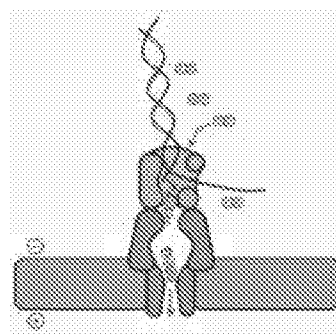


Resumo

Um método para determinar o sequenciamento de um ácido nucleico alvo, incluindo passos de contacto de um ácido nucleico alvo com uma polimerase para remover, sequencialmente, trifosfatos nucleótidos a partir do ácido nucleico alvo, em que os trifosfatos nucleótidos que são removidos possuem uma variedade de diferentes porções de base; e distinguindo as diferentes porções de base para os trifosfatos nucleótidos que são removidos. É também fornecido um aparelho que inclui um nanoporo posicionado numa barreira impermeável a fluidos para formar uma passagem através da qual um trifosfato nucleótido pode passar de um primeiro reservatório de fluido para um segundo reservatório de fluido, e uma mistura de reação no primeiro reservatório de fluido que inclui uma polimerase, ácido nucleico alvo com duas cadeias e uma concentração pirofosforolítica de pirofosfato.



Descrição

SEQUENCIAMENTO PIROPOSFOROLÍTICO USANDO NANO POROS

ANTECEDENTES

Esta divulgação refere-se, geralmente, a análises de ácidos nucleicos, e mais especificamente à síntese de ácido nucleico usando nano poros.

As plataformas comerciais atualmente disponíveis para sequenciar o DNA são relativamente dispendiosas. Essas plataformas usam uma abordagem de "sequenciamento por síntese", assim chamada porque os polímeros de DNA são sintetizados enquanto se detecta a adição de cada monómero (isto é, nucleotídeo) à estrutura de polímero em crescimento. Porque uma cadeia do DNA modelo direciona, estritamente, a síntese de um novo polímero de DNA, pode-se inferir o sequenciamento do modelo de DNA da série de monómeros nucleotídicos que foram adicionados à cadeia de crescimento durante a síntese. A capacidade de detectar adições de monómeros é facilitada por variantes especialmente criadas em engenharia dos componentes bioquímicos que normalmente realizam a síntese de DNA em sistemas biológicos. Esses componentes de engenharia são caros de fazer e são consumidos em quantidades relativamente elevadas durante o sequenciamento por síntese. Além disso, a monitorização da reação usa hardware relativamente caro, como lasers, ótica de detecção e sistemas complexos de fornecimento de fluidos. As plataformas comerciais mais bem-sucedidas até à data também

exigem reagentes e hardware caros para amplificar os modelos de DNA antes que o sequenciamento por síntese possa até começar.

Outros métodos de sequenciamento foram considerados, de modo a reduzir custos, aumentar o rendimento e/ou simplificar o processo. Uma dessas abordagens baseia-se no rosquear de uma única cadeia de DNA através de um nanoporo e na identificação do seu sequenciamento a partir da variação na corrente iônica que flui através do poro à medida que a cadeia é roscada. Uma alternativa a esta abordagem de sequenciamento de "nanoporo-cadeia" é o sequenciamento de "nanoporo-exonuclease", que envolve a remoção catalisada por exonuclease de monofosfatos de nucleótidos, um de cada vez, a partir de uma cadeia de DNA e passando sequencialmente os monofosfatos de nucleótidos libertados através de um nanoporo. No entanto, as variações resultantes na corrente iônica que flui através dos nanoporos são bastante pequenas e é difícil distinguir um nucleótido do outro. Foram feitas tentativas para modificar o DNA antes da digestão ou para modificar os monofosfatos nucleótidos, uma vez que foram libertados. No entanto, apesar destes esforços, o sequenciamento de nanoporo-exonuclease ainda não foi demonstrado a um nível comercialmente viável até à data.

Assim, existe a necessidade de plataformas mais econômicas, rápidas e convenientes que forneçam uma alternativa àquelas atualmente disponíveis para o sequenciamento de ácidos nucleicos. A presente divulgação aborda esta necessidade e também fornece outras vantagens.

Os seguintes documentos podem ser úteis para entender a presente divulgação.

O documento WO 2003080861 A1 refere-se a um processo para o sequenciamento de ácidos nucleicos em que a molécula de ácido nucleico a ser sequenciada é degradada sequencialmente na presença de um reagente marcado com fluorescência em que um nucleósido marcado com fluorescência é formado com características de fluorescência específicas de nucleobase. Preferencialmente, o método é um procedimento de sequenciamento de molécula única, compreendendo um passo de detecção resolvido espacialmente, por exemplo, detecção confocal. Preferencialmente, o método compreende um passo de detecção confocal. Além disso, são fornecidos novos corantes de fluorescência específicos de nucleobase e reagentes contendo os ditos corantes.

O documento US 6268146 B1 descreve processos espectroscópicos de massa, espectroscópicos de absorvência e espectroscópicos de fluorescência para detetar a despolimerização de um híbrido de ácido nucleico de modo a ensaiar, qualitativa e quantitativamente, para a presença de um ácido nucleico alvo predeterminado. As aplicações desses processos incluem a detecção de polimorfismos de nucleotídeos únicos, identificação de mudanças de base única, especificação, determinação de carga viral, genotipagem, diagnóstico de marcadores médicos e similares, incluindo ensaios multiplexados.

SUMÁRIO BREVE

O método e o aparelho da invenção estão definidos nas reivindicações.

A presente descrição fornece um método para determinar o sequenciamento de um ácido nucleico alvo. O método pode incluir as etapas de (a) fornecer um ácido nucleico alvo; (b) contactar o ácido nucleico alvo com uma polimerase para remover sequencialmente trifosfatos nucleótidos do ácido nucleico alvo, em que os trifosfatos nucleótidos que são removidos possuem uma variedade de porções de base diferentes; e (c) distinguir as diferentes porções de base para os trifosfatos nucleótidos que são removidos, determinando assim o sequenciamento do ácido nucleico alvo.

O método para determinar a sequência de um ácido nucleico alvo é realizado usando as etapas de (a) fornecer um ácido nucleico alvo com duas cadeias;

(b) contactar o ácido nucleico alvo com uma polimerase em condições para remover, sequencialmente, nucleótidos da primeira das duas cadeias por pirofosforólise, produzindo sequencialmente trifosfatos nucleótidos possuindo uma variedade de porções de base diferentes; e (c) distinguir as porções de base diferentes para os trifosfatos nucleótidos produzidos sequencialmente, determinando assim o sequenciamento do ácido nucleico alvo.

A presente invenção também fornece um aparelho que inclui (a) uma barreira impermeável a fluidos que separa um primeiro reservatório de fluido de um segundo reservatório de fluido; (b) um nanoporo posicionado na barreira impermeável a fluidos para formar uma passagem através da qual um trifosfato nucleótido pode passar do primeiro reservatório de fluido para o segundo reservatório de fluido; e (c) uma mistura de reação no primeiro reservatório de fluido, a mistura de reação compreendendo uma polimerase, um ácido nucleico alvo tendo duas cadeias e a concentração pirofosforolítica do pirofosfato.

BREVE DESCRIÇÃO DOS DESENHOS

A FIG. 1 mostra um diagrama de uma reação de sequenciamento pirofosforolítico usando uma polimerase anexada a um nanoporo.

A FIG. 2 mostra um diagrama de uma reação de sequenciamento pirofosforolítico usando uma polimerase anexada a um nanoporo e um modelo de ácido nucleico anexado a uma barreira impermeável a fluidos.

A FIG. 3. mostra um sequenciamento pirofosforolítico com semente de membrana de múltiplos modelos de ácido nucleico.

DESCRIÇÃO DETALHADA

A presente descrição fornece um método de sequenciamento de ácidos nucleicos de uma forma inversa em comparação com técnicas de sequenciamento padrão por síntese (SBS). Em formas de realização particulares, o método da presente divulgação explora uma função catalítica da polimerase conhecida como pirofosforólise que é, tipicamente, caluniada como culpada por artefactos indesejados em técnicas de SBS. A pirofosforólise resulta na remoção de trifosfatos nucleótidos a partir de uma cadeia de ácido nucleico por uma polimerase e, como tal, é o inverso da reação de polimerização que impulsiona técnicas padrão de SBS.

A pirofosforólise pode ser distinguida da atividade da exonuclease (que está presente em algumas polimerases), por exemplo, com base no mecanismo catalítico diferente para as duas reações, diferentes locais ativos na estrutura da polimerase onde ocorrem as duas reações e os diferentes produtos para as reações. Em relação ao mecanismo catalítico e às diferenças de locais ativos, sabe-se que a atividade de exonuclease pode ser removida de muitas espécies de polimerase por deleção de certos domínios, enquanto que a atividade de pirofosforólise acredita-se que seja catalisada pelo mesmo domínio que catalisa a polimerização. Além disso, o ciclo de reação catalisado pela exonuclease é a conversão de DNA_n (DNA de comprimento n) para DNA_{n-1} (DNA que é um nucleotídeo mais curto do que o DNA de comprimento n) e um nucleótido monofosfato. Em contraste, um ciclo de pirofosforólise produz DNA_{n-1} e um trifosfato nucleótido do DNA_{n-1} e pirofosfato. Notavelmente, o pirofosfato é consumido numa reação de pirofosforólise, mas não é consumido numa reação de exonuclease.

Os métodos de sequenciamento de pirofosforólise utilizam a detecção de nanoporos. Os nanoporos podem ser usados para detetar sequencialmente os trifosfatos nucleótidos que são libertados de um ácido nucleico, de modo a determinar o sequenciamento do ácido nucleico. Tais formas de realização fornecem uma combinação de vantagens que são, tipicamente, apenas parcialmente satisfeitas por sequenciamento de nanoporo-exonuclease ou por sequenciamento de nanoporo-cadeia. Especificamente, os métodos de sequenciamento pirofosforolíticos da presente invenção, abordam alguns dos desafios incorridos no sequenciamento de nanoporo-exonuclease, ou seja, poucas probabilidades de captura e detecção, mantendo a sua principal vantagem em relação ao sequenciamento de cadeias, nomeadamente a resolução de base única. Esta vantagem deriva do facto de que a afinidade de nanoporos para monofosfatos nucleótidos é bastante fraca (na ordem da afinidade micromolar), mesmo na presença de um adaptador de ciclodextrina- α 6 que tenha sido usado para melhorar o sinal em alguns sistemas de nanoporos. Ver Clarke et al., Nat. Nanotechnol. Apr;4(4):265-70 (2009). Para uma distinção bem-sucedida de diferentes tipos de nucleótidos num contexto de sequenciamento, é desejada uma afinidade no alcance nanomolar. O ATP tem uma afinidade que está num bom alcance, mesmo sem o uso de um adaptador no poro. Ver Cheley et al., Chem. Biol. 9,829-838 (2002). Sem querer ser vinculado pela teoria, postula-se que a afinidade melhorada do ATP sobre o monofostato nucleótido é devido à carga tripla negativa carregada pelo ATP, o que pode fazer com que ele se apegue mais fortemente dentro do nanoporo. Além disso, a carga tripla negativa pode ajudar a capturar a molécula na presença de um campo elétrico, especialmente quando o campo é muito fraco, como é o caso fora do nanoporo onde os nucleotídeos são realmente libertados.

Além da captura e detecção melhoradas de trifosfatos nucleótidos, existe um número de outras vantagens que podem ser fornecidas por formas de realização aqui expostas, tais como o uso de uma polimerase em oposição a uma exonuclease. As polimerases demonstraram formar um bom "encaixe" com nanoporos, com o objetivo do sequenciamento de nanoporo-cadeia (Cherf et al, Nat. Biotech. 30:344-348 (2012); Manrao et al., Nat. Biotech. 30:349-353 (2012)). Um encaixe semelhante ainda está para ser demonstrado com exonucleases. Além disso, o substrato para polimerases é um DNA de cadeia dupla que geralmente não entra no nanoporo. Em contraste, o DNA de cadeia única, o substrato para a maioria das exonucleases, pode representar o problema de entrar e bloquear o nanoporo. Finalmente, ao contrário de ambos no sequenciamento baseado em exonuclease ou no sequenciamento baseado em cadeia baseado em polimerase, a taxa de uma reação de sequenciamento pirofosforolítica pode ser controlada ajustando a concentração do pirofosfato.

Os termos aqui usados serão entendidos para assumir o seu significado comum, a menos que especificado de outra forma. Os exemplos de vários termos aqui usados e suas definições são apresentados abaixo.

Tal como aqui usado, o termo "anexado" pretende significar conectado por forças que impedem a separação por difusão. O termo pode incluir conexões que são covalentes ou não covalentes por natureza. Por exemplo, duas proteínas podem ser unidas de forma covalente através da sua sequência primária (por exemplo, a ligação de péptido ou fusão de proteína) ou entre as suas sequências primárias (por exemplo, ligação de dissulfureto ou reticulação química via grupos de aminoácidos R). Duas proteínas podem ser unidas

não covalentemente, por exemplo, através de uma ou mais ligações de hidrogénio, ligações iónicas, forças de van der Waals, ligações hidrofóbicas ou semelhantes.

Conforme usado aqui, o termo "cada um", quando usado em referência a uma coleção de itens, destina-se a identificar um item individual na coleção, mas não se refere necessariamente a cada item da coleção. As exceções podem ocorrer se a divulgação ou o contexto explícito ditarem de forma inequívoca o contrário.

Conforme usado aqui, o termo "atividade da exonuclease" pretende significar a hidrólise da ligação fosfodiéster que anexa um nucleótido à extremidade de um ácido nucleico de comprimento n para produzir um monofosfato nucleótido e um ácido nucleico de comprimento $n-1$. A hidrólise pode ocorrer na ligação que anexa o nucleótido 5' ao ácido nucleico (isto é, atividade de exonuclease de 5' a 3') ou na ligação que anexa o nucleótido 3' ao ácido nucleico (isto é, atividade de exonuclease de 3' a 5').

Tal como aqui usado, o termo "barreira impermeável a fluidos" pretende significar qualquer coisa que impeça a passagem de um fluido. Por exemplo, uma barreira impermeável a fluidos pode estar presente entre dois reservatórios de modo que um fluido no primeiro reservatório seja separado do fluido no segundo reservatório e os fluidos não se misturam. Uma barreira impermeável a fluidos pode incluir um poro ou abertura que permite a passagem de um fluido que de outra forma está obstruído pelo restante da barreira. Em formas de

realização particulares, o fluido pode ser um fluido aquoso e a barreira pode ser impermeável aos fluídos aquosos.

Tal como aqui usado, o termo "falta de atividade de exonuclease" pretende significar que não há atividade de exonuclease mensurável. Por exemplo, uma polimerase ou outro agente que não possui atividade de exonuclease de 3' a 5' não terá atividade de exonucleases mensurável de 3' a 5'.

De modo semelhante, uma polimerase ou outro agente que não possui a atividade de exonuclease de 5' a 3' não terá atividade de exonuclease de 5' a 3' mensurável. As polimerases conhecidas na técnica como "exominus" (ou "exo-"), sejam de ocorrência natural ou de engenharia, são exemplos não limitantes de polimerases que não possuem atividade de exonuclease. As variantes conhecidas incluem aquelas que são 5' a 3' exominus e/ou 3' a 5' exominus.

Tal como aqui usado, o termo "nanoporo" pretende significar um pequeno orifício que permite a passagem de trifosfatos nucleótidos através de uma barreira impermeável. A barreira é tipicamente uma camada eletricamente isolante e o nanoporo tipicamente permite que os iões fluam de um lado da barreira para o outro, impulsionados por um potencial aplicado.

O nanoporo permite, preferencialmente, que os nucleotídeos fluam através do lúmen do nanoporo ao longo do potencial aplicado. O nanoporo também pode permitir que um ácido nucleico, como DNA ou RNA, seja empurrado ou puxado através do lúmen do nanoporo. No entanto, em formas de realização

particulares, o nanoporo não precisa permitir a passagem de um ácido nucleico de cadeia dupla ou de cadeia simples.

Um nanoporo usado numa forma de realização particular pode ter um diâmetro mínimo do lúmen não superior a 10 nm, 5 nm, 4 nm, 3 nm, 2 nm, 1 nm, 0.5 nm ou menor. Um tipo de nanoporo é o "nanoporo de proteína" que é um polipéptido ou uma coleção de polipéptidos que forma o pequeno orifício para permitir a passagem de trifosfatos nucleótidos através de uma barreira tal como uma bicamada lipídica. Os exemplos de nanoporos de proteína incluem o nanoporo de alfa-hemolisina, *mycobacterium smegmatis* porin A (MspA) e suas variantes. Outro tipo de nanoporo é o "nanoporo de estado sólido", que é um pequeno orifício fabricado através de um material sólido. O material sólido pode ser, por exemplo, grafeno ou silício.

Tal como aqui usado, o termo "nucleótido" pretende incluir ribonucleótidos, desoxinucleótidos ou seus análogos. Por exemplo, o termo é aqui usado para se referir, geralmente, a uma porção de nucleósido (seja ribose, desoxirribose ou um análogo dos mesmos) incluindo uma porção de base e opcionalmente ligada a uma ou mais porções de fosfato. Os nucleótidos exemplares incluem, mas não estão limitados a, monofosfato de ribonucleótido (às vezes referido como um monofosfato de ribonucleósido), difosfato de ribonucleótido (às vezes referido como um difosfato de ribonucleósido), trifosfato de ribonucleótido (às vezes referido como um trifosfato de ribonucleósido), monofosfato de desoxinucleótido (às vezes referido como um monofosfato de desoxinucleósido), desoxinucleótido difosfato (às vezes referido como um desoxinucleósido difosfato) e desoxinucleótido trifosfato (às vezes referido como

desoxinucleósido trifosfato). Para maior clareza, quando se deseja distinguir componentes de RNA de componentes de DNA, o termo "ribonucleótido" pode ser usado para especificar nucleótidos de RNA, tais como trifosfato de riburidina, trifosfato de riboguanidina, trifosfato de ribocitidina ou trifosfato de riboadenosina; e o termo "desoxinucleótido" pode ser usado para especificar nucleótidos de DNA, tais como trifosfato de desoxitimidina, trifosfato de desoxiguanidina, trifosfato de desoxicitidina e trifosfato de desoxiadenosina. Em formas de realização particulares, os nucleótidos são "extensíveis", por exemplo, sem uma porção de bloqueio de extensão no hidroxilo 3' ou em qualquer outra posição no nucleótido.

Tal como aqui usado, o termo "pirofosforólise" pretende significar a reação entre o pirofosfato e a unidade de nucleótido 3' de um ácido nucleico para libertar o nucleótido na forma do trifosfato nucleótido correspondente. Um produto adicional da reação é o ácido nucleico que não possui o nucleótido correspondente (isto é, a reação converte um ácido nucleico de comprimento n num ácido nucleico de comprimento $n-1$). A reação é tipicamente catalisada por uma polimerase e é o inverso da reação de polimerização realizada pela polimerase sob condições biológicas padrão.

Tal como aqui usado, o termo "concentração pirofosforolítica", quando usado em referência ao pirofosfato, pretende significar uma concentração de pirofosfato que provoca uma reação de pirofosforólise a um nível substancial. Por exemplo, uma concentração pirofosforolítica de pirofosfato pode resultar numa polimerase exibindo uma maior taxa de pirofosforólise do

que a polimerização. Assim, uma concentração pirofosforolítica de pirofosfato pode resultar numa reversão substancial da atividade de polimerização que de outra forma seria catalisada por uma polimerase.

Tal como aqui usado, o termo "reservatório" pretende significar um recetáculo ou câmara para manter ou restringir o fluxo do fluido. Um reservatório pode ser completamente fechado, pelo menos temporariamente. Alternativamente, um reservatório pode ser parcialmente fechado, por exemplo, tendo uma ou mais aberturas (por exemplo, uma ou mais entradas ou saídas). Um fluido pode fluir através de um reservatório, por exemplo, em formas de realização em que o reservatório está numa célula de fluxo.

Tal como aqui usado, o termo "espécie" é usado para identificar moléculas que compartilham a mesma estrutura química. Por exemplo, uma mistura de nucleótidos pode incluir várias moléculas de dCTP. As moléculas de dCTP serão entendidas como sendo da mesma espécie umas das outras.

Da mesma forma, as moléculas de DNA individuais que têm a mesma sequência de nucleótidos são a mesma espécie.

As formas de realização apresentadas abaixo e citadas nas reivindicações podem ser entendidas em vista das definições acima.

A presente descrição fornece um método para determinar a sequência de um ácido nucleico alvo. O método pode incluir

as etapas de (a) fornecer um ácido nucleico alvo; (b) contactar o ácido nucleico alvo com uma polimerase para remover sequencialmente os trifosfatos nucleótidos do ácido nucleico alvo, em que os trifosfatos nucleótidos que são removidos possuem uma variedade de diferentes porções de base; e (c) distinguir as diferentes porções de base para os trifosfatos nucleótidos que são removidos, determinando assim a sequência do ácido nucleico alvo.

Em algumas partes da descrição, um método para determinar a sequência de um ácido nucleico alvo pode ser realizado usando as etapas de (a) fornecer um ácido nucleico alvo com duas cadeias; (b) contactar o ácido nucleico alvo com uma polimerase em condições para a remoção sequencial de nucleótidos da primeira das duas cadeias por pirofosforólise, produzindo sequencialmente trifosfatos nucleótidos possuindo uma variedade de diferentes porções de base; e (c) distinguir as diferentes porções de base para os trifosfatos nucleótidos produzidos sequencialmente, determinando assim a sequência do ácido nucleico alvo.

Um exemplo de uma forma de realização é mostrado na **FIG. 1**. Como mostrado, uma polimerase liga-se a uma molécula de DNA de cadeia dupla e, na presença de pirofosfato em excesso, catalisa uma reação de pirofosforólise para libertar trifosfatos nucleótidos da extremidade 3' de uma das cadeias. Neste exemplo, um trifosfato de desoxiguanidina foi produzido seguido por um trifosfato de desoxitimidina. A polimerase é acoplada a um nanoporo e o trifosfato de desoxiguanidina está no lúmen do nanoporo, enquanto que o trifosfato de desoxitimidina está no processo de libertação no lúmen do nanoporo. Como tal, os trifosfatos desoxinucleótidos entram no nanoporo sequencialmente, na

mesma ordem em que foram libertados da cadeia de DNA pela ação pirofosforolítica da polimerase. Os trifosfatos desoxinucleótidos têm uma carga negativa líquida devido aos trifosfatos e são conduzidos através do nanoporo por um potencial através da membrana (como indicado pelo sinal positivo no lado da membrana onde ocorre a pirofosforólise e um sinal negativo no lado oposto da membrana). Tipicamente, os trifosfatos nucleótidos reagentes não estão presentes numa reação de pirofosforólise. Em alguns casos, existem vestígios de trifosfatos nucleótidos, mas em quantidades que não catalisam, substancialmente, uma reação de extensão primária, adiante, pela polimerase. Assim, na maioria das formas de realização, os únicos trifosfatos nucleótidos que estão, substancialmente, presentes numa reação de pirofosforólise são aqueles produzidos pela ação catalítica da polimerase no ácido nucleico.

Uma reação semelhante é exemplificada na **FIG. 2**, exceto que a cadeia modelo (isto é, a cadeia que não está sendo cortada pela pirofosforólise) está ligada à membrana. Neste caso, a cadeia modelo tem uma porção de esterol anexada covalentemente (por exemplo, colesterol) que se liga ao interior hidrofóbico da bicamada de lipídios da membrana.

Um método da presente divulgação pode ser usado com qualquer uma de uma variedade de ácidos nucleicos alvo. O ácido nucleico alvo pode ter uma estrutura natural ocorrente como encontrada, por exemplo, no DNA ou RNA. O DNA contém naturalmente monómeros com bases de timina, guanina, citosina ou adenina. Uma cadeia de DNA que é submetida à pirofosforólise pode produzir desoxitimidina trifosfato, desoxiguanidina trifosfato, desoxicitidina trifosfato e desoxiadenosina trifosfato, respetivamente. O

DNA também pode conter algumas variantes das quatro bases nucleotídicas, tais como 5-metil citosina, 5-hidroximetilcitosina e outras bases metiladas. Os trifosfatos desoxinucleótidos possuindo estas bases variantes podem ser produzidos e/ou detetados usando um método ou aparelho aqui apresentado. Por exemplo, a presença ou ausência de metilação na citosina pode ser distinguida para facilitar análises epigenéticas. O RNA contém naturalmente monómeros com bases de uracilo, guanina, citosina ou adenina. Uma cadeia de RNA que é submetida à pirofosforólise pode produzir trifosfato de ribouridina, trifosfato de riboguanidina, trifosfato de ribocitridina ou trifosfato de riboadenosina, respetivamente.

Um ácido nucleico pode incluir modificações que não ocorrem naturalmente, tais como bases não nativas, modificações nas porções de fosfato ou modificações nas porções de açúcar. As bases não-nativas exemplificativas que podem ser incluídas num ácido nucleico, seja com um esqueleto nativo ou estrutura análoga, incluem, sem limitação, inosina, xatanina, hipoxatanina, isocitosina, isoguanina, 2-aminoadenina, 6-metilo adenina, 6-metilo guanina, 2-propilo guanina, 2-propilo adenina, 2-tioLiracil, 2-tiotimina, 2-tiocitosina, 15-halouracilo, 15-halocitosina, 5-propinilo uracilo, 5-propinilo citosina, 6-azo uracilo, 6-azo citosina, 5-uracilo, 4-tiouracilo, 8-halo adenina ou guanina, 8-amino adenina ou guanina, 8-tiol adenina ou guanina, 8-tioalquilo adenina ou guanina, 8-hidroxilo adenina ou guanina, citosina ou uracilo substituído de 5-halo, 7-metiloguanina, 7-metiloadenina, 8-azaguanina, 8-azaadenina, 7-deazaguanina, 7-deazaadenina, 3-deazaguanina, 3-desazaadenina ou semelhantes.

As bases não nativas podem ser selecionadas, por exemplo, para conferir o tamanho maior ou menor, ou para conferir a carga aumentada ou diminuída, de modo a influenciar a capacidade dos análogos de trifosfato nucleótido resultantes a serem distinguidos por um nanoporo ou outro componente de proteção. Do mesmo modo, tais bases podem ser benéficas se selecionadas para conferir uma taxa desejada de pirofosforólise. As bases não nativas podem ser incorporadas num ácido nucleico usando métodos conhecidos tais como amplificação ou replicação de um ácido nucleico modelo na presença dos análogos de nucleótidos. Uma ou mais das cópias de ácido nucleico resultantes podem então ser usadas como ácido(s) nucleico(s) alvo num aparelho ou método de sequenciamento aqui apresentado.

Em formas de realização particulares, um ácido nucleico que é usado num método ou aparelho aqui usado, não terá uma ou mais das bases não nativas ou outras porções não nativas aqui apresentadas. Por exemplo, em formas de realização particulares, os métodos não são usados para remover um nucleótido terminador da extremidade 3' de uma cadeia de ácido nucleico. Consequentemente, em algumas formas de realização, um aparelho ou método pode não incluir qualquer(s) ácido(s) nucleico(s) com um nucleótido terminador na sua extremidade 3'. Alternativa, ou adicionalmente, em algumas formas de realização, um aparelho ou método pode não incluir qualquer(s) nucleótido(s) terminador(s).

Como exemplificado na **FIG. 1** e em outro lugar aqui, um ácido nucleico alvo pode ser um DNA de cadeia dupla, por exemplo, quando se usa uma polimerase de DNA para a pirofosforólise. Um heteroduplex, formado entre uma cadeia

de DNA e uma cadeia de RNA, também pode ser usado. Por exemplo, uma polimerase de RNA pode ser usada para catalisar a pirofosforólise na extremidade 3' de uma cadeia de RNA que é hibridizada com uma cadeia de DNA modelo, produzindo deste modo os trifosfatos ribonucleótidos.

Um ácido nucleico que é usado num método ou aparelho aqui pode ser isolado de uma fonte biológica e usado diretamente ou processado para produzir um produto amplificado ou modificado. Alternativamente, um ácido nucleico sintético pode ser usado, de novo, diretamente ou após o processamento. O processamento pode incluir, por exemplo, um ou mais isolamentos de componentes nativos, clivagem para formar fragmentos, amplificação (por exemplo, via PCR, Rolling Circle ou outras técnicas conhecidas), ligação de sequências de adaptadores ou sequências de etiquetas, marcação usando uma transposição ou métodos de "preparação de amostras" usados antes das técnicas de sequenciamento de ácido nucleico.

As técnicas de processamento úteis são conhecidas na técnica, tal como estabelecido, por exemplo, em Sambrook et al., Molecular Cloning: A Laboratory Manual, 3ª edição, Cold Spring Harbor Laboratory, New York (2001) ou em Ausubel et al., Current Protocols in Molecular Biology, John Wiley and Sons, Baltimore, Md. (1998), cada um dos quais é incorporado aqui por referência.

Os exemplos de métodos de preparação de amostras que podem ser usados para processar ácidos nucleicos antes do sequenciamento baseado na pirofosforólise incluem métodos que foram desenvolvidos para a amplificação de genoma total

ou ampliação total de exoma em combinação com técnicas de sequenciamento massivamente paralelas. Por exemplo, técnicas de preparação de amostras comercialmente disponíveis de Illumina, Inc. (San Diego, CA), Life Technologies (Carlsbad, CA), 454 Life Sciences (a subsidiar of Roche, Basel, Switzerland) ou pode ser usado New England Biolabs (Ipswich, MA). Os métodos de preparação de amostras úteis que podem ser usados, são descritos, por exemplo, nas Patentes US N°s 6,107,023 ou 7,741,463; ou a Publicação da Aplicação da Patente US N°. 2010/0120098 A1. Os métodos de preparação de amostras alvo, também podem ser usados para isolar um subconjunto do conteúdo de sequência de uma amostra complexa para o sequenciamento posterior. Os métodos comerciais exemplificativos que podem ser usados para captura alvo de um subconjunto de fragmentos de ácido nucleico incluem, mas não estão limitados a SureSelect™ kits (Agilent, Santa Clara, CA), TruSeq® Enrichment Kits (Illumina, Inc., San Diego, CA) ou Nextera® Enrichment Kits (Illumina, Inc., San Diego, CA).

Os ácidos nucleicos podem ser isolados de qualquer uma das várias fontes, incluindo, sem limitação, um mamífero como um roedor, rato, ratazana, coelho, porquinho-da-índia, ungulado, cavalo, ovelha, porco, cabra, vaca, gato, cachorro, primata, primata humano ou não humano; uma planta como a *Arabidopsis thaliana*, milho (*Zea mays*), sorgo, aveia (*oryza sativa*), trigo, arroz, canola ou soja; uma alga como a *Chlamydomonas reinhardtii*; um nematodo como a *Caenorhabditis elegans*; um inseto como o *Drosophila melanogaster*, mosquito, mosca da fruta, abelha de mel ou aranha; um peixe como o peixe-zebra (*Danio rerio*); um réptil; um anfíbio como um sapo ou *Xenopus laevis*; um *dictyostelium discoideum*; um fungo como o *pneumocystis*

carinii, *Takifugu rubripes*, leveduras, *Saccharomyces cerevisiae* ou *Schizosaccharomyces pombe*; ou um *plasmodium falciparum*.

Os ácidos nucleicos também podem ser derivados de genomas menores, tais como as de um procarionte, como uma bactéria, *Escherichia coli*, *staphylococci* ou *mycoplasma pneumoniae*; uma archaea; um vírus como o vírus da hepatite C ou o vírus da imunodeficiência humana; ou um viroide.

Qualquer uma das várias polimerases pode ser usada num método ou aparelho aqui apresentado, incluindo, por exemplo, enzimas baseadas em proteína, isoladas de sistemas biológicos e suas variantes funcionais. A referência a uma polimerase particular, tal como aquelas exemplificadas abaixo, será entendida como incluir variantes funcionais da mesma, a menos que seja indicado de outra forma. Os exemplos de polimerases úteis incluem polimerases de DNA e polimerases de RNA. Os exemplos de polimerases de DNA incluem aqueles que foram classificados por homologia estrutural em famílias identificadas como A, B, C, D, X, Y e RT. As polimerases de DNA na família A incluem, por exemplo, a polimerase de DNA T7, a polimerase γ de DNA mitocondrial eucariótico, *E.coli* DNA Pol I, *Thermus aquaticus* Pol I e *Bacillus stearothermophilus* Pol I. As polimerases de DNA na família B incluem, por exemplo, polimerases de DNA eucarióticas polimerases α , δ e ϵ ; a polimerase de DNA ζ ; a polimerase de DNA T4, a polimerase de DNA *Phi29* e a polimerase de DNA de bacteriófago RB69. A família C inclui, por exemplo, a subunidade alfa da polimerase de DNA III de *E. coli*. A família D inclui, por exemplo, polimerases derivadas do subdomínio Euryarchaeota de Archaea. As polimerases de DNA

na Família X incluem, por exemplo, polimerases eucarióticas Pol β , Pol σ , Pol λ e Pol μ e *S. cerevisiae* Pol4. As polimerases de DNA na Família Y incluem, por exemplo, Pol η , Pol iota, Pol kappa, *E. coli* Pol IV (DINB) e *E. coli* Pol V (UmuD'2C). A família RT (transcriptase reversa) de polimerases de DNA inclui, por exemplo, transcriptases reversas de retrovírus e telomerases eucarióticas. As polimerases de RNA exemplares incluem, mas não estão limitadas a polimerases de RNA virais tais como a polimerase de RNA T7; a polimerase de RNA eucariótica, tais como a polimerase de RNA I, a polimerase de RNA II, a polimerase de RNA III, a polimerase de RNA IV e a polimerase de RNA V; e a polimerase de RNA Arcaea.

As classificações acima são fornecidas para fins ilustrativos. Deve entender-se que as variações no sistema de classificação são possíveis. Por exemplo, em pelo menos um sistema de classificação, as polimerases da Família C foram categorizadas como uma subcategoria da Família X. Além disso, as polimerases podem ser classificadas de acordo com outras características, quer funcionais ou estruturais, que podem ou não se sobrepor às características estruturais exemplificadas acima. Algumas características exemplares são apresentadas com mais detalhe abaixo.

Muitas polimerases têm uma atividade de exonuclease de revisão intrínseca de 3' a 5' que pode ser útil para algumas formas de realização. As polimerases que lhes falta, substancialmente, a atividade de exonuclease de revisão de 3' a 5' são preferidas em algumas formas de realização, por exemplo, na maioria das formas de realização de sequenciamento. A ausência de atividade de

exonuclease pode ser uma característica de tipo selvagem ou uma característica conferida por uma polimerase variante ou modificada.

Por exemplo, o fragmento *exo minus* Klenow é uma versão mutada do fragmento de Klenow que não possui atividade de exonuclease de revisão de 3' a 5'. O fragmento Klenow e a sua variante *exo minus* podem ser úteis num método ou aparelho aqui apresentado. As polimerases que catalisam a pirofosforólise, a reversão direta da polimerização no mesmo local ativo, são particularmente úteis.

Dependendo da forma de realização a ser usada, uma polimerase pode ser termófila ou inativável pelo calor. As polimerases termófilas são tipicamente úteis para condições de alta temperatura ou em condições de termociclagem tais como as empregues para técnicas de reação em cadeia da polimerase (PCR). Os exemplos de polimerases termófilas incluem, mas não estão limitados à polimerase de DNA 9°N, polimerase de DNA Taq, polimerase de DNA Phusion®, polimerase de DNA Pfu, polimerase de DNA RB69, polimerase de DNA KOD e polimerase de DNA VentR®. A maioria das polimerases isoladas de organismos não-termófilos são inativáveis pelo calor. A inativação por calor pode ser útil para parar uma reação de sequenciamento aqui estabelecida. A reação pode ser reativada opcionalmente, adicionando um novo fornecimento de polimerase à reação a uma temperatura apropriadamente permissiva. Os exemplos de polimerases inativáveis pelo calor são as do fago. Deve entender-se que as polimerases de qualquer uma variedade de fontes podem ser modificadas para aumentar ou diminuir a sua tolerância a condições de alta temperatura.

Pode ser usada uma variante modificada de uma polimerase com uma atividade de pirofosforólise aumentada. As variantes exemplares são aquelas que foram criadas para uso em técnicas de PCR incluindo, mas não se limitando às variantes descritas na Patente US N° 8,071,536, que é aqui incorporada por referência.

Uma polimerase pode ser induzida a remover, sequencialmente, nucleótidos de uma das duas cadeias de ácido nucleico por pirofosforólise num método aqui apresentado. A polimerase pode ser colocada sob condições para ocorrer a pirofosforólise. Por exemplo, a polimerase pode ser contactada com um ácido nucleico de cadeia dupla e uma concentração pirofosforolítica de pirofosfato. Pode ser usada qualquer concentração de pirofosfato que provoca uma reação de pirofosforólise a um nível substancial, incluindo, mas não limitado a, pelo menos cerca de 100 μM de pirofosfato. Podem ser empregues concentrações mais elevadas de pirofosfato, por exemplo, para conduzir a pirofosforólise a uma velocidade mais rápida. Consequentemente, pode ser usada uma concentração de pelo menos cerca de 250 μM de pirofosfato, pelo menos cerca de 500 μM de pirofosfato, pelo menos cerca de 750 μM de pirofosfato, pelo menos cerca de 1 mM de pirofosfato, pelo menos cerca de 2 mM de pirofosfato, pelo menos cerca de 5 mM de pirofosfato, pelo menos cerca de 10mM de pirofosfato, pelo menos cerca de 20mM de pirofosfato ou superior.

A capacidade de alterar a taxa da pirofosforólise é uma vantagem para ajustar a taxa da reação de sequenciamento, por exemplo, para acomodar uma taxa ótima ou desejada de deteção de trifosfato nucleótido para um determinado dispositivo de deteção. Por exemplo, a taxa da

pirofosforólise pode ser diminuída usando uma menor concentração de pirofosfato do que as estabelecidas acima. Assim, como uma alternativa ou adição às concentrações limiares estabelecidas acima, uma concentração máxima de pirofosfato presente num aparelho ou método aqui apresentado pode ter no máximo cerca de 20 mM de pirofosfato, no máximo cerca de 10 mM de pirofosfato, no máximo cerca de 5 mM de pirofosfato, no máximo cerca de 2 mM de pirofosfato, no máximo cerca de 1 mM de pirofosfato, no máximo cerca de 750 μ M de pirofosfato, no máximo cerca de 500 μ M de pirofosfato, no máximo cerca de 250 μ M de pirofosfato, no máximo cerca de 100 μ M de pirofosfato, ou menos.

Os trifosfatos nucleótidos reagentes estão tipicamente ausentes em condições da pirofosforólise. Assim, na maioria das formas de realização, os únicos trifosfatos nucleótidos que estão substancialmente presentes numa reação de pirofosforólise são aqueles produzidos pela ação catalítica da polimerase no ácido nucleico. Se os trifosfatos nucleótidos estiverem presentes em condições de pirofosforólise, os trifosfatos nucleótidos estarão presentes no que pode ser considerado uma concentração não catalítica. Uma concentração não catalítica de trifosfato nucleótido é uma concentração que é insuficiente para permitir que a atividade de extensão da polimerase substancial ou detetável ocorra. Por exemplo, uma concentração não catalítica de trifosfato nucleótido é uma concentração que está substancialmente abaixo da constante de ligação da associação para a ligação dos trifosfatos nucleótidos à polimerase. Consequentemente, a pirofosforólise pode ser realizada contactando uma polimerase com um ácido nucleico de cadeia dupla na

presença de uma concentração pirofosforolítica de pirofosfato e não mais do que uma concentração não catalítica do trifosfato nucleótido.

As condições exemplares para induzir a pirofosforólise que podem ser aqui usadas são apresentadas em Patel et al. *Biochem.* 30:511-525 (1991), ou nas Patentes US. N°s 7,090,975; 7,914,995 ou 8,071,536. Em vários casos, estas referências descrevem as reações que também incluem componentes usados para a extensão de um ácido nucleico. Será entendido que a extensão não é utilizada em formas de realização particulares da presente divulgação. Um ou mais dos componentes que são descritos em Patel et al. *Biochem.* 30:511-525 (1991) ou nas Patentes US. N°s 7,090,975; 7,914,995 ou 8,071,536, incluindo, por exemplo, componentes usados para a reação de extensão da polimerase, podem ser excluídos de um método ou aparelho aqui estabelecido.

As substâncias, sais, íons metálicos, glicerol, DMSO e outros componentes tipicamente incluídos no armazenamento da polimerase ou misturas de reação podem ser incluídos num aparelho ou método da presente divulgação, conforme desejado. As quantidades de componentes a incluir serão evidentes para os especialistas na técnica ou podem ser determinadas, por exemplo, através de técnicas de titulação de rotina.

Um aspecto benéfico de algumas formas de realização é a capacidade de parar ou pausar o método de sequenciamento alterando as condições de reação para inibir a pirofosforólise. O método de sequenciamento pode então ser reiniciado, opcionalmente, alterando as condições da reação

para permitir que a pirofosforólise seja retomada. Qualquer uma das várias condições de reação aqui expostas, ou nas referências aqui citadas, pode ser alterada entre pirofosforólise em pausa e pirofosforólise retomada, permitindo, assim, uma alternância efetiva entre o sequenciamento pausado e ativo, respetivamente.

Em formas de realização particulares, um método da presente divulgação pode incluir um passo de pausa na produção sequencial de trifosfatos nucleótidos removendo o pirofosfato do contato com uma polimerase que está a catalisar a pirofosforólise, seguida por um passo de retoma da produção sequencial dos trifosfatos nucleótidos contactando a polimerase com pirofosfato. O pirofosfato pode ser fornecido a uma reação usando técnicas adequadas para o sistema de fluidos a ser usado, incluindo, por exemplo, a pipetagem de alíquotas de fluidos, o movimento de pílulas grandes de fluido sob pressão positiva ou negativa (por exemplo, através de bombas ou gravidade), eletroforese, isotacoforese, manipulação de gotículas (por exemplo, electrowetting) ou similares. As técnicas fluídicas similares podem ser usadas para remover o pirofosfato, por exemplo, por deslocamento e/ou substituição com uma solução de lavagem. Naturalmente, tais técnicas fluídicas podem ser usadas para adicionar ou remover outros componentes usados nos métodos e aparelhos aqui estabelecidos.

Alternativamente, ou adicionalmente aos métodos fluídicos expostos acima, o pirofosfato pode ser removido da reação por sequestro, degradação ou inativação. Por exemplo, podem ser usadas manipulações físicas, como adsorção a um agente sequestrador ou degradação por calor, luz ou eletricidade.

Os métodos químicos podem ser usados para modificar a estrutura ou atividade do pirofosfato ou para degradar a molécula. Os métodos enzimáticos também podem ser usados, tais como a degradação por pirofosfatase, como mostrado para as reações de PCR nos documentos US 4,800,159 e US 5,498,523 e para reações de sequenciamento baseadas em gel no documento US 4,962,020.

Alternativamente, ou adicionalmente às técnicas acima descritas, a pirofosforólise pode ser parada ou pausada pela remoção de outros componentes da reação. Por exemplo, a polimerase pode ser removida de uma reação e opcionalmente substituída ou retomada num estado ativo. Por exemplo, a polimerase pode ser removida por métodos fluídicos, de sequestro, de degradação ou de inativação, tais como aqueles exemplificados acima para o pirofosfato. Em formas de realização particulares, uma polimerase sensível ao calor (não termófila) pode ser usada numa reação de pirofosforólise e depois inativada por calor. De modo semelhante, uma polimerase pode ser degradada por modificação química ou degradação enzimática (por exemplo, através de uma protease). Seja degradada por técnicas físicas, químicas ou enzimáticas, a polimerase gastada pode ser lavada e, desse modo, a pirofosforólise pode ser retomada por adição de uma nova polimerase ao ácido nucleico sendo sequenciado.

A atividade da polimerase também pode ser modificada pela adição e remoção de inibidores, alternando entre temperaturas permissivas e não permissivas para as polimerases termicamente estáveis, ou a presença e ausência de um agente sequestrador ou substrato competitivo. A pirofosforólise também pode ser interrompida e iniciada

por desnaturação e renaturação, respetivamente, do ácido nucleico que está sendo sequenciado.

Embora os métodos e aparelhos tenham sido exemplificados aqui para formas de realização que usam pirofosfato para conduzir a pirofosforólise, será entendido que os análogos de pirofosfato podem ser usados em vez disso. Um análogo exemplar é o pirovanadato, que pode ser usado, por exemplo, como descrito em Akabayov et al. *J. Biol. Chem.* 286:29146-29157 (2011).

Como outros exemplos, podem ser usados análogos de pirofosfato com porções adicionais. Geralmente, são selecionados análogos de pirofosfato que não inibem inteiramente a pirofosforólise ou a passagem dos trifosfatos nucleótidos resultantes, ou seus análogos, através de um nanoporo. No entanto, os análogos de pirofosfatos podem alterar as características da pirofosforólise e/ou a detecção do nanoporo. Por exemplo, pode ser benéfico usar um análogo de pirofosfato para abrandar ou acelerar a pirofosforólise para fornecer uma taxa de detecção desejada ou ideal. Da mesma forma, os análogos de trifosfatos nucleótidos que resultam quando um análogo de pirofosfato é usado numa reação da pirofosforólise também podem fornecer características desejadas para detecção de nanoporos. Por exemplo, as porções que alteram a carga ou o tamanho, em comparação com o difosfato sozinho, podem aumentar ou diminuir a taxa de passagem de análogos trifosfatos nucleótidos através de um nanoporo, ou de outra forma alterarem as interações dos análogos trifosfatos nucleótidos com o nanoporo, para fornecer resultados de sequenciamento melhorados.

No entanto, em algumas formas de realização, um método ou aparelho da presente divulgação irá excluir o pirofosfato com quaisquer porções adicionadas. Por exemplo, o pirofosfato que não possui uma porção detetável, opticamente, tal como uma porção fluorescente, pode ser usado.

Os trifosfatos nucleótidos são detetados usando nanoporos. Os trifosfatos nucleótidos que são sequencialmente removidos de um ácido nucleico via pirofosforólise podem ser passados através de um nanoporo para detecção.

Através do uso de um nanoporo apropriado, diferentes porções de base dos trifosfatos nucleótidos podem ser distinguidas para permitir a detecção de sequenciamento. Geralmente, pode ser usado um aparelho que inclui um primeiro e um segundo compartimento, separados por uma barreira física, em que a barreira tem um ou mais nanoporos. O primeiro compartimento pode incluir componentes usados para uma reação de pirofosforólise. O aparelho pode ser configurado para aplicar um campo elétrico através da barreira de modo que os trifosfatos nucleótidos sejam movidos do primeiro compartimento através do poro para o segundo compartimento. O aparelho pode ser configurado para medir a assinatura eletrônica de um trifosfato nucleótido passando através do nanoporo. Consequentemente, um aparelho útil pode incluir um circuito elétrico capaz de aplicar um potencial e medir um sinal elétrico através de uma barreira e de um nanoporo. Os métodos podem ser realizados usando um patch-clamp ou um grampeamento de voltagem.

Um método da presente divulgação pode ser realizado usando qualquer sistema adequado no qual um poro penetra através de uma barreira. A barreira em muitas formas de realização é preferencialmente uma bicamada lipídica. As bicamadas lipídicas podem ser feitas usando métodos conhecidos na técnica, por exemplo, como descrito em Montai e Mueller *Proc. Natl. Acad. Sci. USA* 69:3561-3566 (1972) ou WO 2008/102120.

As bicamadas lipídicas podem ser formadas a partir de qualquer uma de uma variedade de lípidos incluindo, mas não se limitando a, fosfolípidos, glicolípidos, colesterol e suas misturas. Os nanoporos exemplares que podem ser usados incluem, por exemplo, nanoporos à base de proteínas tais como o nanoporo alfa de hemolisina, *mycobacterium smegmatis* porin A (MspA) e suas variantes. O nanoporo de emolisina alfa e as variantes do nanoporo nativo que são particularmente úteis, são descritas, por exemplo, na Publicação da Aplicação da Patente US N° 2011/0229877 A1, ou nas Patentes US N°s. 6,916,665; 7,867,716; 7,947,454; ou 8,105,846. A MspA e as variantes do nanoporo nativo que são particularmente úteis, são descritas, por exemplo, na Publicação da Aplicação da Patente US N° 2012/0055792 A1. Os nanoporos em estado sólido também podem ser úteis, incluindo, por exemplo, os descritos nas Patentes. US N°s 6,413,792; 7,444,053; ou 7,582,490.

A detecção de trifosfatos de nucleotídeos pode explorar a interação com um nanoporo que resulta em mudanças na corrente que flui através do nanoporo de uma maneira específica para cada espécie de trifosfato nucleótido. Por exemplo, uma primeira espécie de trifosfato nucleótido pode reduzir a corrente que flui através do nanoporo para um

período de tempo médio específico ou para uma extensão particular. Uma segunda espécie de trifosfato nucleótido pode ser distinguida em virtude de um período de tempo médio diferente ou de uma extensão diferente da alteração atual ao passar pelo nanoporo. Assim, diferentes espécies de trifosfato nucleótido podem ser distinguidas com base em alterações distintivas da corrente que flui através de um nanoporo.

A detecção de nanoporos pode ser realizada usando qualquer um de uma variedade de aparelhos conhecidos na técnica incluindo, por exemplo, os descritos na Publicação da Aplicação da Patente US N°s. 2011/0229877 A1; ou 2012/0055792 A1; ou Patentes US N°s. 6,413,792; 6,916,665; 7,867,716; 7,444,053; 7,582,490; 7,947,454; ou 8,105,846.

Uma polimerase que é usada num aparelho ou método aqui apresentado pode estar presente em solução de modo que seja relativamente livre para difundir, pelo menos dentro de uma câmara de reação ou pode ser relativamente limitada na sua capacidade de difusão por estar ligada a um suporte de fase sólida, nanoporo, barreira ou outro componente de um método ou aparelho aqui apresentado. A limitação de difusão por anexação pode fornecer uma vantagem de acoplar de perto o ponto da produção de trifosfato nucleótido (por exemplo, uma pirofosforossólise catalisadora de polimerase) com o ponto de detecção de trifosfato nucleótido (por exemplo, um nanoporo através do qual os trifosfatos nucleótidos passam).

Uma polimerase pode ser anexada a um nanoporo, por exemplo, através de uma fusão de proteína recombinante para uma

subunidade de um nanoporo, uma reticulação química ou uma porção de adaptador. Os métodos úteis para anexar as polimerases a nanoporos e componentes de polimerase-nanoporo são apresentados, por exemplo, na Publicação da Aplicação da Patente US N°s. 2011/0229877 A1; ou 2012/0055792 A1; ou na Patente US N°. 7,947,454.

Uma polimerase pode ser anexada a um grânulo ou outro suporte sólido que reside numa câmara onde ocorre a pirofosforólise. Os ligadores químicos que são úteis para ligar as polimerases a grânulos ou suportes sólidos incluem aqueles que estão comercialmente disponíveis, por exemplo, em Thermo Fisher (Rockford, IL) ou Sigma Aldrich (St. Louis, MO) ou de outra forma conhecidos na técnica.

Uma polimerase pode ser anexada a uma barreira usada num aparelho de sequenciamento de nanoporos. Por exemplo, em formas de realização que usam uma bicamada lipídica como barreira, uma porção lipofílica pode ser anexada à polimerase para localizar a polimerase na proximidade da bicamada devido a interações entre a bicamada e a porção lipofílica.

Os exemplos de porções lipofílicas incluem, mas não estão limitados a esteróis ou lipídios. Um outro exemplo de uma porção lipofílica é uma proteína de membrana (ou sua porção) que pode ser anexada a uma polimerase, por exemplo, por meio de fusão de proteína recombinante. As ligações como as acima descritas para grânulos e outros suportes sólidos podem ser usados para anexar uma polimerase a uma barreira usada em sistemas de nanoporos de estado sólido.

Um ácido nucleico que é sequenciado num método aqui apresentado ou presente num aparelho da presente divulgação pode estar em solução de modo que seja relativamente livre para difundir ou pode ser relativamente limitado na sua capacidade de difusão por estar anexado a um suporte de fase sólida, nanoporo, barreira ou outro componente de um método ou aparelho aqui apresentado. Os anexos semelhantes aos estabelecidos acima para polimerases podem ser usados para ácidos nucleicos. Por exemplo, um esterol, lípido ou outra porção lipofílica pode ser anexado a um ácido nucleico para o localizar para uma bicamada lipídica. Um exemplo é mostrado na FIG. 2, em que o ácido nucleico é localizado na membrana por meio de uma porção de esterol anexada à cadeia de molde. Tal como exemplificado pela figura, o ácido nucleico pode ser ligado através da cadeia modelo, por exemplo, na extremidade 5' da cadeia modelo. O anexo também pode ser feito num ponto no modelo que está entre o local onde a polimerase está ligada ao modelo e a extremidade 5' do modelo.

Uma porção lipofílica pode ser anexada a um ácido nucleico usando métodos conhecidos na técnica para anexar outras porções tais como a biotina ou os fluoróforos. Por exemplo, um iniciador com a porção lipofílica pode ser usado numa amplificação, extensão do iniciador ou reação de ligação. Alternativa ou adicionalmente, os nucleótidos que possuem a porção podem ser usados numa reação de extensão ou de amplificação. Se desejado, uma porção lipofílica pode ser introduzida antes do sequenciamento e durante um passo de preparação da amostra, tal como os apresentados aqui anteriormente. Por exemplo, pode ser empregue uma técnica de sequenciamento alvo em que um subconjunto de ácido nucleico alvo com sequências desejadas deve ser selecionado

de um fundo de sequenciamento mais complexo. Neste exemplo, uma porção lipofílica pode ser introduzida seletivamente no subconjunto de ácidos nucleicos alvo usando a técnica de direcionamento e isso pode permitir que os alvos sejam capturados seletivamente por uma bicamada lipídica, enquanto que outras sequências não direcionadas são lavadas devido a não terem sido modificadas para incluir a porção lipofílica.

Pode ser benéfico em algumas formas de realização limitar a difusão, ambas da polimerase e do ácido nucleico, em relação a um nanoporo, por exemplo, usando um ou mais dos meios de anexação estabelecidos acima.

Conforme estabelecido aqui anteriormente, um método da presente divulgação inclui um passo de contato com um ácido nucleico alvo com uma polimerase em condições para remover sequencialmente nucleótidos, produzindo sequencialmente trifosfatos nucleotídeos com uma variedade de diferentes porções de base. A variedade de diferentes porções de base produzidas dependerá do conteúdo do ácido nucleico alvo que é contactado com a polimerase. Por exemplo, o DNA inclui, normalmente, as quatro bases comuns de guanina, citosina, adenina e timina de tal modo que a pirofosforólise irá produzir desoxiguanidina trifosfato, desoxicitidina trifosfato, desoxiadenosina trifosfato e desoxitimidina trifosfato. Em alguns casos, o ácido nucleico alvo pode não incluir os quatro desses tipos de base, de modo que não mais do que 3, 2 ou mesmo 1 tipo de trifosfato desoxinucleótido seja produzido pela pirofosforólise. Em alguns casos, as variantes de um ou mais desses quatro tipos de bases podem estar presentes no DNA alvo e, conseqüentemente, a pirofosforólise pode produzir

trifosfatos desoxinucleótidos variantes, tendo, por exemplo, metilo, hidroximetilo ou outras porções adicionadas. Outras bases de variantes conhecidas na técnica tais como aquelas apresentadas aqui, também podem estar presentes nos trifosfatos desoxinucleótidos produzidos pela pirofosforólise.

Outro exemplo é o RNA, que tipicamente inclui as quatro bases comuns de guanina, citosina, adenina e uracilo, de modo que a pirofosforólise produzirá trifosfato de riboguanidina, trifosfato de ribocitidina, trifosfato de riboadenosina e trifosfato de ribotimidina. Em alguns casos, o ácido nucleico alvo pode não incluir os quatro desses tipos de base, de modo que não mais do que 3, 2 ou mesmo 1 tipo de ribonucleótido trifosfato sejam produzidos pela pirofosforólise. As bases de variante, tais como aquelas aqui exemplificadas, por exemplo, com respeito aos desoxinucleótidos trifosfatos, ou conhecidos de outra forma na técnica, podem estar presentes nos ribossucleótidos trifosfatos. Geralmente, os trifosfatos nucleótidos produzidos pela pirofosforólise (ambos o desoxinucleótidos trifosfatos ou ribonucleótidos trifosfatos) incluirão um, pelo menos dois, pelo menos três, pelo menos quatro ou mais diferentes tipos de base.

Um método da presente divulgação pode ser realizado sob condições que removem sequencialmente uma série de nucleótidos a partir de um ácido nucleico alvo, produzindo sequencialmente o mesmo número de trifosfatos nucleótidos. Além disso, pelo menos o mesmo número de trifosfatos nucleótidos pode ser distinguido, por exemplo, através da passagem através de um nanoporo, para permitir a determinação de uma sequência com um comprimento pelo menos

equivalente ao número de nucleótidos removidos do ácido nucleico alvo. Em formas de realização particulares, o número é pelo menos 1, 2, 3, 4, 5, 10, 25, 50, 100, 200, 500, 1000, 10 000 ou mais até e incluindo o comprimento do ácido nucleico alvo. Alternativa, ou adicionalmente, o número não pode ser superior a 1, 2, 3, 4, 5, 10, 25, 50, 100, 200, 500, 1000 ou 10 000. O número pode ser, mas não precisa ser, entre dois desses valores. Conforme aqui estabelecido anteriormente, uma variedade de técnicas pode ser usada para pausar a pirofosforólise. Isto pode fornecer o controlo do comprimento da sequência determinada usando formas de realização dos presentes métodos.

O número de trifosfatos nucleótidos libertados pela pirofosforólise e/ou detetados num método aqui apresentado pode ser maior que o número de diferentes tipos de trifosfatos nucleótidos detetados. No entanto, a ordem e o número dos diferentes trifosfatos nucleótidos detetados podem ser correlacionados com a sequência do ácido nucleico.

Em algumas formas de realização, pode ser benéfico sequenciar repetidamente um ácido nucleico alvo particular. A repetição pode ser alcançada, por exemplo, ao processar repetidamente uma molécula de ácido nucleico alvo num método aqui apresentado. Por exemplo, um método pode incluir as etapas de (a) contactar um ácido nucleico alvo com uma polimerase para remover sequencialmente trifosfatos nucleótidos do ácido nucleico alvo, em que os trifosfatos nucleótidos que são removidos possuem uma variedade de diferentes porções de base; (b) distinguir as diferentes porções de base para os trifosfatos nucleótidos que são removidos, determinando assim a sequência do ácido nucleico

alvo; (c) regenerar pelo menos uma porção do ácido nucleico alvo; e as etapas repetitivas (a) e (b) usando o ácido nucleico alvo regenerado. O ácido nucleico alvo pode ser regenerado, por exemplo, adicionando trifosfatos nucleótidos sob condições para a polimerase (ou uma polimerase recém-adicionada) para realizar uma reação de polimerização para regenerar pelo menos uma porção de uma cadeia do ácido nucleico alvo que foi previamente removida pela pirofosforólise. Tipicamente, o pirofosfato estará, substancialmente, ausente durante a reação de polimerização.

Alternativa, ou adicionalmente à forma de realização de processamento repetida acima, um ácido nucleico alvo pode ser amplificado ou copiado para criar cópias múltiplas que são processadas usando um método da presente divulgação. Um exemplo esquemático é mostrado na FIG. 3. Múltiplas cópias de um ácido nucleico modelo de cadeia dupla são localizadas para uma barreira numa câmara com uma fusão de nanoporo polimerase (passo 1), uma das cadeias é capturada pela polimerase (passo 2), ocorre um sequenciamento baseado na pirofosforólise (passo 3), a cadeia modelo, sendo uma cadeia simples, pode então ser puxada através do nanoporo através da força elétrica (passo 4) até que seja removida da câmara onde as outras cópias do ácido nucleico permanecem (n.b. as outras cópias permanecem devido a serem de cadeia dupla e, portanto, resistentes à passagem através do nanoporo) (passo 5) e, em seguida, é capturada outra cópia do ácido nucleico modelo de cadeia dupla para iniciar a repetição dos passos 2 *et seq.*

Qualquer um de uma variedade de métodos conhecidos na técnica para amplificar ácidos nucleicos, tais como os

apresentados aqui anteriormente, pode ser usado para criar as cópias múltiplas do ácido nucleico alvo.

Embora o sistema da FIG. 3 seja exemplificado para cópias de um único modelo, será entendido que as espécies de ácido nucleico com sequências diferentes podem ser usadas de forma semelhante. Assim, uma variedade de diferentes espécies de ácidos nucleicos de cadeia dupla pode ser localizada numa barreira numa câmara com uma fusão de nanoporo-polimerase (passo 1), uma cadeia de uma primeira espécie pode ser capturada pela polimerase (passo 2), pode ocorrer um sequenciamento baseado na pirofosforólise (passo 3), a cadeia modelo da primeira cadeia pode ser puxada através do nanoporo através da força elétrica (passo 4) até que seja removida da câmara onde permanecem as outras espécies de ácido nucleico (passo 5) e então outras espécies de ácidos nucleicos de cadeia dupla podem ser capturadas para iniciar a repetição dos passos 2 *et seq.*

A presente divulgação também fornece um aparelho que inclui (a) uma barreira impermeável a fluidos que separa um primeiro reservatório de fluido de um segundo reservatório de fluido; (b) um nanoporo posicionado na barreira impermeável a fluidos para formar uma passagem através da qual um trifosfato nucleótido pode passar do primeiro reservatório de fluido para o segundo reservatório de fluido; e (c) uma mistura de reação no primeiro reservatório de fluido, a mistura de reação incluindo uma polimerase, ácido nucleico alvo tendo duas cadeias e uma concentração pirofosforolítica de pirofosfato. Os componentes usados no aparelho podem ser um ou mais dos exemplificados acima no contexto de vários métodos.

Componentes adicionais e configurações são exemplificados abaixo para fins ilustrativos.

Uma barreira impermeável a fluidos pode ser configurada para separar dois reservatórios e ter um nanoporo colocado na barreira para fornecer uma conexão fluida entre os reservatórios. Os exemplos de nanoporos e barreiras são apresentados acima e em várias referências apresentadas aqui, acima. Geralmente, os dois reservatórios estarão em comunicação fluida através de um único nanoporo. Assim, os trifosfatos nucleótidos produzidos num dos reservatórios terá um e apenas um caminho de fluido para o segundo reservatório. O uso de um único nanoporo dessa forma permite uma medição conveniente de cada trifosfato nucleótido que passa de um reservatório para o outro devido a mudanças nas propriedades elétricas no nanoporo, barreira e/ou reservatórios. No entanto, também é possível, em algumas formas de realização, incluir mais de que um nanoporo na barreira que separa dois reservatórios. Quando nanoporos múltiplos conectam de forma fluida dois reservatórios, a passagem dos trifosfatos nucleótidos pode ser medida no nanoporo individual usando, por exemplo, medições óticas ou elétricas que resolvem cada nanoporo.

Um reservatório pode criar uma câmara onde o fluido permanece contido por pelo menos uma parte do tempo. Por exemplo, uma câmara pode ser configurada para formar um poço, cavidade, compartimento, etc., que restringe o fluxo de fluido. Alternativamente, um reservatório pode ser configurado para fluxo de fluido. Por exemplo, o reservatório pode ser configurado como um tubo, canal ou célula de fluxo, permitindo assim o fluxo de fluidos para entrega e remoção convenientes de componentes usados num

método de sequenciamento. Em formas de realização particulares, um primeiro reservatório que contém ácido nucleico modelo, polimerase e pirofosfato será configurado para o fluxo de fluido, enquanto que o segundo reservatório, que está conectado à primeira câmara através de um nanoporo, pode ser configurado como uma câmara. O segundo reservatório não precisa ser configurado para o fluxo de fluido, mas pode ser, opcionalmente.

A presente divulgação fornece formas de realização multiplex. Por exemplo, as sequências para uma pluralidade de ácidos nucleicos alvo podem ser determinadas em paralelo. Um método multiplex pode incluir as etapas de (a) fornecer uma pluralidade de ácidos nucleicos alvo; (b) contactar cada um dos ácidos nucleicos alvo com uma polimerase para remover, sequencialmente, trifosfatos nucleótidos de cada ácido nucleico alvo, em que os trifosfatos nucleótidos que são removidos possuem uma variedade de diferentes porções de base; e (c) distinguir as diferentes porções de base para os trifosfatos nucleótidos que são removidos de cada ácido nucleico, determinando assim as sequências dos ácidos nucleicos alvo.

Um outro exemplo de um método multiplex é aquele que inclui os passos de (a) fornecer uma pluralidade de ácidos nucleicos alvo tendo cada um duas cadeias; (b) contactar cada um dos ácidos nucleicos alvo com uma polimerase em condições para remover, sequencialmente, os nucleótidos da primeira de cada uma das duas cadeias pela pirofosforólise, produzindo, sequencialmente, trifosfatos nucleotídeos com uma variedade de diferentes porções de base; e (c) distinguir as diferentes porções de base para os

trifosfatos nucleótidos produzidos sequencialmente, determinando assim a sequência dos ácidos nucleicos alvo.

Um aparelho multiplex pode incluir (a) uma pluralidade de barreiras impermeáveis a fluidos que cada uma separa um primeiro reservatório de fluido de um segundo reservatório de fluido; (b) um nanoporo posicionado em cada uma das barreiras impermeáveis a fluidos para formar uma passagem através da qual um trifosfato nucleótido pode passar do primeiro reservatório de fluido para o segundo reservatório de fluido; e (c) uma mistura de reação em cada um dos primeiros reservatórios de fluido, cada uma das misturas de reação incluindo uma polimerase, ácido nucleico alvo tendo duas cadeias e uma concentração pirofosforolítica de pirofosfato.

A plexidade (isto é, o nível de multiplexagem) de um método ou aparelho pode ser selecionada para satisfazer um uso particular. Por exemplo, o número de ácidos nucleicos alvo que são processados ou presentes em conjunto pode ser determinado a partir da complexidade da amostra a ser avaliada. As estimativas de complexidade exemplares para alguns dos genomas que podem ser avaliados usando métodos ou aparelhos da presente divulgação são cerca de 3.1 Gbp (humano), 2.7 Gbp (rato), 2.8 Gbp (ratazana), 1.7 Gbp (peixe zebra), 165 Mbp (mosca da fruta), 13.5 Mbp (*S. cerevisiae*), 390 Mbp (fugu), 278 Mbp (mosquito) ou 103 Mbp (*C. elegans*). Os especialistas na técnica reconhecerão que os genomas que possuem tamanhos diferentes dos exemplificados acima, incluindo, por exemplo, genomas menores ou maiores, podem ser usados num método da invenção. Tipicamente, uma amostra de ácido nucleico é fragmentada antes do uso.

O número de fragmentos a serem manipulados em paralelo dependerá da complexidade do genoma, do tamanho médio do fragmento e da cobertura desejada. Por exemplo, a cobertura 1x de um genoma humano (3.1 Gbp) que é fragmentada para um tamanho médio de 1000 nucleótidos pode ser alcançada usando uma plexidade de 3 milhões de fragmentos (isto é $((3.1 \text{ bilhões}/1000) \times 1) = 1 \text{ milhão}$). Usando cálculos semelhantes, pode-se determinar que uma plexidade de 30 milhões de fragmentos (de 1000 nucleótidos cada) é suficiente para fornecer uma cobertura de 30x de um genoma humano.

Os métodos e aparelhos aqui estabelecidos podem ser configurados a um nível de plexidade suficiente para fornecer pelo menos 1x, 2x, 5x, 10x, 20x, 30x, 50x ou mais cobertura de qualquer uma variedade de genomas, incluindo, mas não limitado aqueles exemplificados aqui. A plexidade pode ser uma função do número de vários componentes aqui apresentados, tal como o número de fragmentos de ácido nucleico alvo como exemplificado acima. Outros componentes que podem ser multiplexados incluem, por exemplo, o número de nanoporos usados, o número de polimerases, o número de câmaras que possuem uma barreira e nanoporo, etc. O nível de multiplexação destes ou de outros componentes, pode ser, pelo menos, 2, 5, 10, 100, 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^9 ou superior. Alternativa, ou adicionalmente, o nível de multiplexagem pode ser selecionado para não ser superior a 2, 5, 10, 100, 1×10^3 , 1×10^4 , 1×10^5 , 1×10^6 , 1×10^9 .

O termo "compreender" está destinado aqui a ser aberto, incluindo não apenas os elementos recitados, mas abrangendo ainda outros elementos adicionais.

Embora a invenção tenha sido descrita com referência aos exemplos fornecidos acima, deve entender-se que podem ser feitas várias modificações sem se afastar da invenção. Consequentemente, a invenção é limitada apenas pelas reivindicações.

Lisboa, 26 de Fevereiro de 2018

REFERÊNCIAS CITADAS NA DESCRIÇÃO

Esta lista de referências citadas pelo requerente é apenas para a conveniência do leitor. A mesma não faz parte do documento de Patente Europeia. Embora muito cuidado tenha sido tomado na compilação das referências, erros e omissões não podem ser excluídos e o EPO nega qualquer responsabilidade neste sentido.

Documentos de Patente citados na descrição

WO 2003080861 A1
US 6268146 B1
US 6107023 A
US 7741463 B
US 20100120098 A1
US 8071356 B
US 7090975 B
US 7914995 B
US 4800159 A
US 5498523 A
US 4962020 A1
WO 2008102120 A
US 20110229877 A1
US 6916665 B
US 7867716 B
US 7947454 B
US 8105846 B
US 20120055792 A1
US 6413792 B
US 7444053 B
US 7582490 B

Documentos que não são patente citados na descrição

CLARKE et al. Nat. Nanotechnol., 2009, vol. 4 (4), 265-70

CHELEY et al. Chem. Biol., 2002, vol. 9, 829-838

CHERF et al. Nat. Biotech., 2012, vol. 30, 344-348

MANRAO et al. Nat. Biotech., 2012, vol. 30, 349-353

SAMBROOK et al. Molecular Cloning: A Laboratory Manual.
Cold Spring Harbor Laboratory, 2001

AUSUBEL et al. Current Protocols in Molecular Biology. John
Wiley and Sons, 1998

PATEL et al. Biochem., 1991, vol. 30, 511-525

AKABAYOV et al. J. Biol. Chem., 2011, vol. 286, 29146-29157

MONTAL; MUELLER. Proc. Natl. Acad. Sci. USA, 1972, vol. 69,
3561-3566

Reivindicações

1. Um método para determinar a sequência de um ácido nucleico alvo, compreendendo;

(a) fornecer um ácido nucleico alvo possuindo duas cadeias;

(b) contactar o ácido nucleico alvo com uma polimerase em condições para remover, sequencialmente, nucleótidos da primeira das duas cadeias pela pirofosforólise, produzindo, sequencialmente, trifosfatos nucleótidos possuindo uma variedade de diferentes porções de base; e

(c) distinguir as diferentes porções de base para os trifosfatos nucleótidos produzidos sequencialmente, determinando assim a sequência do ácido nucleico alvo; em que a distinção das diferentes porções de base para os trifosfatos nucleótidos produzidos, sequencialmente, compreende a passagem dos trifosfatos nucleótidos através de um nanoporo.

2. O método da reivindicação 1, em que a polimerase está anexada ao nanoporo.

3. O método da reivindicação 1, em que as condições para remover sequencialmente os nucleótidos de uma das duas cadeias pela pirofosforólise compreendem o contacto da polimerase com uma concentração pirofosforolítica de pirofosfato.

4. O método da reivindicação 3, em que o método compreende ainda um passo de pausar a produção sequencial dos trifosfatos nucleótidos, removendo o pirofosfato do contacto com a polimerase e depois retomando a produção

sequencial dos trifosfatos nucleótidos, contactando a polimerase com o pirofosfato.

5. O método da reivindicação 1, em que a variedade de diferentes porções de base compreende pelo menos duas espécies diferentes de porções de base e no máximo quatro espécies diferentes de porções de base.

6. O método da reivindicação 5, em que a sequência que é determinada é superior a quatro nucleótidos.

7. O método da reivindicação 1, em que as porções de base compreendem adenina, guanina, citosina ou timina de ocorrência natural.

8. O método da reivindicação 1, em que pelo menos uma das porções de base compreende uma porção que não ocorre naturalmente no DNA ou RNA.

9. Um aparelho, que compreende;

(a) uma barreira impermeável a fluidos que separa um primeiro reservatório de fluido de um segundo reservatório de fluido;

(b) um nanoporo posicionado na barreira impermeável a fluidos para formar uma passagem através da qual um trifosfato nucleótido pode passar do primeiro reservatório de fluido para o segundo reservatório de fluido; e

(c) uma mistura de reação no primeiro reservatório de fluido, a mistura de reação compreendendo uma polimerase,