

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 994 071**

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 20183774 (7)**

97 Fecha y número de publicación de la concesión europea: **28.08.2024** **EP 3825415**

54 Título: **Secuenciación pirofosforolítica usando nanoporos**

30 Prioridad:

24.05.2013 US 201361827175 P

45 Fecha de publicación y mención en BOPI de la
traducción de la patente:
16.01.2025

73 Titular/es:

ILLUMINA CAMBRIDGE LIMITED (100.00%)
19 Granta ParkGreat Abington
Cambridge CB21 6DF, GB

72 Inventor/es:

MEULEMAN, WOUTER

74 Agente/Representante:

DEL VALLE VALIENTE, Sonia

ES 2 994 071 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

Secuenciación pirofosforolítica usando nanoporos

5 **Antecedentes**

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

10 Las plataformas disponibles comerciales actualmente para la secuenciación del ADN son relativamente costosas. Estas plataformas utilizan un enfoque 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 molde dirige estrictamente la síntesis de un nuevo polímero de ADN, se puede inferir la secuencia del ADN molde a partir de la serie de monómeros de nucleótidos que se añadieron a la cadena en crecimiento durante la síntesis. La capacidad de detectar adiciones de monómeros se ve facilitada por variantes especialmente diseñadas de los componentes bioquímicos que normalmente llevan a cabo la síntesis de ADN en los sistemas biológicos. Estos componentes diseñados son caros de fabricar y se consumen en cantidades relativamente grandes durante la secuenciación por síntesis. Además, la monitorización de la reacción utiliza equipos relativamente caros, tales como láseres, ópticas de detección y sistemas complejos de suministro de fluidos. Las plataformas comerciales más exitosas hasta la fecha también requieren reactivos y equipos costosos para amplificar los moldes de ADN antes de que la secuenciación por síntesis pueda siquiera comenzar.

Se han considerado otros métodos de secuenciación para reducir el coste, aumentar la productividad y/o simplificar el proceso. Uno de estos enfoques se basa en hacer pasar una sola cadena de ADN a través de un nanoporo e identificar su secuencia a partir de la variación en la corriente iónica que fluye a través del poro a medida que pasa la cadena. 25 Una alternativa a este enfoque de secuenciación “nanoporo-cadena” es la secuenciación “nanoporo-exonucleasa”, que implica la retirada catalizada por exonucleasas de monofosfatos de nucleótidos, de uno en uno, de una cadena de ADN y el paso secuencial de los monofosfatos de nucleótidos liberados a través de un nanoporo. Sin embargo, las variaciones resultantes en la corriente iónica que fluye 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. Sin embargo, a pesar de estos esfuerzos, la secuenciación nanoporo-exonucleasa aún no se ha demostrado a un nivel comercialmente viable hasta la fecha.

Por lo tanto, existe la necesidad de plataformas más rentables, rápidas y prácticas que proporcionen una alternativa a las disponibles actualmente para la secuenciación de ácidos nucleicos. La presente descripción aborda esta 35 necesidad y también proporciona otras ventajas.

Breve resumen

La presente invención es un método para determinar la secuencia de un ácido nucleico diana que comprende (a) proporcionar un ácido nucleico diana que tiene dos cadenas; (b) poner en contacto el ácido nucleico diana con una polimerasa unida a un nanoporo en condiciones para eliminar secuencialmente nucleótidos de la primera de las dos cadenas mediante pirofosforólisis y de esta manera producir secuencialmente trifosfatos de nucleótidos; y (c) distinguir los trifosfatos de nucleótidos producidos secuencialmente mediante la detección de cambios en la corriente que fluye a través del nanoporo.

45 La presente descripción también proporciona realizaciones no cubiertas por la invención, tales como 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 colocado en la barrera impermeable a los fluidos para formar un paso a través del cual un trifosfato de nucleótido puede pasar del primer depósito de fluido al segundo depósito de fluido; y (c) una mezcla de reacción en el primer depósito de fluido, estando comprendida en la mezcla de reacción una polimerasa, un ácido nucleico diana que tiene dos cadenas, y una concentración pirofosforolítica de pirofosfato.

Breve descripción de los dibujos

55 La FIGURA 1 muestra un diagrama de una reacción de secuenciación pirofosforolítica que utiliza una polimerasa unida a un nanoporo.

La FIGURA 2 muestra un diagrama de una reacción de secuenciación pirofosforolítica que utiliza una polimerasa unida a un nanoporo y un ácido nucleico molde unido a una barrera impermeable a los fluidos.

60 La FIGURA 3 que muestra la secuenciación pirofosforolítica con siembra en membrana de múltiples moldes de ácido nucleico.

Descripción detallada

65

La presente descripción proporciona un método para secuenciar ácidos nucleicos en modo inverso en comparación con las técnicas estándar de secuenciación por síntesis (SBS). El método de la presente descripción aprovecha una función catalítica de la polimerasa conocida como pirofosforólisis que típicamente se considera culpable de artefactos no deseados en las técnicas de SBS. La pirofosforólisis da como resultado la retirada de los trifosfatos de nucleótidos de una cadena de ácido nucleico por una polimerasa, y como tal es lo inverso de la reacción de polimerización empleadas en las técnicas estándar de SBS.

La pirofosforólisis se puede distinguir de la actividad de la exonucleasa (que está presente en algunas polimerasas), por ejemplo, según el diferente mecanismo catalítico de las dos reacciones, los diferentes sitios activos en la estructura de la polimerasa donde se producen las dos reacciones, y los diferentes productos de las reacciones. Respecto al mecanismo catalítico y las diferencias en el sitio activo, se tiene conocimiento de que la actividad de la exonucleasa se puede eliminar de muchas especies de polimerasas mediante la delección de ciertos dominios, mientras que se cree que la actividad de la pirofosforólisis está catalizada por el mismo dominio que cataliza la polimerización. Además, el ciclo de reacción catalizado por la exonucleasa es la conversión de ADN_n (ADN de longitud n) en ADN_{n-1} (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_{n-1} y un trifosfato de nucleótido a partir de ADN_n y pirofosfato. Notablemente, el pirofosfato se consume en una reacción de pirofosforólisis, pero no se consume en una reacción de exonucleasa.

Las realizaciones particulares de los métodos de secuenciación pirofosforolítica utilizan la detección de nanoporos. Los nanoporos se pueden utilizar para detectar secuencialmente los trifosfatos de nucleótidos que se liberan de un ácido nucleico para determinar la secuencia del ácido nucleico. Dichas realizaciones proporcionan una combinación de ventajas que típicamente solo se satisfacen parcialmente mediante la secuenciación de nanoporo-exonucleasa o la secuenciación nanoporo-cadena. Específicamente, los métodos de secuenciación pirofosforolítica de la presente descripción abordan algunos de los desafíos incurridos en la secuenciación nanoporo-exonucleasa, especialmente bajas probabilidades de captura y detección, al tiempo que conservan su principal ventaja sobre la secuenciación de cadenas, especialmente la resolución de una sola base. Esta ventaja se deriva del hecho de que la afinidad de los nanoporos por los monofosfatos de nucleótidos es bastante débil (del orden de afinidad micromolar), incluso en presencia de un adaptador de am6-ciclodextrina que se ha utilizado para mejorar la señal en algunos sistemas de nanoporos. Véase Clarke et al., *Nat. Nanotechnol.* abril;4(4):265-70 (2009). Para una distinción acertada de diferentes tipos de nucleótidos en un contexto de secuenciación, se desea una afinidad de rango nanomolar. El ATP tiene una afinidad que se encuentra en un buen rango, incluso sin la utilización de un adaptador en el poro. Véase Cheley et al., *Chem. Biol.* 9829-838 (2002). Sin pretender imponer ninguna teoría, se postula que la mejora de la afinidad del ATP sobre el monofosfato de nucleótido se debe a la triple carga negativa que transporta el ATP, lo que puede provocar que se una más fuertemente dentro del nanoporo. Además, la triple carga negativa puede ayudar a capturar la molécula en presencia de un campo eléctrico, especialmente cuando el campo es muy débil, como es el caso en el exterior del nanoporo, donde se liberan realmente los nucleótidos.

Además de la captura y detección mejoradas de los trifosfatos de nucleótidos, hay otras numerosas ventajas que pueden proporcionar las realizaciones expuestas en la presente memoria, tal como la utilización de una polimerasa en lugar de una exonucleasa. Se ha demostrado que las polimerasas forman un buen “ajuste” con los nanoporos con el fin de la secuenciación nanoporo-cadena (Cherf et al., *Nat. Biotech.* 30:344-348 (2012); Manrao et al., *Nat. Biotech.* 30:349-353 (2012). Aún no se ha demostrado un ajuste similarmente bueno con exonucleasas. Además, el sustrato de las polimerasas es ADN bicatenario que generalmente no entra en el nanoporo. Por el contrario, el ADN monocatenario, el sustrato de la mayoría de exonucleasas, puede plantear el problema de entrar y bloquear el nanoporo. Finalmente, a diferencia de la secuenciación basada en exonucleasa o la secuenciación de cadenas basada en polimerasa, la velocidad de una reacción de secuenciación pirofosforolítica se puede controlar ajustando la concentración de pirofosfato.

Debe entenderse que los términos utilizados en la presente memoria adoptarán su significado habitual a menos que se especifique lo contrario. A continuación, se exponen ejemplos de varios términos utilizados en la presente memoria y sus definiciones.

Como se usa en la presente descripción, el término “unido/a” pretende significar conectado por fuerzas que impiden la separación por difusión. El término puede incluir conexiones que son de naturaleza covalente o no covalente. Por ejemplo, dos proteínas pueden unirse covalentemente a través de su secuencia primaria (p. ej., un enlace peptídico o una fusión de proteínas) o entre sus secuencias primarias (p. ej., un enlace disulfuro o un enlace cruzado químico a través de grupos R de aminoácidos). Dos proteínas pueden unirse no covalentemente, por ejemplo, mediante uno o más enlaces de hidrógeno, enlaces iónicos, fuerzas de Van der Waals, enlaces hidrófobos o similares.

Como se utiliza en la presente memoria, el término “cada uno”, cuando se utiliza en referencia a una colección de artículos, pretende identificar un artículo individual en la colección, pero no se refiere necesariamente a cada artículo de la colección. Puede haber excepciones si la descripción explícita o el contexto inequívocamente dictan lo contrario.

Como se utiliza en la presente memoria, el término “actividad de la exonucleasa” pretende significar la 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 producirse en el enlace que une el nucleótido 5' al ácido nucleico (es decir, la actividad de la exonucleasa 5' a 3') o en el enlace que une el nucleótido 3' al ácido nucleico (es decir, la actividad de la exonucleasa 3' a 5').

Como se utiliza en la presente memoria, el término “barrera impermeable a los fluidos” pretende significar cualquier elemento 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 contenido en el primer depósito esté separado del fluido contenido 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 cualquier otra manera estaría obstruido por el resto de la barrera. En realizaciones particulares, el fluido puede ser un fluido acuoso y la barrera puede ser impermeable a los fluidos acuosos.

Como se utiliza en la presente memoria, el término “carece de actividad de exonucleasa” pretende significar que no tiene actividad medible de exonucleasa. Por ejemplo, una polimerasa u otro agente que carece de actividad de exonucleasa 3' a 5' no tendrá actividad medible de exonucleasa 3' a 5'. Similarmente, una polimerasa u otro agente que carece de actividad de exonucleasa 5' a 3' no tendrá actividad medible de exonucleasa 5' a 3'. Las polimerasas conocidas en la técnica como “exo menos” (o “exo-”), ya sean naturales o diseñadas, son ejemplos no limitativos de polimerasas que carecen de actividad de exonucleasa. Las variantes conocidas incluyen aquellas que son 5' a 3' exo menos y/o 3' a 5' exo menos.

Como se utiliza en la presente memoria, el término “nanoporo” pretende significar un pequeño orificio que permite el paso de los trifosfatos de nucleótidos a través de una barrera que de cualquier otra manera sería impermeable. La barrera es típicamente una capa eléctricamente aislante y el nanoporo típicamente permite que los iones fluyan de un lado de la barrera al otro, impulsados por un potencial aplicado. El nanoporo permite preferentemente que los nucleótidos fluyan a través de la luz del nanoporo a lo largo del potencial aplicado. El nanoporo también puede permitir que un ácido nucleico, tal como el ADN o el ARN, sea empujado o arrastrado a través de la luz del nanoporo. Sin embargo, en realizaciones particulares, el nanoporo no necesita permitir el paso de un ácido nucleico bicatenario o monocatenario. Un nanoporo utilizado en una realización particular 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 un “nanoporo de proteína”, que es un polipéptido o una colección de polipéptidos que forma el pequeño orificio que permite el paso de los trifosfatos de nucleótidos a través de una barrera tal como una bicapa lipídica. Entre los ejemplos de nanoporos de proteína se incluyen los nanoporos de alfa hemolisina, porina A de *Mycobacterium smegmatis* (MspA) y variantes de los mismos. Otro tipo de nanoporo es un “nanoporo en 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 utiliza en la presente memoria, el término “nucleótido” pretende incluir los ribonucleótidos, desoxinucleótidos o análogos de los mismos. Por ejemplo, el término se usa en la presente memoria para referirse generalmente a un resto de nucleósido (ya sea ribosa, desoxirribosa o un análogo de las mismas) que incluye un resto de base y opcionalmente unido a uno o más restos de fosfato. Los nucleótidos ilustrativos incluyen, aunque no de forma limitativa, 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 a la hora de distinguir los componentes de ARN de los componentes de ADN, el término “ribonucleótido” puede utilizarse 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 utilizarse para especificar nucleótidos de ADN, tales como trifosfato de desoxitimidina, trifosfato de desoxiguanidina, trifosfato de desoxicitidina y trifosfato de desoxiadenosina. En realizaciones particulares, los nucleótidos son “extensibles”, por ejemplo, carecen de un resto que bloquee la extensión en el hidroxilo 3' o en cualquier otra posición del nucleótido.

Como se utiliza en la presente memoria, el término “pirofosforólisis” pretende significar la reacción entre el pirofosfato y la unidad 3'-nucleotido de un ácido nucleico para liberar el nucleótido en forma del trifosfato de nucleótido correspondiente. 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 estándar.

Como se utiliza en la presente memoria, el término “concentración pirofosforolítica”, cuando se usa en referencia al pirofosfato, pretende significar una concentración de pirofosfato que provoca que se produzca una reacción de pirofosforólisis a un nivel sustancial. Por ejemplo, una concentración pirofosforolítica de pirofosfato puede dar como resultado que una polimerasa muestre una mayor velocidad de pirofosforólisis que de polimerización. Por lo tanto, una concentración pirofosforolítica de pirofosfato puede dar como resultado una inversión sustancial de la actividad de polimerización que de cualquier otra manera sería catalizada por una polimerasa.

Como se utiliza en la presente memoria, el término “depósito” pretende significar un receptáculo o cámara para contener o restringir el flujo de fluido. Un depósito puede estar totalmente cerrado, al menos temporalmente. Alternativamente, un depósito puede estar parcialmente cerrado, por ejemplo, que tenga una o más aberturas (p. ej., una o más entradas o salidas). Un fluido puede fluir a través de un depósito, por ejemplo, en realizaciones en los que el depósito está en una célula de flujo.

Como se utiliza en la presente memoria, el término “especie” se usa para identificar moléculas que comparten 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 entre sí. Similarmente, las moléculas de ADN individuales que tienen la misma secuencia de nucleótidos son de la misma especie.

Las realizaciones expuestas a continuación y mencionadas en las reivindicaciones pueden entenderse en vista de las definiciones anteriores.

La presente descripción proporciona un método para determinar la secuencia de un ácido nucleico diana tal como se define en las reivindicaciones.

Se muestra una realización ilustrativa en la **FIGURA 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 trifosfato de desoxiguanidina seguido de 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. Por tanto, los trifosfatos de desoxinucleótidos entran en el nanoporo secuencialmente, en el mismo orden en que fueron liberados de la cadena de ADN por la acción pirofosforolítica de la polimerasa. Los trifosfatos de desoxinucleótidos tienen una carga neta negativa debido a los trifosfatos y son conducidos a través del nanoporo por un potencial a través de la membrana (como indican el signo positivo en el lado de la membrana donde se produce la pirofosforólisis y un signo negativo en el lado opuesto de la membrana). De forma típica, los trifosfatos de nucleótidos reactivos no están presentes en una reacción de pirofosforólisis. En algunos casos, pueden estar presentes 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.

Una reacción similar se ilustra en la **FIGURA 2** salvo que la cadena molde (es decir, la cadena que no se está escindiendo por pirofosforólisis) está unida a la membrana. En este caso, la cadena molde tiene un resto de esterol unido covalentemente (p. ej., colesterol) que se une al interior hidrófobo de la bicapa lipídica de la membrana.

Un método de la presente descripción se puede utilizar con cualquiera de una variedad de ácidos nucleicos diana. El ácido nucleico diana puede tener una estructura natural como la que se encuentra, por ejemplo, en el ADN o el 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-metilcitosina, 5-hidroximetilcitosina y otras bases metiladas. Los trifosfatos de desoxinucleótidos que tienen estas bases variantes pueden producirse y/o detectarse con un método o aparato expuesto en la presente memoria. Por ejemplo, la presencia o ausencia de metilación en la citosina se puede distinguir para facilitar los análisis epigenéticos. El ARN contiene naturalmente 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.

Un ácido nucleico puede incluir modificaciones no naturales, tales como bases no nativas, modificaciones en los restos de fosfato o modificaciones en los restos de azúcares. Las bases no nativas ilustrativas que pueden incluirse en un ácido nucleico, tanto si tiene una cadena principal natural como una estructura análoga, incluyen, aunque no de forma limitativa, inosina, xantina, hipoxantina, isocitocina, isoguanina, 2-aminoadenina, 6-metil adenina, 6-metil guanina, 2-propil guanina, 2-propil adenina, 2-tiouracilo, 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-azaguanina, 8-azaadenina, 7-deazaguanina, 7-deazaadenina, 3-deazaguanina, 3-deazaadenina o similares.

Las bases no nativas se pueden seleccionar, por ejemplo, para transmitir un tamaño mayor o menor, o para transmitir una carga aumentada o disminuida, a fin de influir en la capacidad de los análogos de trifosfato de nucleótido resultantes para ser distinguidos por un nanoporo u otro componente de detección. Similarmente, tales bases pueden ser beneficiosas si se seleccionan para transmitir una velocidad deseada de pirofosforólisis. Las bases no nativas se pueden incorporar a un ácido nucleico mediante métodos conocidos tales como la amplificación o replicación de un ácido nucleico molde en presencia de los análogos de nucleótidos. Una o más de las copias de ácido nucleico resultantes se pueden utilizar a continuación como ácido(s) nucleico(s) diana en el aparato o método de secuenciación expuesto en la presente memoria.

En realizaciones particulares, un ácido nucleico que se utiliza en un método o aparato expuesto en la presente memoria carecerá de una o más de las bases no nativas u otros restos no nativos expuestos en la presente memoria. Por ejemplo, en realizaciones particulares, los métodos no se utilizan para eliminar un nucleótido terminador del extremo 3' de una cadena de ácido nucleico. En consecuencia, en algunas realizaciones, un aparato o método puede no incluir ningún ácido nucleico que tenga un nucleótido terminador en su extremo 3'. Alternativamente o de forma adicional, en algunas realizaciones, un aparato o método puede no incluir ningún nucleótido terminador.

Como se ilustra en la **FIGURA 1** y en cualquier otro sitio de la presente memoria, un ácido nucleico diana puede ser ADN bicatenario, por ejemplo, cuando se usa ADN-polimerasa para la pirofosforólisis. También se puede utilizar un heterodúplex, formado entre una cadena de ADN y una cadena de ARN. Por ejemplo, se puede utilizar ARN-polimerasa para catalizar la pirofosforólisis en el extremo 3' de una cadena de ARN que se hibrida con una cadena molde de ADN, produciendo de esta manera trifosfatos de ribonucleótidos.

Un ácido nucleico que se usa en un método o aparato en la presente memoria puede aislarse de una fuente biológica y utilizarse directamente o procesarse para producir un producto amplificado o modificado. Alternativamente, se puede usar un ácido nucleico sintético, de nuevo, directamente o después del procesamiento. El procesamiento puede incluir, por ejemplo, uno o más aislamientos a partir de componentes naturales, escisión para formar fragmentos, amplificación (p. ej., mediante PCR, círculo rodante u otras técnicas conocidas), ligadura de secuencias adaptadoras o secuencias de etiquetas, tagmentación mediante un transposón o métodos de "preparación de muestras" utilizados antes de las técnicas de secuenciación de ácidos nucleicos. En la técnica, se conocen técnicas de procesamiento útiles tal como se expone, por ejemplo, en Sambrook et al., *Molecular Cloning: A Laboratory Manual*, 3.^a 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).

Entre los ejemplos de métodos de preparación de muestras que se pueden utilizar para procesar ácidos nucleicos antes de la secuenciación basada en pirofosforólisis, se incluyen métodos que se han desarrollado para la amplificación del genoma completo o la amplificación del exoma completo junto con técnicas de secuenciación paralela masiva. Por ejemplo, se pueden usar 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. núm. 6.107.023 o 7.741.463; o la publicación de solicitud de patente de EE. UU. núm. 2010/0120098 A1. También se pueden usar métodos de preparación de muestras dirigidos para aislar un subconjunto del contenido de secuencia de una muestra compleja para su posterior secuenciación. Los ejemplos de métodos comerciales que pueden utilizarse para la captura dirigida de un subconjunto de fragmentos de ácido nucleico incluyen, aunque no de forma limitativa, los kits SureSelect™ (Agilent, Santa Clara, CA), los kits de enriquecimiento TruSeq® (Illumina, Inc., San Diego, CA) o los kits de enriquecimiento Nextera® (Illumina, Inc., San Diego, CA).

Los ácidos nucleicos pueden aislarse de cualquiera de una variedad de fuentes que incluyen, aunque no de forma limitativa, un mamífero tal como un roedor, ratón, rata, conejo, cobaya, ungulado, caballo, oveja, cerdo, cabra, vaca, gato, perro, primate, humano o primate no humano; una planta, tal como *Arabidopsis thaliana*, maíz (*Zea mays*), sorgo, avena (*Oryza sativa*), trigo, arroz, colza o soja; un 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, tales 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 obtenerse de genomas más pequeños tal como los de un organismo procariota tal como una bacteria, *Escherichia coli*, *staphylococci* 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.

Cualquiera de una variedad de polimerasas se puede utilizar en un método o aparato expuesto en la presente memoria; incluyen, por ejemplo, enzimas basadas en proteínas aisladas de sistemas biológicos y variantes funcionales de las mismas. Se entenderá que la referencia a una polimerasa particular, tales como las que se ilustran a continuación, incluye variantes funcionales de la misma, salvo que se indique lo contrario. Los ejemplos de polimerasas útiles incluyen las ADN-polimerasas y las ARN-polimerasas. Las ADN-polimerasas ilustrativas incluyen aquellas que se han clasificado por homología estructural en familias identificadas como A, B, C, D, X, Y y RT. Las ADN-polimerasas de la familia A incluyen, por ejemplo, ADN-polimerasa T7, ADN-polimerasa y mitocondrial eucariótica, ADN Pol I de *E. coli*, *Thermus aquaticus* Pol I, y *Bacillus stearothermophilus* Pol I. Las ADN-polimerasas de la familia B incluyen, por ejemplo, ADN-polimerasas eucarióticas α , δ , y ϵ ; ADN-polimerasa ζ ; ADN-polimerasa T4, ADN-polimerasa *Phi29*, y ADN-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 de la familia X incluyen, por ejemplo, las polimerasas eucarióticas Pol β , pol σ , Pol λ y Pol μ , y Pol4 de *S. cerevisiae*. Las ADN-polimerasas de la familia Y incluyen, por ejemplo, Pol η , 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 ilustrativas incluyen, aunque no de forma limitativa, ARN-polimerasas víricas tales como ARN-polimerasa T7; ARN-polimerasas eucarióticas tales como ARN-polimerasa I, ARN-polimerasa II, ARN-polimerasa III, ARN-polimerasa IV y ARN-polimerasa V; y ARN-polimerasa de Archaea.

Las clasificaciones anteriores se proporcionan con fines ilustrativos. Se entenderá que las variaciones en el sistema de clasificación son posibles. Por ejemplo, en al menos un sistema de clasificación, las polimerasas de la familia C se han categorizado como una subcategoría de la familia X. Además, las polimerasas se pueden clasificar de acuerdo con otras características, ya sean funcionales o estructurales, que se pueden solapar o no con las características estructurales ilustradas anteriormente. A continuación, se exponen algunas características ilustrativas con mayor detalle.

Muchas polimerasas tienen una actividad de exonucleasa de corrección 3' a 5' intrínseca que puede ser útil en algunas realizaciones. En algunas realizaciones, se prefieren las polimerasas que carecen sustancialmente de actividad de la exonucleasa de corrección 3' a 5', por ejemplo, en la mayoría de realizaciones de secuenciación. La ausencia de actividad de la exonucleasa puede ser una característica de tipo natural o una característica transmitida por una variante o polimerasa diseñada. Por ejemplo, el fragmento de Klenow exo menos es una versión mutada del fragmento de Klenow que carece de actividad de exonucleasa de corrección 3' a 5'. El fragmento de Klenow y su variante exo menos pueden ser útiles en un método o aparato expuesto en la presente memoria. Las polimerasas que catalizan la pirofosforólisis, la inversión directa de la polimerización en el mismo sitio activo, son particularmente útiles.

Dependiendo de la realización que se vaya a usar, una polimerasa puede ser termófila o inactivable por calor. Las polimerasas termófilas son típicamente útiles para condiciones de alta temperatura o en condiciones de termociclado tales como las empleadas para las técnicas de reacción en cadena de la polimerasa (PCR). Los ejemplos de polimerasas termófilas incluyen, aunque no de forma limitativa, ADN-polimerasa 9°N, ADN-polimerasa Taq, ADN-polimerasa Phusion®, ADN-polimerasa Pfu, ADN-polimerasa RB69, ADN-polimerasa KOD y ADN-polimerasa VentR®. La mayoría de 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. Los ejemplos de polimerasas inactivables por calor son las de los fagos. Se entenderá que las polimerasas de cualquiera de una variedad de fuentes se pueden modificar para aumentar o disminuir su tolerancia a condiciones de alta temperatura.

Se puede usar una variante diseñada de una polimerasa que tenga una actividad de pirofosforólisis aumentada. Las variantes ilustrativas son aquellas que se han creado para su uso en técnicas de PCR, incluidas, aunque no de forma limitativa, las variantes descritas en la patente de EE. UU. núm. 8.071.536.

Se puede inducir una polimerasa para eliminar secuencialmente los nucleótidos de una de las dos cadenas de ácido nucleico mediante pirofosforólisis en un método expuesto en la presente memoria. La polimerasa se puede colocar 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. Se puede utilizar una concentración de pirofosfato que provoque que se produzca una reacción de pirofosforólisis a un nivel sustancial, incluido, aunque no de forma limitativa, al menos aproximadamente 100 pM de pirofosfato. Se pueden emplear concentraciones más altas de pirofosfato, por ejemplo, para impulsar la pirofosforólisis a una velocidad mayor. En consecuencia, se puede usar una concentración de al menos aproximadamente 250 pM de pirofosfato, al menos aproximadamente 500 pM de pirofosfato, al menos aproximadamente 750 pM 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 la 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 se puede disminuir mediante una concentración menor de pirofosfato que las expuestas anteriormente. Por lo tanto, como 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 20 mM de pirofosfato, como máximo aproximadamente 10 mM de pirofosfato, como máximo aproximadamente 5 mM de pirofosfato, como máximo aproximadamente 2 mM de pirofosfato, como máximo aproximadamente 1 mM de pirofosfato, como máximo aproximadamente 750 µM de pirofosfato, como máximo aproximadamente 500 µM de pirofosfato, como máximo aproximadamente 250 µM de pirofosfato, como máximo aproximadamente 100 µM de pirofosfato, o menos.

Los trifosfatos de nucleótidos reactivos típicamente están 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 se puede considerar una concentración no catalítica. Una concentración no catalítica de trifosfato de nucleótido es una concentración que es insuficiente para permitir que se produzca una 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. En consecuencia, la pirofosforólisis se puede llevar 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 inducir la pirofosforólisis que se pueden utilizar en la presente memoria se exponen en Patel et al. Biochem. 30:511-525 (1991), o la patente de EE. UU. núm. 7.090.975; 7.914.995 o 8.071.536. En varios casos, estas referencias describen reacciones que también incluyen componentes utilizados para la extensión de un ácido nucleico. Se entenderá que la extensión no se utiliza en realizaciones particulares de la presente descripción. Uno o más de los componentes que se describen en Patel et al. Biochem. 30:511-525 (1991) o la patente de EE. UU. núm. 7.090.975; 7.914.995 o 8.071.536, incluidos, por ejemplo, los componentes utilizados para la reacción de extensión de la polimerasa, pueden excluirse de un método o aparato expuesto en la presente memoria.

Los tampones, sales, iones de metal, glicerol, DMSO y otros componentes típicamente incluidos en las mezclas de reacción o almacenamiento de la polimerasa se pueden incluir en un aparato o método de la presente descripción, según se desee. Las cantidades y proporciones de componentes que se van a incluir resultarán evidentes para los expertos en la técnica o se podrán determinar, por ejemplo, mediante técnicas de valoración volumétrica habituales.

Un aspecto beneficioso de algunas realizaciones es la capacidad de detener o pausar el método de secuenciación alterando las condiciones de reacción para inhibir la pirofosforólisis. El método de secuenciación puede reiniciarse después opcionalmente, alterando las condiciones de reacción para permitir que se reanude la pirofosforólisis. Cualquiera de una variedad de condiciones de reacción expuestas en la presente memoria, o en las referencias citadas en la presente memoria, puede alterarse entre la pirofosforólisis pausada y la pirofosforólisis reanudada, permitiéndose de esta manera un cambio efectivo entre la secuenciación pausada y la activa, respectivamente.

En realizaciones particulares, un método de la presente descripción puede incluir una etapa para pausar la producción secuencial de trifosfatos de nucleótidos al hacer que el pirofosfato deje de estar en contacto con una polimerasa que está catalizando la pirofosforólisis, seguida de una etapa para reanudar la producción secuencial de los trifosfatos de nucleótidos poniendo en contacto la polimerasa con pirofosfato. El pirofosfato se puede suministrar a una reacción mediante técnicas apropiadas para el sistema de fluido que se utiliza, incluido, por ejemplo, pipetear alícuotas de fluido, mover bolos de fluido bajo presión positiva o negativa (p. ej., mediante bombas o gravedad), electroforesis, isotacoforesis, manipulación de gotículas (p. ej., electrohumectación) o similares. Se pueden usar técnicas fluidicas similares para eliminar el pirofosfato, por ejemplo, mediante desplazamiento y/o reemplazo con una solución de lavado. Por supuesto, tales técnicas fluidicas se pueden utilizar para añadir o eliminar otros componentes utilizados en los métodos y aparatos expuestos en la presente memoria.

Alternativamente o de forma adicional a los métodos fluidicos expuestos anteriormente, el pirofosfato se puede eliminar de la reacción mediante secuestro, degradación o inactivación. Por ejemplo, se pueden usar manipulaciones físicas tales como la adsorción a un agente secuestrante o la degradación por calor, luz o electricidad. Se pueden utilizar métodos químicos para modificar la estructura o actividad del pirofosfato o para degradar la molécula. También se pueden utilizar métodos enzimáticos, tales como la degradación por pirofosfatasa, como se muestra para las reacciones de PCR en EE. UU. 4.800.159 y EE. UU. 5.498.523 y para las reacciones de secuenciación basadas en gel en EE. UU. 4.962.020.

Alternativamente o de forma adicional a las técnicas expuestas anteriormente, la pirofosforólisis se puede detener o pausar al eliminar otros componentes de la reacción. Por ejemplo, la polimerasa puede eliminarse de una reacción y opcionalmente reemplazarse o devolverse a un estado activo. Por ejemplo, la polimerasa se puede eliminar mediante métodos fluidicos, secuestro, degradación o inactivación, tales como los ilustrados anteriormente para el pirofosfato. En realizaciones particulares, puede utilizarse una polimerasa sensible al calor (no termófila) en una reacción de pirofosforólisis y, después, inactivarse por calor. Similarmente, una polimerasa se puede degradar mediante modificación química o degradación enzimática (p. ej., mediante una proteasa). Ya sea degradada mediante técnicas físicas, químicas o enzimáticas, la polimerasa agotada se puede eliminar por lavado y, a continuación, se puede reanudar la pirofosforólisis mediante la adición de polimerasa nueva al ácido nucleico que se está secuenciando. La actividad de la polimerasa también se puede alternar mediante la adición y retirada de inhibidores, alternando entre temperaturas permisivas y no permisivas para las polimerasas termoestables, o mediante la presencia y ausencia de un agente secuestrante o sustrato competitivo. La pirofosforólisis también se puede detener e iniciar mediante la desnaturalización y la renaturalización, respectivamente, del ácido nucleico que se está secuenciando.

Aunque en la presente memoria se han ilustrado métodos y aparatos para realizaciones que utilizan pirofosfato para impulsar la pirofosforólisis, se entenderá que se pueden utilizar análogos del pirofosfato en su lugar. Un análogo ilustrativo es el pirovanadato, que puede utilizarse, por ejemplo, como se describe en Akabayov et al. J. Biol. Chem. 286:29146-29157 (2011). Como ejemplos adicionales, se pueden usar análogos de pirofosfato que tengan restos adicionales. Generalmente, se seleccionan análogos de pirofosfato que no inhiban completamente la pirofosforólisis o el paso de los trifosfatos de nucleótidos resultantes, o análogos de los mismos, 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 utilizar un análogo de pirofosfato para ralentizar o acelerar la pirofosforólisis para proporcionar una velocidad de detección deseada u óptima. Similarmente, los análogos de los trifosfatos de nucleótidos que se producen cuando se utiliza un análogo de pirofosfato en una reacción de pirofosforólisis también pueden transmitir las características deseadas para la detección de nanoporos. Por ejemplo, los restos que alteran la carga o el tamaño, en comparación con el difosfato por sí solo, pueden aumentar o disminuir la velocidad de paso de los análogos de trifosfato de nucleótido a través de un nanoporo, o alterar de cualquier otra manera 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 el pirofosfato que tenga cualquier resto añadido. Por ejemplo, se puede utilizar pirofosfato que carezca de un resto ópticamente detectable, tal como un resto fluorescente.

En realizaciones particulares, los trifosfatos de nucleótidos se detectan mediante nanoporos. Por ejemplo, los trifosfatos de nucleótidos que se eliminan secuencialmente de un ácido nucleico mediante pirofosforólisis pueden hacerse pasar a través

de un nanoporo para su detección. Mediante el uso de un nanoporo apropiado, se pueden distinguir diferentes restos de bases de trifosfatos de nucleótidos para permitir la detección de secuencias. Generalmente se puede utilizar un aparato que incluye un primer y un segundo compartimento separados por una barrera física, donde la barrera tiene uno o más nanoporos. El primer compartimento puede incluir componentes utilizados para una reacción de pirofosforólisis. El aparato se puede configurar para aplicar un campo eléctrico a través de la barrera, de modo que los trifosfatos de nucleótidos sean conducidos desde el primer compartimento a través del poro hasta el 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 un nanoporo. Los métodos pueden llevarse a cabo mediante fijación en parche de membrana o fijación de voltaje.

Un método de la presente descripción puede llevarse a cabo mediante cualquier sistema adecuado en el que un poro penetre a través de una barrera. La barrera en muchas realizaciones es preferiblemente una bicapa lipídica. Las bicapas lipídicas se pueden fabricar con métodos conocidos en la técnica, por ejemplo, como se describe en Montal y Mueller Proc. Natl. Acad. Sci. USA 69:3561-3566 (1972) o WO 2008/102120. Las bicapas lipídicas se pueden formar a partir de cualquiera de una variedad de lípidos, incluidos, aunque no de forma limitativa, fosfolípidos, glicolípidos, colesterol y mezclas de estos.

Los nanoporos ilustrativos que se pueden utilizar incluyen, por ejemplo, nanoporos basados en proteínas tales como el nanoporo de alfa hemolisina, porina A de *Mycobacterium smegmatis* (MspA) y variantes de los mismos. El nanoporo de alfa hemolisina y las variantes del nanoporo natural que son particularmente útiles se describen, por ejemplo, en la publicación de solicitud de patente de EE. UU. núm. 2011/0229877 A1, o las patentes de EE. UU. núm. US 6.916.665; 7.867.716; 7.947.454; o 8.105.846. MspA y las variantes del nanoporo natural que son particularmente útiles se describen, por ejemplo, en la publicación de solicitud de patente de EE. UU. núm. 2012/0055792 A1. Los nanoporos en estado sólido también pueden ser útiles, incluidos, por ejemplo, los descritos en las patentes de EE. UU. núm. 6.413.792; 7.444.053; o 7.582.490.

La detección de trifosfatos de nucleótidos puede aprovechar la interacción con un nanoporo que resulta en cambios en la corriente que fluye a través del nanoporo de una manera específica para cada especie de trifosfato de nucleótido. Por ejemplo, una primera especie de trifosfato de nucleótido puede reducir la corriente que fluye a través del nanoporo durante un período de tiempo medio particular o en un grado particular. Una segunda especie de trifosfato de nucleótido se puede distinguir en virtud de un período de tiempo medio diferente o de un grado diferente de alteración de la corriente al pasar a través del nanoporo. Por lo tanto, se pueden distinguir diferentes especies de trifosfato de nucleótido en función de las alteraciones distintivas de la corriente que fluye a través de un nanoporo.

La detección de nanoporos se puede llevar a cabo mediante cualquiera de una variedad de aparatos conocidos en la técnica, incluidos, por ejemplo, los descritos en las publicaciones de solicitud de patente de EE. UU. núm. 2011/0229877 A1; o 2012/0055792 A1; o las patentes de EE. UU. núm. US 6.413.792; 6.916.665; 7.867.716; 7.444.053; 7.582.490; 7.947.454; o 8.105.846.

Una polimerasa que se usa 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 nucleótido (p. ej., una polimerasa que cataliza la 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, mediante una fusión de proteínas recombinantes con una subunidad de un nanoporo, enlaces cruzados químicos o resto adaptador. Se exponen métodos útiles para unir polimerasas a nanoporos y componentes de polimerasa-nanoporo, por ejemplo, en la publicación de solicitud de patente de EE. UU. núm. 2011/0229877 A1; o 2012/0055792 A1; o la patente de EE. UU. núm. 7.947.454.

Una polimerasa se puede unir a una perla u otro soporte sólido que reside en una cámara donde se produce la pirofosforólisis. Las porciones químicas que son útiles para unir polimerasas a perlas o soportes sólidos incluyen aquellas que están disponibles comercialmente, por ejemplo, de Thermo Fisher (Rockford, IL) o Sigma Aldrich (St. Louis, MO) o conocidas de otro modo en la técnica.

Una polimerasa se puede unir a una barrera usada en un aparato de secuenciación de nanoporos. Por ejemplo, en las 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 ilustrativos incluyen, aunque no de forma limitativa, 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, mediante la fusión de proteínas recombinantes. Se pueden usar enlaces tales como los expuestos anteriormente para perlas y otros soportes sólidos para unir una polimerasa a una barrera utilizada en sistemas de nanoporos en estado sólido.

Un ácido nucleico que se secuencia en un método expuesto en la presente memoria o presente en un aparato de la presente descripción puede estar en solución de manera que sea relativamente libre de difundirse o puede estar relativamente limitado en su capacidad de 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. Se pueden utilizar uniones similares a las expuestas anteriormente

para las polimerasas para los ácidos nucleicos. Por ejemplo, un estero, lípido u otro resto lipófilo puede unirse a un ácido nucleico para localizarlo en una bicapa lipídica. Se muestra un ejemplo en la **FIGURA 2**, donde el ácido nucleico se localiza en la membrana a través de un resto de estero unido a la cadena molde. Como se ilustra en la figura, el ácido nucleico se puede unir a través de la cadena molde, por ejemplo, en el extremo 5' de la cadena molde. La unión también se puede realizar en un punto del molde que esté entre la ubicación donde la polimerasa está unida al molde y el extremo 5' del molde.

Un resto lipófilo se puede unir a un ácido nucleico mediante métodos conocidos en la técnica para unir otros restos tales como biotina o fluoróforos. Por ejemplo, un cebador que tiene el resto lipófilo se puede usar en una reacción de amplificación, extensión del cebador o ligadura. Alternativamente o de forma adicional, los nucleótidos que tienen el resto se pueden utilizar en una reacción de extensión o amplificación. Si se desea, se puede introducir un resto lipófilo antes de la secuenciación y durante una etapa de preparación de la muestra, tal como las expuestas anteriormente en la presente memoria. Por ejemplo, se puede emplear una técnica de secuenciación dirigida donde un subconjunto de ácidos nucleicos diana que tienen las secuencias deseadas se debe seleccionar de entre un fondo de secuencia más complejo. En este ejemplo, se puede introducir selectivamente un resto lipófilo en el subconjunto de ácidos nucleicos diana mediante la técnica de localización específica, lo que puede permitir que las dianas se capturen selectivamente por una bicapa lipídica, mientras que otras secuencias no diana se eliminan por lavado debido a que no se han modificado para incluir el resto lipófilo.

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

Como se expuso anteriormente en la presente memoria, un método de la presente descripción puede incluir una etapa para poner en contacto un ácido nucleico diana con una polimerasa en condiciones para eliminar secuencialmente los nucleótidos, produciendo de esta manera secuencialmente trifosfatos de nucleótidos que tienen una variedad de diferentes restos de bases. La variedad de diferentes restos de bases producidas dependerá del contenido del ácido nucleico diana que se pone en contacto con la polimerasa. Por ejemplo, el ADN típicamente incluye las cuatro bases comunes guanina, citosina, adenina y timina, de manera 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 todos estos cuatro tipos de bases, de modo que no se producirán más de 3, 2 o incluso 1 tipo de trifosfato de desoxinucleótido mediante pirofosforólisis. En algunos casos, pueden estar presentes variantes de uno o más de estos cuatro tipos de bases en el ADN diana y en consecuencia la pirofosforólisis puede producir variantes de trifosfatos de desoxinucleótidos 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.

Otro ejemplo es el ARN, que típicamente incluye las cuatro bases comunes guanina, citosina, adenina y uracilo de manera que la pirofosforólisis producirá trifosfato de riboguanidina, trifosfato de ribocitidina, trifosfato de riboadenosina y trifosfato de ribotimidina. En algunos casos, el ácido nucleico diana puede no incluir todos estos cuatro tipos de bases, de modo que no se producirán más de 3, 2 o incluso 1 tipo de trifosfato de ribonucleótido mediante pirofosforólisis. Las bases variantes, tales como las ilustradas en la presente memoria, por ejemplo, con respecto a los trifosfatos de desoxinucleótidos, o conocidas de cualquier otra manera 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.

Un método de la presente descripción puede llevarse a cabo en condiciones que eliminan secuencialmente un número de nucleótidos de un ácido nucleico diana, produciendo de esta manera secuencialmente ese mismo número de trifosfatos de nucleótidos. Además, se puede distinguir al menos ese mismo número de trifosfatos de nucleótidos, por ejemplo, mediante el 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 realizaciones particulares, el número es de al menos 1, 2, 3, 4, 5, 10, 25, 50, 100, 200, 500, 1000, 10.000 o más hasta e incluida la longitud del ácido nucleico diana. Alternativamente o de forma adicional, el número puede ser no más de 1, 2, 3, 4, 5, 10, 25, 50, 100, 200, 500, 1000 o 10.000. El número puede estar, pero no tiene por qué estar, entre dos de estos valores. Como se expuso anteriormente en la presente memoria, se pueden usar una variedad de técnicas para pausar la pirofosforólisis. Esto puede proporcionar el control de la longitud de la secuencia determinada mediante las 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 diferentes tipos de trifosfatos de nucleótidos detectados. Sin embargo, el orden y el número de los diferentes trifosfatos de nucleótidos detectados pueden correlacionarse con la secuencia del ácido nucleico.

En algunas realizaciones, puede ser beneficioso secuenciar repetidamente un ácido nucleico diana particular. La repetición se puede lograr, 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 de (a) poner en contacto un ácido nucleico diana con una polimerasa para eliminar secuencialmente los trifosfatos de nucleótidos del ácido nucleico diana, donde los trifosfatos de nucleótidos que se eliminan tienen una variedad de diferentes restos de bases; (b) distinguir los diferentes restos de bases de los trifosfatos de nucleótidos que se eliminan, determinándose de esta manera la secuencia del ácido nucleico diana; (c) regenerar al menos una porción del ácido nucleico diana; y repetir las etapas (a) y (b) con el ácido nucleico diana.

regenerado. El ácido nucleico diana se puede regenerar, 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 mediante pirofosforólisis. Típicamente, el pirofosfato estará sustancialmente ausente durante la reacción de polimerización.

Alternativamente o de forma adicional a la realización de procesamiento repetido anterior, un ácido nucleico diana se puede amplificar o copiar para crear múltiples copias que se procesan con un método de la presente descripción. Se muestra un ejemplo diagramático en la **FIGURA 3**. Múltiples copias de un ácido nucleico molde bicatenario se localizan en una barrera en una cámara que tiene una fusión nanoporo-polimerasa (etapa 1), una de las cadenas es capturada por la polimerasa (etapa 2), se produce la secuenciación basada en pirofosforólisis (etapa 3), la cadena molde, que es monocatenaria, puede arrastrarse después a través del nanoporo mediante fuerza eléctrica (etapa 4) hasta que salga de la cámara donde permanecen las otras copias del ácido nucleico (nota: las otras copias permanecen debido a que son bicatenarias y por lo tanto resistentes al paso a través del nanoporo) (etapa 5) y, después, se captura otra copia del ácido nucleico molde bicatenario para iniciar la repetición de las etapas 2 y siguientes. Se puede utilizar cualquiera de una variedad de métodos conocidos en la técnica para amplificar ácidos nucleicos, tales como los expuestos anteriormente en la presente memoria, para crear las múltiples copias del ácido nucleico diana.

Aunque el sistema de la **FIGURA 3** se ilustra para copias de un único molde, se entenderá que las especies de ácido nucleico que tienen diferentes secuencias se pueden usar similarmente. Por lo tanto, una variedad de diferentes especies de ácidos nucleicos bicatenarios se pueden localizar en una barrera en una cámara que tiene una fusión nanoporo-polimerasa (etapa 1), la polimerasa puede capturar una cadena de una primera especie (etapa 2), se puede producir una secuenciación basada en pirofosforólisis (etapa 3), la cadena molde de la primera cadena puede arrastrarse después a través del nanoporo mediante fuerza eléctrica (etapa 4) hasta que salga de la cámara donde permanecen las otras especies de ácido nucleico (etapa 5) y, después, se puede capturar otra especie de ácido nucleico bicatenario para iniciar 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 colocado en la barrera impermeable a los fluidos para formar un paso a través del cual un trifosfato de nucleótido puede pasar del primer depósito de fluido al segundo depósito de fluido; y (c) una mezcla de reacción en el primer depósito de fluido, la mezcla de reacción incluye una polimerasa, un ácido nucleico diana que tiene dos cadenas, y una concentración pirofosforolítica de pirofosfato. Los componentes utilizados en el aparato pueden ser uno o más de los ilustrados anteriormente en el contexto de varios métodos. Los componentes y configuraciones adicionales se ilustran a continuación a título ilustrativo.

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. Se han expuesto anteriormente nanoporos y barreras ilustrativos y en varias referencias expuestas anteriormente. Generalmente, los dos depósitos estarán en comunicación fluida a través de un único nanoporo. Por lo tanto, los trifosfatos de nucleótidos producidos en uno de los depósitos tendrán una y solo una trayectoria de fluido hacia el segundo depósito. El uso de un único nanoporo de esta manera permite una medición práctica de cada trifosfato de nucleótido que pasa de un depósito al otro debido a cambios en las propiedades eléctricas en el nanoporo, barrera y/o depósitos. Sin embargo, en algunas realizaciones, también es posible incluir más de un nanoporo en la barrera que separa dos depósitos. Cuando varios nanoporos conectan de manera fluida dos depósitos, el paso de trifosfatos de nucleótidos se puede medir en el nanoporo individual con, por ejemplo, mediciones ópticas o eléctricas que resuelven cada nanoporo.

Un depósito puede crear una cámara donde el fluido permanece contenido durante al menos parte del tiempo. Por ejemplo, se puede configurar una cámara para formar una depresión, cavidad, compartimento, etc. que restringe el flujo de fluido. Alternativamente, se puede configurar un depósito para el flujo de fluido. Por ejemplo, el depósito se puede configurar como un tubo, canal o célula de flujo y permitir de esta manera el flujo de fluidos para el suministro y la retirada prácticos de los componentes utilizados en un método de secuenciación. En realizaciones particulares, un primer depósito que contiene ácido nucleico molde, polimerasa y pirofosfato se configurará para el flujo de fluido, mientras que el segundo depósito, que está conectado a la primera cámara a través de un nanoporo, se puede configurar como una cámara. No es necesario configurar el segundo depósito para el flujo de fluido, pero se puede hacer opcionalmente.

La presente descripción proporciona realizaciones multiplex. Por ejemplo, las secuencias para una pluralidad de ácidos nucleicos diana se pueden determinar en paralelo. Un método multiplex puede incluir las etapas de (a) proporcionar una pluralidad de ácidos nucleicos diana; (b) poner en contacto cada uno de los ácidos nucleicos diana con una polimerasa para eliminar secuencialmente los trifosfatos de nucleótidos de cada ácido nucleico diana, 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 de cada ácido nucleico, determinándose de esta manera las secuencias de los ácidos nucleicos diana.

Un ejemplo adicional de un método multiplex es uno que incluye las etapas de (a) proporcionar una pluralidad 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 para eliminar secuencialmente nucleótidos de la primera de cada una de las dos cadenas mediante pirofosforólisis, produciendo de esta manera secuencialmente trifosfatos de nucleótidos que tienen

una variedad de diferentes restos de bases; y (c) distinguir los diferentes restos de bases para los trifosfatos de nucleótidos producidos secuencialmente, determinándose de esta manera la secuencia de los ácidos nucleicos diana.

Un aparato multiplex puede incluir (a) una pluralidad de barreras impermeables a los fluidos, cada una de las cuales separa un primer depósito de fluido de un segundo depósito de fluido; (b) un nanoporo colocado en cada una de las barreras impermeables a los fluidos para formar un paso a través del cual un trifosfato de nucleótido puede pasar del primer depósito de fluido al segundo depósito de fluido; y (c) una mezcla de reacción en cada uno de los primeros depósitos de fluido, en la cual cada una de las mezclas de reacción incluye una polimerasa, ácido nucleico diana que tiene dos cadenas, y una concentración pirofosforolítica de pirofosfato.

La plexidad (es decir, nivel multiplex) de un método o aparato se puede seleccionar para satisfacer un uso particular. Por ejemplo, el número de ácidos nucleicos diana que se procesan o están presentes juntos se puede determinar a partir de la complejidad de la muestra a evaluar. Las estimaciones de complejidad ilustrativas para algunos de los genomas que pueden evaluarse con métodos o aparatos de la presente descripción son 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 de los ilustrados anteriormente incluidos, por ejemplo, genomas más pequeños o más grandes, se pueden utilizar en un método de la invención. Típicamente, una muestra de ácido nucleico se fragmenta antes de su uso. El número de fragmentos que se manejarán en paralelo dependerá de la complejidad del genoma, del tamaño medio de los fragmentos y de la cobertura deseada. Por ejemplo, se puede lograr una cobertura de 1x un genoma humano (3,1 Gbp) que está fragmentado hasta un tamaño medio de 1000 nucleótidos mediante una plexidad de 3 millones de fragmentos (es decir, $((3100 \text{ millones} / 1000) \times 1) = 1 \text{ millón}$). Con cálculos similares, se puede determinar que una plexidad de 30 millones de fragmentos (de 1000 nucleótidos cada uno) es suficiente para proporcionar una cobertura de 30x de un genoma humano.

Los métodos y aparatos expuestos en la presente memoria se pueden configurar a un nivel de plexidad suficiente para proporcionar al menos una cobertura de 1x, 2x, 5x, 10x, 20x, 30x, 50x o más de cualquiera de una variedad de genomas incluidos, aunque no de forma limitativa, los ilustrados en la presente memoria. La plexidad puede ser una función del número de varios componentes expuestos en la presente memoria tal como el número de fragmentos de ácido nucleico diana como se ilustró 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 un nanoporo, etc. El nivel multiplex de estos u otros componentes puede ser, por ejemplo, al menos 2, 5, 10, 100, 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^9 o superior. Alternativamente o de forma adicional, el nivel multiplex se puede seleccionar para ser no más de 2, 5, 10, 100, 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , o 1×10^9 .

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

Aunque la invención se ha descrito haciendo referencia a los ejemplos proporcionados anteriormente, la invención está limitada solo por las reivindicaciones.

REIVINDICACIONES

1. Un método para determinar la secuencia de un ácido nucleico diana, que comprende
 - (a) proporcionar un ácido nucleico diana que tiene dos cadenas;
 - (b) poner en contacto el ácido nucleico diana con una polimerasa unida a un nanoporo en condiciones para eliminar secuencialmente nucleótidos de una primera cadena de las dos cadenas mediante pirofosforólisis y de esta manera producir secuencialmente trifosfatos de nucleótido individuales; y
 - (c) distinguir los trifosfatos de nucleótido producidos secuencialmente mediante la detección de cambios en la corriente que fluye a través del nanoporo.
2. El método de la reivindicación 1, en donde el nanoporo comprende un nanoporo de proteína que está incrustado en una membrana.
3. El método de la reivindicación 1, en donde una segunda cadena de las dos cadenas del ácido nucleico diana está unida a la membrana.
4. El método de la reivindicación 1, en donde el nanoporo es un poro en estado sólido.
5. El método de la reivindicación 1, en donde las condiciones para eliminar secuencialmente trifosfatos de nucleótidos comprenden poner en contacto la polimerasa con una concentración pirofosforolítica de pirofosfato
6. El método de la reivindicación 5, en donde la concentración pirofosforolítica de pirofosfato comprende una concentración de al menos 100 pM.
7. El método de la reivindicación 5, en donde el método comprende además una etapa para pausar la producción secuencial de los trifosfatos de nucleótido al hacer que el pirofosfato deje de estar en contacto con la polimerasa y, a continuación, reanudar la producción secuencial de los trifosfatos de nucleótidos poniendo en contacto la polimerasa con el pirofosfato.
8. El método de la reivindicación 1, en donde los trifosfatos de nucleótido comprenden adenina, guanina, citosina o timina naturales.
9. El método de la reivindicación 1, en donde el ácido nucleico diana es el ADN.
10. El método de la reivindicación 1, en donde el ácido nucleico diana incluye al menos un resto base que no es natural en el ADN o el ARN.
11. El método de la reivindicación 10, en donde al menos uno de los restos base comprende 5-metilcitosina o 5-hidroximetilcitosina.
12. El método de la reivindicación 1, en donde el ácido nucleico diana es un heterodúplex formado por una cadena de ADN y una cadena de ARN.
13. El método de la reivindicación 1, en donde las condiciones para eliminar secuencialmente trifosfatos de nucleótido de la primera cadena comprenden poner en contacto la polimerasa con una concentración pirofosforolítica de un análogo de pirofosfato.
14. El método de la reivindicación 13, en donde el análogo de pirofosfato comprende un pirofosfato que tiene un resto adicional.
15. El método de la reivindicación 13, en donde el análogo de pirofosfato comprende pirovanadato.

Figura 1

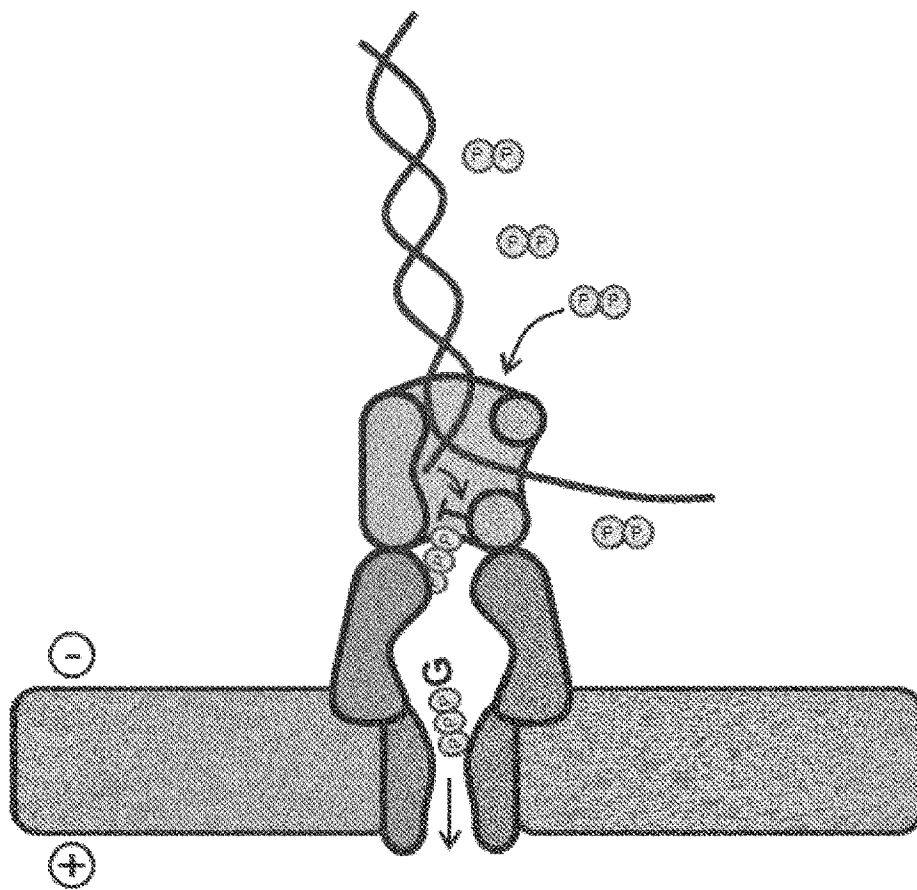


Figura 2

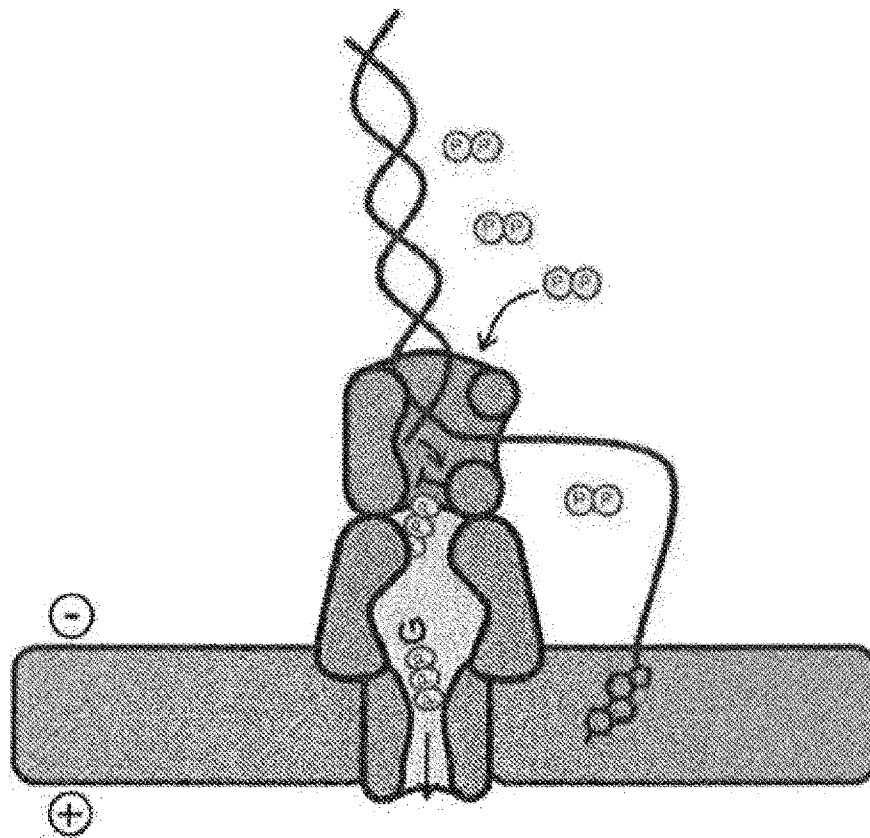


Figura 3

