilizando nanoporos. Por ejemplo, los trifosfatos de nucleótidos que se eliminan sucesivamente de un ácido nucleico por pirofosforólisis pueden pasarse a través de un nanoporo para su detección. Mediante la utilización de un nanoporo apropiado, diferentes restos de bases de los trifosfatos de nucleótido se pueden distinguir para permitir la detección de secuencias. Generalmente, se puede utilizar un aparato que incluye un primer y un segundo compartimentos separados por una barrera física, en donde la barrera tiene uno o más nanoporos. El primer compartimento puede incluir componentes utilizados para una reacción de pirofosforólisis. El aparato puede configurarse para aplicar un campo eléctrico a través de la barrera de manera que los trifosfatos de nucleótidos se dirijan desde el primer compartimento a través del poro al segundo compartimento. El aparato puede configurarse para medir la firma electrónica de un trifosfato de nucleótido que pasa a través del nanoporo. En consecuencia, un aparato útil puede incluir un circuito eléctrico capaz de aplicar un potencial y medir una señal eléctrica a través de una barrera y nanoporo. Los procedimientos pueden llevarse a cabo utilizando una técnica de fijación de membranas o una fijación de voltaje.

Un procedimiento de la presente descripción puede llevarse a cabo utilizando cualquier sistema adecuado en el que un poro penetra a través de una barrera. La barrera en muchas realizaciones es preferiblemente una bicapa lipídica. Las bicapas lipídicas pueden prepararse utilizando métodos conocidos en la técnica, por ejemplo, como se describe en Montal y Mueller *Proc. Natl. Acad. Sci. USA* 69:3561-3566 (1972) o en el documento WO 2008/102120. Pueden formarse bicapas lipídicas a partir de cualquiera de una variedad de lípidos incluidos, pero no limitados a, fosfolípidos, glicolípidos, colesterol y sus mezclas.

Ejemplos de nanoporos que se pueden utilizar incluyen, por ejemplo, nanoporos de proteínas tales como nanoporo de alfa hemolisina, porin A de *Mycobacterium smegmatis* (MspA) y una de sus variantes. El nanoporo de alfa hemolisina y las variantes del nanoporo natural que son particularmente útiles se describen, por ejemplo, en la solicitud de patente de EE.UU. Pub. nº 2011/0229877 A1, o las patentes de EE. UU. nº US 6,916,665; nº 7,867,716; nº 7,947,454 o nº 8.105.846. MspA y las variantes del nanoporo natural que son especialmente útiles se describen, por ejemplo, en la solicitud de patente de EE.UU. Pub. nº 2012/0055792 A1. Los nanoporos en estado sólido también pueden ser útiles, incluidos, por ejemplo, los descritos en las patentes de EE.UU. nº 6,413,792; nº 7,444,053; o nº 7,582,490.

La detección de trifosfatos de nucleótidos puede aprovechar la interacción con un nanoporo que da como resultado cambios en la corriente que circula a través del nanoporo de una manera que es específica de cada especie de trifosfato de nucleótido. Por ejemplo, una primera especie de trifosfato de nucleótido puede reducir la corriente que circula a través del nanoporo durante un período medio determinado o en un grado determinado. Una segunda especie de trifosfato de nucleótido puede distinguirse en virtud de un período medio diferente o un grado diferente de alteración de la corriente cuando pasa a través del nanoporo. Por lo tanto, se pueden distinguir diferentes especies de trifosfato de nucleótido basadas en alteraciones distintivas de la corriente que circula a través de un nanoporo.

La detección de nanoporos se puede llevar a cabo usando cualquiera de una variedad de aparatos conocidos en la técnica incluidos, por ejemplo, los descritos en las solicitudes de patente de EE.UU. nº 2011/0229877 A1 o nº 2012/0055792 A1; o las patentes de EE.UU. nº US 6,413,792; nº 6,916,665; nº 7,867,716; nº 7,444,053; nº 7,582,490; nº 7,947,454 o nº 8,105,846.

Una polimerasa que se utiliza en un aparato o método expuesto en la presente memoria puede estar presente en solución de manera que sea relativamente libre de difundirse, al menos dentro de una cámara de reacción o puede estar relativamente limitada en su capacidad de difundirse por estar unida a un soporte en fase sólida, nanoporo, barrera u otro componente de un método o aparato expuesto en la presente memoria. Limitar la difusión por unión puede proporcionar una ventaja de acoplar estrechamente el punto de producción de trifosfato de nucleótidos (p. ej., una polimerasa que cataliza pirofosforólisis) con el punto de detección de trifosfato de nucleótido (p. ej., un nanoporo a través del cual pasan los trifosfatos de nucleótidos). Una polimerasa se puede unir a un nanoporo, por ejemplo, a través de una fusión de proteína recombinada a una subunidad de un nanoporo, reticulación química o resto adaptador. Se exponen métodos útiles para unir polimerasas a nanoporos y componentes de polimerasa-nanoporo, por ejemplo, en las Pub. de solicitud de patente de EE.UU. nº 2011/0229877 A1; 2012/0055792 A1 o la patente de EE.UU. nº 7,947,454.

Una polimerasa se puede unir a una perla u otro soporte sólido que reside en una cámara donde se produce

pirofosforólisis. Los enlazadores químicos que son útiles para unir polimerasas a perlas o soportes sólidos incluyen los que están disponibles en el mercado, por ejemplo, de Thermo Fisher (Rockford, IL), Sigma Aldrich (St. Louis, MO) o de lo contrario conocidos en la técnica.

5 Una polimerasa se puede unir a una barrera utilizada en un aparato de secuenciación de nanoporos. Por ejemplo, en realizaciones que usan una bicapa lipídica como barrera, se puede unir un resto lipófilo a la polimerasa para localizar la polimerasa en la proximidad de la bicapa debido a las interacciones entre la bicapa y el resto lipófilo. Los restos lipófilos ejemplares incluyen, pero no se limitan a, esteroides o lípidos. Un ejemplo adicional de un resto lipófilo es una proteína de membrana (o una porción de la misma) que se puede unir a una polimerasa, por ejemplo, por fusión de proteína recombinada. Los enlaces tales como los expuestos anteriormente para perlas y otros soportes sólidos pueden usarse para unir una polimerasa a una barrera usada en sistemas de nanoporos en estado sólido.

10 Un ácido nucleico que se secuencía en un método expuesto en la presente memoria o presente en un aparato de la presente descripción puede estar en solución de modo que sea relativamente libre de difundirse o puede estar relativamente limitado en su capacidad para difundirse por estar unido a un soporte en fase sólida, nanoporo, barrera u otro componente de un método o aparato expuesto en la presente memoria. Las uniones similares a las expuestas anteriormente para polimerasas pueden usarse para ácidos nucleicos. Por ejemplo, un esteroide, lípido u otro resto lipófilo puede unirse a un ácido nucleico para localizarlo en una bicapa lipídica. Un ejemplo se muestra en la FIG. 2, donde el ácido nucleico se localiza en la membrana mediante un resto de esteroide unido a la cadena plantilla. Como se ejemplifica en la figura, el ácido nucleico se puede unir mediante la cadena plantilla, por ejemplo, en el extremo 5' de la cadena plantilla. La unión también puede hacerse en un punto en la plantilla que está entre la posición donde la polimerasa se une a la plantilla y al extremo 5' de la plantilla.

15 Un resto lipófilo se puede unir a un ácido nucleico usando métodos conocidos en la técnica para unir otros restos tales como biotina o fluoróforos. Por ejemplo, puede usarse un cebador que tenga el resto lipófilo en una amplificación, extensión del cebador o reacción de ligadura. Alternativamente o además, los nucleótidos que tienen el resto se pueden utilizar en una reacción de extensión o amplificación. Si se desea, puede introducirse un resto lipófilo antes de la secuenciación y durante una etapa de preparación de muestra, tal como las expuestas anteriormente en la presente memoria. Por ejemplo, puede emplearse una técnica de secuenciación dirigida en donde un subconjunto de ácido nucleico diana que tiene las secuencias deseadas debe seleccionarse de un fondo de secuencias más complejo. En este ejemplo, un resto lipófilo puede introducirse selectivamente en el subconjunto de ácidos nucleicos diana usando la técnica de focalización y esto puede permitir que las dianas sean capturadas selectivamente por una bicapa lipídica, mientras que otras secuencias no focalizadas se eliminan por lavado debido a que no se han modificado para incluir el resto.

Puede ser beneficioso en algunas realizaciones limitar la difusión tanto de la polimerasa como del ácido nucleico con respecto a un nanoporo, por ejemplo, usando uno o más de los medios de unión expuestos anteriormente.

20 Como se expone anteriormente en la presente memoria, un método de la presente descripción puede incluir una etapa de puesta en contacto de un ácido nucleico diana con una polimerasa en condiciones para eliminar consecutivamente nucleótidos, produciendo de esta forma consecutivamente trifosfatos de nucleótidos con una variedad de diferentes restos de bases. La variedad de restos de bases diferentes producidos dependerá del contenido del ácido nucleico diana que se pone en contacto con la polimerasa. Por ejemplo, el ADN incluye normalmente las cuatro bases corrientes guanina, citosina, adenina y timina de modo que la pirofosforólisis producirá trifosfato de desoxiguanidina, trifosfato de desoxicitidina, trifosfato de desoxiadenosina y trifosfato de desoxitimidina. En algunos casos, el ácido nucleico diana puede no incluir los cuatro tipos de bases, de modo que la pirofosforólisis no producirá más de 3, 2 o incluso 1 tipo de trifosfato de desoxinucleótido. En algunos casos, las variantes de uno o más de estos cuatro tipos de bases pueden estar presentes en el ADN diana y, por consiguiente, la pirofosforólisis puede producir trifosfatos de desoxinucleótidos variantes que tienen, por ejemplo, metilo, hidroximetilo u otros restos añadidos. Otras bases variantes conocidas en la técnica, tales como las expuestas en la presente memoria, también pueden estar presentes en los trifosfatos de desoxinucleótidos producidos por pirofosforólisis.

25 Otro ejemplo es el ARN, que normalmente incluye las cuatro bases corrientes guanina, citosina, adenina y uracilo de manera que la pirofosforólisis producirá trifosfato de riboguanidina, trifosfato de ribocitidina, trifosfato de riboadenosina y trifosfato de ribouracilo. En algunos casos, el ácido nucleico diana puede no incluir los cuatro tipos de bases de manera que no se produzca más de 3, 2 o incluso 1 tipo de trifosfato de ribonucleótido por pirofosforólisis. Las bases variantes, tales como las ejemplificadas en la presente memoria, por ejemplo, con respecto a trifosfatos de desoxinucleótido, o conocidas de otro modo en la técnica, pueden estar presentes en los trifosfatos de ribonucleótidos. Generalmente, los trifosfatos de nucleótidos producidos por pirofosforólisis (ya sean trifosfatos de desoxinucleótidos o trifosfatos de ribonucleótidos) incluirán uno, al menos dos, al menos tres, al menos cuatro o más tipos de bases diferentes.

30 Un método de la presente descripción se puede llevar a cabo en condiciones que eliminan sucesivamente un número de nucleótidos de un ácido nucleico diana, produciendo de ese modo sucesivamente el mismo número de trifosfatos de nucleótidos. Además, se puede distinguir al menos el mismo número de trifosfatos de nucleótidos, por ejemplo, por paso a través de un nanoporo, para permitir la determinación de una secuencia que tiene una longitud

que es al menos equivalente al número de nucleótidos eliminados del ácido nucleico diana. En determinadas realizaciones, el número es al menos 1, 2, 3, 4, 5, 10, 25, 50, 100, 200, 500, 1.000, 10.000 o más hasta e incluida la longitud del ácido nucleico diana. Alternativamente o además, el número puede ser no más de 1, 2, 3, 4, 5, 10, 25, 50, 100, 200, 500, 1.000 o 10.000. El número puede ser, pero no necesariamente, entre dos cualquiera de estos valores. Como se expuso anteriormente en la presente memoria, se puede utilizar una variedad de técnicas para interrumpir la pirofosforólisis. Esto puede proporcionar el control de la longitud de la secuencia determinada utilizando realizaciones de los presentes métodos.

El número de trifosfatos de nucleótidos liberados por pirofosforólisis y/o detectados en un método expuesto en la presente memoria puede ser mayor que el número de tipos diferentes de trifosfatos de nucleótidos detectados. Sin embargo, el orden y el número de los diferentes trifosfatos de nucleótidos detectados se pueden correlacionar con la secuencia del ácido nucleico.

En algunas realizaciones, puede ser beneficioso secuenciar repetidamente un determinado ácido nucleico diana. La repetición puede conseguirse, por ejemplo, procesando repetidamente una molécula de ácido nucleico diana en un método expuesto en la presente memoria. Por ejemplo, un método puede incluir las etapas siguientes: (a) poner en contacto un ácido nucleico diana con una polimerasa para eliminar sucesivamente los trifosfatos de nucleótidos del ácido nucleico diana, en donde los trifosfatos de nucleótidos que se eliminan tienen una variedad de diferentes restos de bases; (b) distinguir los diferentes restos de bases para los trifosfatos de nucleótidos que se eliminan, determinando de este modo la secuencia del ácido nucleico diana; (c) regenerar al menos una porción del ácido nucleico diana; y repitiendo las etapas (a) y (b) usar el ácido nucleico diana regenerado. El ácido nucleico diana puede regenerarse, por ejemplo, añadiendo trifosfatos de nucleótidos en condiciones para que la polimerasa (o una polimerasa recién añadida) lleve a cabo una reacción de polimerización para regenerar al menos una porción de una cadena del ácido nucleico diana que se eliminó previamente por pirofosforólisis. Generalmente, el pirofosfato estará sustancialmente ausente durante la reacción de polimerización.

Alternativamente o además de la realización de tratamiento repetido anterior, un ácido nucleico diana puede amplificarse o copiarse para crear copias múltiples que se procesan usando un método de la presente descripción. Un ejemplo diagramático se muestra en la FIG. 3. Varias copias de un ácido nucleico plantilla bicatenario se localizan en una barrera en una cámara que tiene una fusión de nanoporo-polimerasa (etapa 1), la polimerasa captura una de las cadenas (etapa 2), se produce la secuenciación en pirofosforólisis (etapa 3), la cadena plantilla, que es monocatenaria, puede ser impulsada a través del nanoporo mediante la fuerza eléctrica (etapa 4) hasta que se elimine de la cámara donde permanecen las otras copias del ácido nucleico (n.b. las otras copias permanecen debido a que son bicatenarias y por lo tanto resistentes al paso a través del nanoporo) (etapa 5), y luego se captura otra copia del ácido nucleico plantilla bicatenario para iniciar la repetición de las etapas 2 y siguientes. Cualquiera de una variedad de métodos conocidos en la técnica para amplificar ácidos nucleicos, tales como los expuestos anteriormente en la presente memoria, pueden utilizarse para crear las copias múltiples del ácido nucleico diana.

Aunque el sistema de la FIG. 3 se ejemplifica para las copias de una única plantilla, se entenderá que las especies de ácido nucleico que tienen secuencias diferentes se pueden usar de manera similar. Por lo tanto, una variedad de diferentes especies de ácido nucleico bicatenario puede localizarse en una barrera en una cámara que tiene una fusión de nanoporo-polimerasa (paso 1), una cadena de una primera especie puede ser capturada por la polimerasa (paso 2), puede producirse secuenciación basada en pirofosforólisis (paso 3), la cadena plantilla de la primera cadena se puede tirar a través del nanoporo mediante fuerza eléctrica (paso 4) hasta que se elimine de la cámara donde permanecen las otras especies de ácido nucleico (paso 5) , y luego se puede capturar otra especie de ácido nucleico bicatenario para iniciar la repetición de las etapas 2 y siguientes.

La presente descripción también proporciona un aparato que incluye (a) una barrera impermeable a los fluidos que separa un primer depósito de fluido de un segundo depósito de fluido; (b) un nanoporo situado en la barrera impermeable a los fluidos para formar un paso a través del cual puede pasar un trifosfato de nucleótido desde el primer depósito de fluido al segundo depósito de fluido; y (c) una mezcla de reacción en el primer depósito de fluido, incluyendo la mezcla de reacción una polimerasa, ácido nucleico diana bicatenario y concentración pirofosforolítica de pirofosfato. Los componentes usados en el aparato pueden ser uno o más de los ejemplificados anteriormente en el contexto de varios métodos. Se ejemplifican a continuación más componentes y configuraciones con fines de ilustración.

Se puede configurar una barrera impermeable a los fluidos para separar dos depósitos y tener un nanoporo colocado en la barrera para proporcionar una conexión fluida entre los depósitos. Nanoporos y barreras a modo de ejemplo se expusieron anteriormente y en diversas referencias expuestas anteriormente. En general, los dos depósitos estarán en comunicación fluida a través de un solo nanoporo. Por lo tanto, los trifosfatos de nucleótidos producidos en uno de los depósitos tendrán una y solo una vía de fluido hacia el segundo depósito. El uso de un solo nanoporo de esta manera permite la medición conveniente de cada trifosfato de nucleótido que pasa de un depósito a otro debido a cambios en las propiedades eléctricas en el nanoporo, barrera y/o depósitos. Sin embargo, también es posible en algunas realizaciones incluir más de un nanoporo en la barrera que separa dos depósitos. Cuando varios nanoporos conectan de forma fluida dos depósitos, el paso de trifosfatos de nucleótidos se puede medir en cada nanoporo usando, por ejemplo, mediciones ópticas o eléctricas que resuelven cada nanoporo.

Un depósito puede crear una cámara donde el fluido queda contenido durante al menos parte del tiempo. Por ejemplo, una cámara puede configurarse para formar un pozo, una cavidad, un compartimiento, etc. que restrinja la circulación de fluido. Alternativamente, puede configurarse un depósito para la circulación de fluido. Por ejemplo, el depósito puede configurarse como un tubo, canal o celda de flujo, permitiendo de este modo la circulación de fluidos para el suministro y la eliminación conveniente de los componentes utilizados en un método de secuenciación. En determinadas realizaciones, un primer depósito que contiene ácido nucleico plantilla, polimerasa y pirofosfato se configurará para la circulación de fluido, mientras que el segundo depósito, que está conectado a la primera cámara por un nanoporo, puede configurarse como una cámara. El segundo depósito no necesita configurarse para circulación de fluido, pero opcionalmente puede estar.

La presente descripción proporciona realizaciones múltiplex. Por ejemplo, las secuencias para un gran número de ácidos nucleicos diana se pueden determinar en paralelo. Un método múltiplex puede incluir las etapas siguientes: (a) proporcionar un gran número de ácidos nucleicos diana; (b) poner en contacto cada uno de los ácidos nucleicos diana con una polimerasa para eliminar sucesivamente los trifosfatos de nucleótidos de cada ácido nucleico diana, en donde los trifosfatos de nucleótidos que se eliminan tienen una variedad de restos de bases diferentes; y (c) distinguir los diferentes restos de bases para los trifosfatos de nucleótidos que se eliminan de cada ácido nucleico, determinando de este modo las secuencias de los ácidos nucleicos diana.

Un ejemplo más de método múltiplex es el que incluye las etapas siguientes: (a) proporcionar un gran número de ácidos nucleicos diana; (b) poner en contacto cada uno de los ácidos nucleicos diana con una polimerasa para eliminar sucesivamente los trifosfatos de nucleótidos de cada ácido nucleico diana, en donde los trifosfatos de nucleótidos que se eliminan tienen una variedad de restos de bases diferentes; y (c) distinguir los diferentes restos de bases para los trifosfatos de nucleótidos que se eliminan de cada ácido nucleico, determinando de este modo las secuencias de los ácidos nucleicos diana. Un método múltiplex puede incluir las etapas siguientes: (a) proporcionar un gran número de ácidos nucleicos diana; (b) poner en contacto cada uno de los ácidos nucleicos diana con una polimerasa para eliminar sucesivamente los trifosfatos de nucleótidos de cada ácido nucleico diana, en donde los trifosfatos de nucleótidos que se eliminan tienen una variedad de restos de bases diferentes; y (c) distinguir los diferentes restos de bases para los trifosfatos de nucleótidos que se eliminan de cada ácido nucleico, determinando de ese modo las secuencias de los ácidos nucleicos diana.

Un ejemplo adicional de un método múltiplex es uno que incluye las etapas siguientes: (a) proporcionar un gran número de ácidos nucleicos diana cada uno con dos cadenas; (b) poner en contacto cada uno de los ácidos nucleicos diana con una polimerasa en condiciones que eliminan sucesivamente nucleótidos de la primera de cada una de las dos cadenas por pirofosforólisis, produciendo de este modo sucesivamente trifosfatos de nucleótido que tienen una variedad de diferentes restos de bases; y (c) distinguir los diferentes restos de bases de los trifosfatos de nucleótido producidos sucesivamente, determinando de ese modo la secuencia de los ácidos nucleicos diana.

Un aparato múltiplex puede incluir (a) un gran número de barreras impermeables a los fluidos que separan cada uno un primer depósito de fluido de un segundo depósito de fluido; (b) un nanoporo situado en cada una de las barreras impermeables a los fluidos para formar un paso a través del cual puede pasar un trifosfato de nucleótido desde el primer depósito de fluido al segundo depósito de fluido; y (c) una mezcla de reacción en cada uno de los primeros depósitos de fluido, incluyendo cada una de las mezclas de reacción una polimerasa, ácido nucleico diana bicatenario, y concentración pirofosforolítica de pirofosfato.

La plexidad (es decir, nivel múltiplex) de un método o aparato puede seleccionarse para satisfacer un uso determinado. Por ejemplo, el número de ácidos nucleicos diana que se tratan o presentan juntos se puede determinar a partir de la complejidad de la muestra que se va a evaluar. Las estimaciones de complejidad a modo de ejemplo para algunos de los genomas que pueden evaluarse utilizando métodos o aparatos de la presente descripción son de aproximadamente 3,1 Gbp (humano), 2,7 Gbp (ratón), 2,8 Gbp (rata), 1,7 Gbp (pez cebra), 165 Mbp (mosca de la fruta), 13,5 Mbp (*S. cerevisiae*), 390 Mbp (pez globo), 278 Mbp (mosquito) o 103 Mbp (*C. elegans*). Los expertos en la técnica reconocerán que los genomas que tienen tamaños distintos a los ejemplificados anteriormente incluidos, por ejemplo, genomas más pequeños o más grandes, pueden usarse en un método de la invención. Generalmente, una muestra de ácido nucleico se fragmenta antes del uso. El número de fragmentos a manejar en paralelo dependerá de la complejidad del genoma, el tamaño medio del fragmento y la cobertura deseada. Por ejemplo, se puede conseguir una cobertura 1x de un genoma humano (3,1 Gbp) que está fragmentado a un tamaño medio de 1.000 nucleótidos utilizando la plexidad de 3 millones de fragmentos (es decir,  $((3,1 \text{ billones}/1.000) \times 1) = 1 \text{ millón}$ ). Utilizando cálculos similares, se puede determinar que una plexidad de 30 millones de fragmentos (de 1.000 nucleótidos cada uno) es suficiente para proporcionar una cobertura de 30x de un genoma humano.

Los métodos y aparatos expuestos en la presente memoria pueden configurarse a un nivel de plexidad suficiente para proporcionar al menos 1x, 2x, 5x, 10x, 20x, 30x, 50x o más cobertura de cualquiera de una variedad de genomas incluidos, pero no limitados a, los ejemplificados en la presente memoria. La plexidad puede ser función del número de diversos componentes expuestos en la presente memoria, tal como el número de fragmentos de ácido nucleico diana ejemplificados anteriormente. Otros componentes que se pueden multiplexar incluyen, por ejemplo, el número de nanoporos utilizados, el número de polimerasas, el número de cámaras que tienen una barrera y nanoporo, etc. El nivel múltiplex de estos u otros componentes puede ser, por ejemplo, al menos 2, 5, 10,

100,  $1 \times 10^3$ ,  $1 \times 10^4$ ,  $1 \times 10^5$ ,  $1 \times 10^6$ ,  $1 \times 10^9$  o más alto. Alternativamente o además, el nivel múltiplex puede seleccionarse para que no sea más de 2, 5, 10, 100,  $1 \times 10^3$ ,  $1 \times 10^4$ ,  $1 \times 10^5$ ,  $1 \times 10^6$  o  $1 \times 10^9$ .

La expresión "que comprende" pretende ser ilimitada, incluidos no solo los elementos citados, sino que abarca además cualquier elemento adicional.

**REIVINDICACIONES**

1. Un método para determinar la secuencia de un ácido nucleico diana, que comprende
  - (a) proporcionar un ácido nucleico diana;
  - 5 (b) poner en contacto el ácido nucleico diana con una polimerasa para eliminar sucesivamente trifosfatos de nucleótidos del ácido nucleico diana por pirofosforólisis, en donde los trifosfatos de nucleótido que se eliminan tienen una variedad de diferentes restos de bases; y
  - (c) distinguir los diferentes restos de bases para los trifosfatos de nucleótidos que se han eliminado, determinando de ese modo la secuencia del ácido nucleico diana; y en donde la polimerasa se une a un nanoporo.
- 10 2. El método de la reivindicación 1, en donde las condiciones para eliminar sucesivamente nucleótidos de una de las dos cadenas por pirofosforólisis comprenden poner en contacto la polimerasa con una concentración pirofosforolítica de pirofosfato.
- 15 3. El método de la reivindicación 2, en donde el método comprende además una etapa de detener la producción secuencial de los trifosfatos de nucleótidos eliminando el pirofosfato en contacto con la polimerasa y luego reanudar la producción secuencial de los trifosfatos de nucleótidos poniendo en contacto la polimerasa con pirofosfato.
4. El método de la reivindicación 1, en donde la variedad de diferentes restos de bases comprende al menos dos especies diferentes de restos de bases y como máximo cuatro especies diferentes de restos de bases.
5. El método de la reivindicación 1, en donde los restos de bases comprenden adenina, guanina, citosina o timina de origen natural.
- 20 6. El método de la reivindicación 1, en donde al menos uno de los restos de bases comprende un resto que no se produce de forma natural en ADN o ARN.
7. El aparato de la reivindicación 1, en donde el nanoporo comprende un nanoporo en estado sólido.

FIG. 1

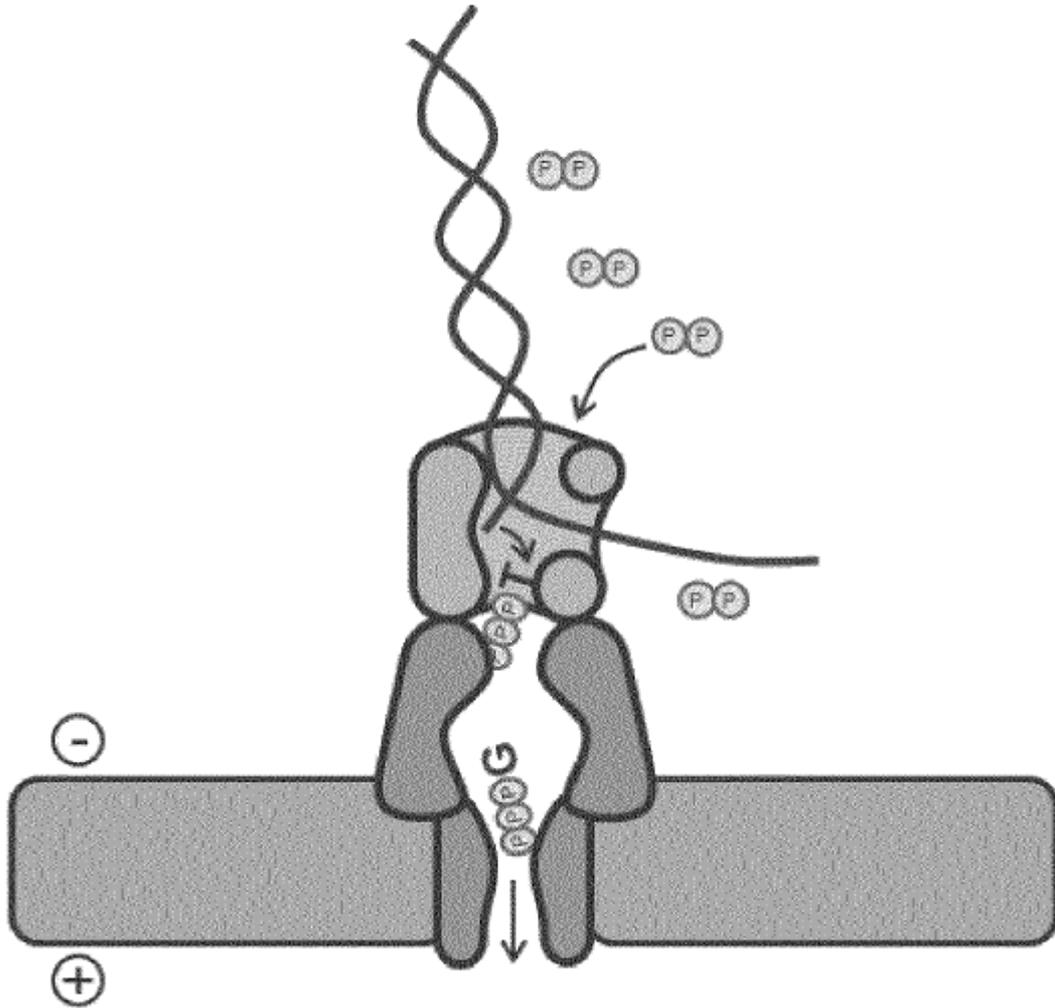


FIG. 2

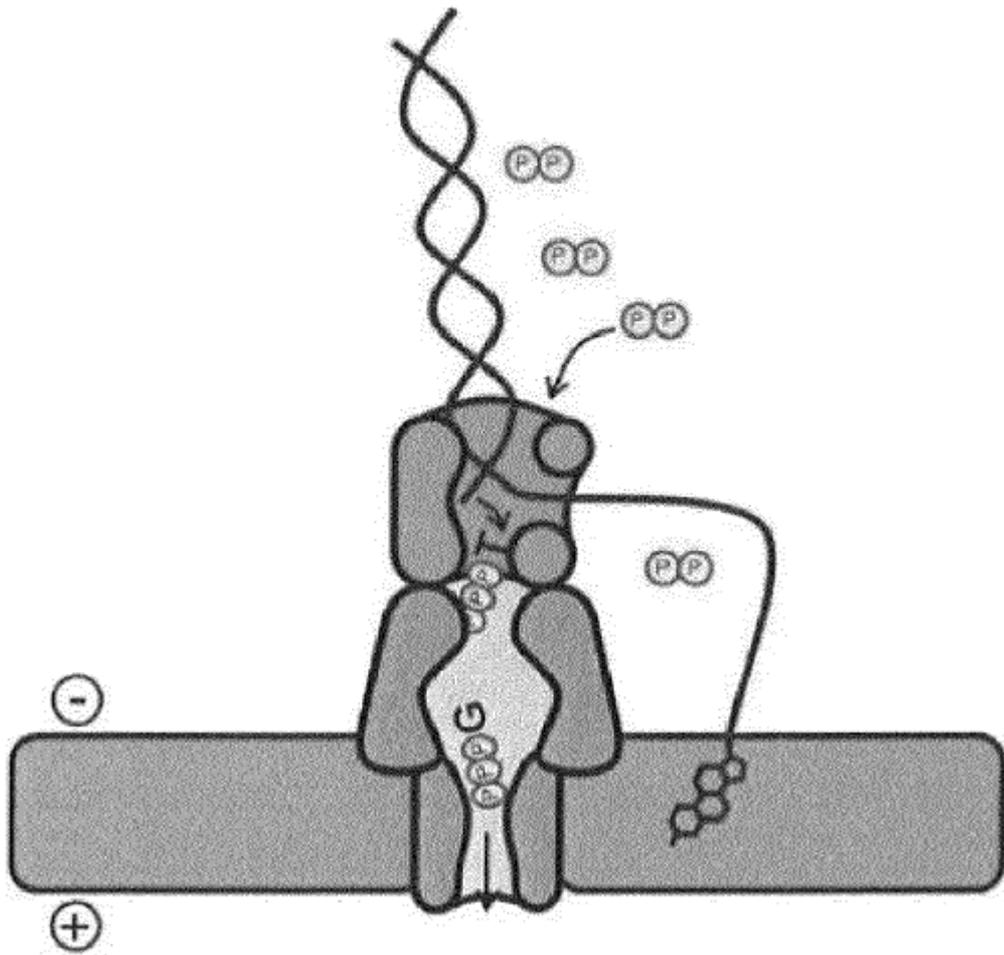


FIG. 3

