

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 768 762**

51 Int. Cl.:

C12N 15/10 (2006.01)

C12Q 1/68 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.10.2015 PCT/US2015/057409**

87 Fecha y número de publicación internacional: **12.05.2016 WO16073237**

96 Fecha de presentación y número de la solicitud europea: **26.10.2015 E 15790422 (8)**

97 Fecha y número de publicación de la concesión europea: **18.12.2019 EP 3215616**

54 Título: **Reducción del daño al ADN durante la preparación de muestras y secuenciación usando agentes quelantes sideróforos**

30 Prioridad:

05.11.2014 US 201462075708 P

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

23.06.2020

73 Titular/es:

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

72 Inventor/es:

**SMITH, VINCENT, PETER y
DE ROZIERES, SOHELA**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 768 762 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

Reducción del daño al ADN durante la preparación de muestras y secuenciación usando agentes quelantes sideróforos

Campo

5 Las realizaciones proporcionadas en la presente memoria se refieren a métodos y composiciones para la secuenciación de nueva generación. Algunas realizaciones se refieren a reactivos para la preparación de muestras relacionadas con la secuenciación que comprenden un sideróforo y un electrolito.

Antecedentes

10 Los agentes quelantes se usan en bioquímica para eliminar o inactivar cationes divalentes inactivos que pueden conducir a efectos perjudiciales y no deseados en la reacción. El uso actual de EDTA como un agente quelante en aplicaciones de preparación de muestras relacionadas con la secuenciación, puede conducir a oxidación importante e irreversible del ADN, en especial en presencia de hierro contaminante, debido a la generación de un complejo de Fe^{2+} (EDTA) que promueve la química de radicales libres perjudicial (Lloyd y Phillips, *Mutation Res.* (1999) 424:23-36; Wang et al., *Nuc. Acid Res.* (2008) 36:e85). Esto da como resultado mutaciones puntuales locales en la muestra de ADN que conducen a casos de polimorfismo erróneo durante el análisis de secuenciación profunda. Este fenotipo es especialmente problemático en aplicaciones de secuenciación profunda relacionadas con el cáncer, donde se deben detectar polimorfismos de baja frecuencia de modo sensible y son de gran importancia biológica.

15 Mike S. Son et al, en "Preparing DNA Libraries for Multiplexed Paired-End Deep Sequencing for Illumina GA Sequencers", *Current Protocols in Microbiology*, Hoboken, NJ, USA, John Wiley & Sons, Inc., (2011-02-01), describen un método para preparar una biblioteca de ácidos nucleicos.

20 Resumen de realizaciones de la invención

Las realizaciones descritas en la presente memoria proporcionan reactivos para la preparación de ácidos nucleicos que comprenden un sideróforo. El sideróforo es una sal de mesilato de desferrioxamina B (DFO-B). En algunas realizaciones, los reactivos pueden comprender una ADN polimerasa. En algunas realizaciones, los reactivos pueden comprender dNTP. Los reactivos no contienen EDTA.

25 Las realizaciones descritas en la presente memoria proporcionan métodos para preparar una biblioteca de ácidos nucleicos, que comprenden: proporcionar una pluralidad de moléculas de ácido nucleico a partir de una muestra; y manipular la pluralidad de moléculas de ácido nucleico en un reactivo para preparar ácidos nucleicos que comprende una sal de mesilato de desferrioxamina B (DFO-B) y el reactivo no contiene EDTA. En algunas realizaciones, la manipulación de la pluralidad de moléculas de ácido nucleico comprende hibridar la pluralidad de moléculas de ácido nucleico con una pluralidad de sondas de oligonucleótidos. En algunas realizaciones, la pluralidad de moléculas de ácido nucleico y/o la pluralidad de sondas de oligonucleótidos se inmovilizan sobre un soporte. En algunas realizaciones, la pluralidad de moléculas de ácido nucleico y/o la pluralidad de sondas de oligonucleótidos se inmovilizan sobre un soporte a través de una pareja de unión al soporte. En algunas realizaciones, el soporte es una perla magnética. En algunas realizaciones, la manipulación de la pluralidad de moléculas de ácido nucleico comprende separar las sondas de oligonucleótidos no unidas específicamente a la pluralidad de moléculas de ácido nucleico. En algunas realizaciones, los métodos comprenden modificar las sondas de oligonucleótidos específicamente unidas a la pluralidad de moléculas de ácido nucleico. En algunas realizaciones, los métodos comprenden fragmentar la pluralidad de moléculas de ácido nucleico. En algunas realizaciones, los métodos comprenden añadir adaptadores a la pluralidad de moléculas de ácido nucleico. En algunas realizaciones, los adaptadores se añaden a la pluralidad de moléculas de ácido nucleico por amplificación.

35 Las realizaciones descritas en la presente memoria proporcionan el uso de métodos para reducir el daño oxidativo a una molécula de ácido nucleico, cuyos métodos comprenden preparar la molécula de ácido nucleico en ausencia de EDTA. Preparar la molécula de ácido nucleico comprende preparar la molécula de ácido nucleico en presencia de una sal de mesilato de desferrioxamina B (DFO-B). En algunas realizaciones, preparar la molécula de ácido nucleico comprende exponer la molécula de ácido nucleico a $Fe(III)$. En algunas realizaciones, preparar la molécula de ácido nucleico comprende exponer la molécula de ácido nucleico a una perla magnética. En algunas realizaciones, el daño oxidativo comprende una mutación puntual en la molécula de ácido nucleico. En algunas realizaciones, la mutación puntual es una transversión de C a A.

45 Las realizaciones descritas en la presente memoria proporcionan el uso de métodos para aumentar la puntuación Q (Phred) de una reacción de secuenciación, cuyos métodos comprenden preparar una molécula de ácido nucleico en ausencia de EDTA. Preparar la molécula de ácido nucleico comprende preparar la molécula de ácido nucleico en presencia de una sal de mesilato de desferrioxamina B (DFO-B). En algunas realizaciones, la puntuación Q es mayor de aproximadamente 34. En algunas realizaciones, la puntuación Q es mayor de aproximadamente 38. En algunas realizaciones, la puntuación Q es mayor de aproximadamente 42. En algunas realizaciones, la reacción de secuenciación es una aplicación de secuenciación profunda. En algunas realizaciones, la aplicación de secuenciación profunda es una aplicación de secuenciación profunda relacionada con el cáncer. En algunas realizaciones, los métodos comprenden secuenciar la molécula de ácido nucleico en ausencia de EDTA.

Breve descripción de los dibujos

La figura 1 muestra los resultados de una realización de ejemplo de un estudio comparativo de puntuaciones Q de experimentos de secuenciación que usan un tampón con y sin EDTA, EGTA o DTPA.

Descripción detallada de realizaciones de la invención

5 Definiciones

Como se usa en la presente memoria, las formas singulares "un", "una" y "el", "la" incluyen referencias plurales salvo que se indique lo contrario, expresamente o por el contexto. Por ejemplo, "un" dímero incluye uno o más dímeros, salvo que se indique lo contrario, expresamente o por el contexto.

10 Los términos "polinucleótido", "oligonucleótido", "ácido nucleico" y "molécula de ácido nucleico" se usan de forma intercambiable en la presente memoria para referirse a una forma polimérica de nucleótidos de cualquier longitud, y puede comprender ribonucleótidos, desoxirribonucleótidos, sus análogos o sus mezclas. Este término se refiere solo a la estructura primaria de la molécula. Por lo tanto, el término incluye ácido desoxirribonucleico ("ADN") de cadena triple, doble y simple, así como ácido ribonucleico ("ARN") de cadena triple, doble y simple. También incluye formas del polinucleótido modificadas, por ejemplo, por alquilación y/o por protección terminal, y no modificadas. Más en particular, los términos "polinucleótido", "oligonucleótido", "ácido nucleico" y "molécula de ácido nucleico" incluyen polidesoxirribonucleótidos (que contienen 2-desoxi-D-ribosa), polirribonucleótidos (que contienen D-ribosa), que incluyen ARNt, ARNr, ARNh y ARNm, sean empalmados o no empalmados, cualquier otro tipo de polinucleótido que es un N- o C-glicósido de una base púrica o pirimidínica, y otros polímeros que contienen cadenas principales no nucleotídicas, por ejemplo, polímeros de poliamida (p. ej., ácidos nucleicos peptídicos ("PNA")) y polimorfolino (disponible en el comercio en Anti-Virals, Inc., Corvallis, OR., como NeuGene®), y otros polímeros de ácido nucleico de secuencia específica sintéticos, con la condición de que los polímeros contienen nucleobases en una configuración que permite el emparejamiento de bases y apilamiento de bases, como se encuentra en el ADN y ARN. Por lo tanto, estos términos incluyen, por ejemplo, 3'-desoxi-2',5'-ADN, oligodesoxirribonucleótido N3' a P5' fosforamidatos, ARN 2'-O-alkil-sustituido, híbridos entre ADN y ARN o entre PNA y ADN o ARN, y también incluyen tipos conocidos de modificaciones, por ejemplo, marcadores, alquilación, "grupos terminales", sustitución de uno o más de los nucleótidos por un análogo, modificaciones internucleótidos tales como, por ejemplo, aquellas con enlaces no cargados (p. ej., metilfosfonatos, fosfotriésteres, fosforamidatos, carbamatos, etc.), con enlaces con carga negativa (p. ej., fosforotioatos, fosforoditioatos, etc.), y con enlaces con carga positiva (p. ej., aminoalquilfosforamidatos, aminoalquilfosfotriésteres), aquellas que contienen restos colgantes, tales como, por ejemplo, proteínas (incluyendo enzimas (p. ej., nucleasas), toxinas, anticuerpos, péptidos señal, poli-L-lisina, etc.), aquellas que tienen intercalantes (p. ej., acridina, psoraleno, etc.), aquellas que contienen quelatos (de, p. ej., metales, metales radiactivos, boro, metales oxidantes, etc.), aquellas que contienen alquilantes, aquellas que tienen enlaces modificados (p. ej., ácidos nucleicos alfa anoméricos, etc.), así como formas no modificadas del polinucleótido u oligonucleótido.

35 "Sonda de ácido nucleico" y "sonda" se usan de forma intercambiable y se refieren a una estructura que comprende un polinucleótido, como se ha definido antes, que contiene una secuencia de ácido nucleico que se puede hibridar con un objetivo correspondiente. Las regiones de polinucleótidos de sondas pueden estar compuestas de ADN, y/o ARN, y/o análogos de nucleótidos sintéticos.

40 En general, la estabilidad de un híbrido es una función de la concentración de iones y la temperatura. Típicamente, una reacción de hibridación se lleva a cabo en condiciones de menor rigurosidad, seguido de lavados de diferente, pero mayor, rigurosidad. La hibridación moderadamente rigurosa se refiere a condiciones que permiten que una molécula de ácido nucleico tal como una sonda se una a una molécula de ácido nucleico complementaria. Las moléculas de ácidos nucleicos hibridadas en general tienen al menos 60% de complementariedad, que incluye por ejemplo al menos cualquiera de 70%, 75%, 80%, 85%, 90%, 95%, 99% o 100% de complementariedad. Las condiciones moderadamente rigurosas son condiciones que equivalen a hibridación en formamida al 50%, 5x solución de Denhardt, 5x SSPE, SDS al 0.2% a 42°C, seguido de lavado en 0.2x SSPE, SDS al 0.2%, a 42°C. Las condiciones de alta rigurosidad se pueden proporcionar, por ejemplo, por hibridación en formamida al 50%, 5x solución de Denhardt, 5x SSPE, SDS al 0.2% a 42°C, seguido de lavado en 0.1x SSPE, y SDS al 0.1% a 65°C.

50 La hibridación de baja rigurosidad se refiere a condiciones que equivalen a hibridación en formamida al 10%, 5x solución de Denhardt, 6x SSPE, SDS al 0.2% a 22°C, seguido de lavado en 1x SSPE, SDS al 0.2%, a 37°C. La solución de Denhardt contiene Ficoll al 1%, polivinilpirrolidona al 1% y albúmina de suero bovino al 1% (BSA). 20x SSPE (cloruro sódico, fosfato sódico, ácido etilendiaminatetraacético (EDTA)) contiene cloruro sódico 3 M, fosfato sódico 0.2 M y (EDTA) 0.025 M. Otros tampones y condiciones de hibridación de rigurosidad moderada y rigurosidad alta adecuados son conocidos por el experto en la técnica.

55 Como se usa en la presente memoria, "muestra biológica" se refiere a cualquier muestra obtenida de una fuente viva o vírica u otra fuente de macromoléculas y biomoléculas, e incluye cualquier tipo de célula o tejido de un sujeto del cual se puede obtener ácido nucleico, proteína u otra macromolécula. La muestra biológica puede ser una muestra obtenida directamente de una fuente biológica o una muestra que está procesada. Por ejemplo, los ácidos nucleicos aislados que se amplifican constituyen una muestra biológica. Las muestras biológicas incluyen, pero no se limitan a

fluidos corporales, tales como sangre, plasma, suero, líquido cefalorraquídeo, líquido sinovial, orina y sudor, muestras de tejidos y órganos de animales y plantas y muestras procesadas derivadas de las mismas. También se incluyen muestras de suelo y agua y otras muestras ambientales, virus, bacterias, hongos, algas, protozoos y sus componentes.

5 Como se usa en la presente memoria, la expresión "se une específicamente" se refiere a la especificidad de unión de una pareja de unión específica. El reconocimiento por un anticuerpo de un objetivo particular en presencia de otros objetivos potenciales es una característica de dicha unión. La unión específica implica dos moléculas diferentes, en donde una de las moléculas se une específicamente a la segunda molécula por medios químicos o físicos. Las dos moléculas están relacionadas en el sentido de que la unión de una con otra es tal que son capaces de distinguir su pareja de unión de otros constituyentes del ensayo que tienen características similares. Los miembros del par de componentes de unión se denominan ligando y receptor (antiligando), miembros del par de unión específica (SBP) y pareja SBP, y similares. Una molécula también puede ser un miembro de SBP para un agregado de moléculas; por ejemplo un anticuerpo generado contra un inmunocomplejo de un segundo anticuerpo y su correspondiente antígeno, se puede considerar que es un miembro de SBP para el inmunocomplejo.

15 Se entiende que aspectos y realizaciones de la invención descritos en la presente memoria, incluyen "que consiste en" y/o "que consiste esencialmente en" aspectos y realizaciones.

Otros objetos, ventajas y características de la presente invención se harán evidentes a partir de la siguiente memoria descriptiva considerada junto con los dibujos que acompañan.

Reactivos para la preparación de ácidos nucleicos

20 Algunas realizaciones descritas en la presente memoria proporcionan reactivos para la preparación de ácidos nucleicos que comprenden un sideróforo,.

25 Los sideróforos son compuestos producidos por bacterias, hongos y plantas gramíneas que se ha sabido que tienen una alta afinidad y selectividad por el Fe(III). Véase, p. ej., Hider y Kong, *Nat. Prod. Rep.* (2010) 27:637-57 para una revisión de la química y biología de los sideróforos. Los sideróforos son compuestos de bajo peso molecular (500-1500 daltons) que tienen una alta afinidad y selectividad por el Fe(III). La biosíntesis de sideróforos típicamente es regulada por los niveles de hierro del entorno donde se encuentra el organismo. Hay más de 500 sideróforos distintos, de los cuales 270 se han caracterizado estructuralmente. Los sideróforos o sus análogos son producidos de forma natural, o se sintetizan.

El sideróforo es una sal de mesilato de desferrioxamina B.

30 Los sideróforos con alta afinidad por el Fe(III) son convenientes para los métodos y composiciones descritos en la presente memoria. La Tabla 2 de Hider y Kong, véase antes, da las constantes de afinidad del Fe(III) de algunos sideróforos. Por lo tanto, en algunas realizaciones, el sideróforo puede tener un $\log K_f(\text{Fe}^{\text{III}})$ que es, es aproximadamente, o es más de, 25, 26, 27, 28, 29, 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, o un $\log K_f(\text{Fe}^{\text{III}})$ que está en el intervalo entre cualquiera de estos valores, por ejemplo, de 25 a 50, de 30 a 40, de 40 a 50, etc.

35 Pueden estar presentes otros agentes quelantes, tal como un sideróforo, o uno de sus análogos, EGTA, DTPA, etc., en los reactivos para la preparación de ácidos nucleicos en una variedad de concentraciones. Por ejemplo, en algunas realizaciones, los reactivos para la preparación de ácidos nucleicos pueden tener una concentración de un sideróforo, o uno de sus análogos, que es, es aproximadamente, o es más de 0.01 mM, 0.02 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 1 mM, o una concentración de un sideróforo, o uno de sus análogos, que es un intervalo entre cualquiera de esos valores, por ejemplo de 0.01 mM a 0.05 mM, de 0.02 mM a 0.5 mM, etc. En algunas realizaciones, los reactivos para la preparación de ácidos nucleicos pueden tener una concentración de EGTA, que es, es aproximadamente, o es más de 0.01 mM, 0.02 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 1 mM, o una concentración de EGTA, que es un intervalo entre cualquiera de esos valores, por ejemplo de 0.01 mM a 0.05 mM, de 0.02 mM a 0.5 mM, etc. En algunas realizaciones, los reactivos para la preparación de ácidos nucleicos pueden tener una concentración de DTPA, que es, es aproximadamente, o es más de 0.01 mM, 0.02 mM, 0.05 mM, 0.1 mM, 0.2 mM, 0.5 mM, 1 mM, o una concentración de DTPA, que es un intervalo entre cualquiera de esos valores, por ejemplo de 0.01 mM a 0.05 mM, de 0.02 mM a 0.5 mM, etc.

Otras características deseables para los sideróforos es la capacidad para reducir la generación de especies de oxígeno reactivas y/o radicales libres en presencia de Fe^{2+} .

50 Los métodos y composiciones descritos en la presente memoria se pueden usar en una variedad de aplicaciones de ácidos nucleicos, por ejemplo, hibridación, amplificación, ligado, extensión, lavado, secuenciación, etc. Se describen reactivos comunes en la Guía de enriquecimiento de TruSeq®, Guía de preparación de muestras de enriquecimiento de Nextera®, Guía de enriquecimiento de captura rápida de Nextera®, Guía de preparación de muestras de enriquecimiento de TruSight™, Guía de preparación de reactivos MiSeq® Reagent Kit v3, Guía de referencia de HiSeq® Cluster Kit v4, Guía de referencia de HiSeq® SBS Kit v4, Guía de usuario de NextSeq® 500 System, Guía v2 de preparación de muestras de ARN TruSeq™, Guía de preparación de bibliotecas de ADN de Nextera XT, Guía de preparación de muestras pares de acoplamiento de Nextera, Guía de preparación de muestras de Nano ADN de

TruSeq, Guía de preparación de muestras de ARN pequeño de TruSeq®, Guía de preparación de muestras de ARN de cadena simple de TruSeq®, Guía de preparación de muestras de ARN total de cadena simple de TruSeq®, Guía de preparación de biblioteca de acceso de ARN de TruSeq®, etc. (Illumina®, Inc., San Diego CA). Sin embargo, los reactivos citados se modifican para reducir o excluir cualquier EDTA y para incluir EGTA, DTPA, un sideróforo, uno de sus análogos o una de sus combinaciones, como se describe en la presente memoria.

Los reactivos y/o ingredientes comunes incluyen uno o más de los siguientes: Tampón de resuspensión (RSB), Tampón objetivo de captura 1 de Nestera® (NCT1), Tampón de objetivo de elución 1 (ET1), Tampón de objetivo de elución 2 (ET2), Tampón de hibridación de enriquecimiento (EHB), Tampón de elución de enriquecimiento 1 (EE1), Solución de lavado de enriquecimiento (EWS), Solución de lavado 1 (WS1), Solución de lavado 2 (WS2), Solución de lavado 3 (WS3), Mezcla maestra para PCR (TC#-PMM), Mezcla para amplificación de enriquecimiento de Nextera® (NEM), Mezcla para amplificación de biblioteca de Nextera® (NLM), Tampón de hibridación HT1, Tampón de lavado HT2, Tampón de lavado PR1, Tampón de lavado PR2, Tampón de lavado PR3, Tampón de lavado SB1, Tampón de lavado SB2, Tampón de lavado SB3, Mezcla de barrido universal USM, Reactivo de barrido SRE, Reactivo de barrido SRM, Tampón de lavado BB2, Tampón de lavado BB3, Tampón de lavado BB4, LNW1 (Lavado de normalización de biblioteca 1), LNS1 (Tampón de almacenamiento de normalización de biblioteca 1), RSB (Tampón de resuspensión), BWB (Tampón de lavado de perlas), Mezcla de PCR potenciada EPM, Tampón de elución ELB, etc. (Illumina®, Inc., San Diego CA).

El sideróforo descrito en la presente memoria se puede usar con una variedad de sales o enzimas y puede ser compatible con la mayoría, si no todos, los reactivos de secuenciación y/o preparación de bibliotecas, y cualquier enzima que no requiera hierro para funcionar.

En algunas realizaciones, los reactivos descritos en la presente memoria pueden comprender una ADN polimerasa. En algunas realizaciones, los reactivos descritos en la presente memoria pueden comprender dNTP. Otros reactivos incluyen reactivos comunes en aplicaciones de ácidos nucleicos, tales como preparación de muestras y/o secuenciación.

Métodos para preparar una biblioteca de ácidos nucleicos

Algunas realizaciones proporcionan métodos para preparar una biblioteca de ácidos nucleicos usando los reactivos descritos en la presente memoria, que comprenden: proporcionar una pluralidad de moléculas de ácido nucleico de una muestra; y manipular la pluralidad de moléculas de ácido nucleico en el reactivo.

En algunas realizaciones, la manipulación de la pluralidad de moléculas de ácido nucleico comprende hibridar la pluralidad de moléculas de ácido nucleico con una pluralidad de sondas de oligonucleótidos. En algunas realizaciones, la manipulación de la pluralidad de moléculas de ácido nucleico comprende separar sondas de oligonucleótidos no unidas específicamente a la pluralidad de moléculas de ácido nucleico.

En algunas realizaciones proporcionadas en la presente memoria, se fracciona selectivamente una población de ácidos nucleicos, las especies monocatenarias y extremos salientes se eliminan de la población de ácidos nucleicos y se reparan los ácidos nucleicos. En algunas realizaciones, una población de ácidos nucleicos se puede procesar llevando a cabo dos o más fraccionamientos selectivos, eliminación de especies monocatenarias, eliminación de extremos salientes de ácidos nucleicos y reparación de ácidos nucleicos antes de la preparación de la biblioteca. Algunas de dichas realizaciones pueden proporcionar un sustrato para la preparación de bibliotecas de ácidos nucleicos adicionales. Los métodos de preparación de una biblioteca de ácidos nucleicos son bien conocidos en la técnica. Los ejemplos incluyen adaptadores de ligado a los ácidos nucleicos en una población de ácidos nucleicos. En algunas realizaciones, los fragmentos de ácidos nucleicos pueden ser de extremos romos, fosforilados, acoplados a colas-A y/o acoplados a adaptadores para dar una biblioteca de ácidos nucleicos. En algunas realizaciones, la biblioteca de ácidos nucleicos se puede amplificar más. Los ejemplos de protocolos de preparación de bibliotecas incluyen, pero no se limitan a métodos para el kit de preparación de muestras de ADN de Nextera™ (Epicentre® Biotechnologies, Madison WI), Kit de preparación de bibliotecas GL FLX Titanium (454 Life Sciences, Branford CT), y similares. La muestra, como se describe en la presente memoria, se puede amplificar más por secuenciación o ensayos de micromatrices, por ejemplo, por técnicas de amplificación por desplazamiento múltiple de cadena (MDA). Para la secuenciación después de la MDA, se prepara una biblioteca de muestras amplificadas, por ejemplo, creando una biblioteca de ADN como se describe en el kit de preparación de biblioteca de pares de acoplamiento, kits de preparación de muestras de ADN genómico o preparación de muestras de TruSeq™ o kits de enriquecimiento de exoma (Illumina®, Inc., San Diego CA). Otro método útil para amplificar ácidos nucleicos es la amplificación por círculo rodante (RCA), por ejemplo, como se describe en Lizardi et al., *Nat. Genet.* (1998) 19:225-232 y documento US 2007/0099208. Los métodos de PCR en emulsión también son útiles, y se describen métodos de ejemplo en Dressman et al., *Proc. Natl. Acad. Sci. USA* (2003) 100:8817-8822, WO 05/010145 o publicación de patente de EE.UU. N° 2005/0130173 o 2005/0064460. Los métodos de la presente descripción no están limitados por ninguna preparación de biblioteca particular o método de amplificación.

En algunas realizaciones, la pluralidad de moléculas de ácido nucleico y/o la pluralidad de sondas de oligonucleótidos se inmovilizan sobre un soporte. La pluralidad de moléculas de ácido nucleico o sondas de oligonucleótidos se pueden inmovilizar sobre un soporte en una serie de formas. En algunas realizaciones, se usan marcadores de purificación.

Por "marcador de purificación" se entiende en la presente memoria un resto que se puede usar para purificar una molécula de ácido nucleico, normalmente por unión a un soporte sólido como se indica en la presente memoria. Los marcadores de purificación adecuados incluyen miembros de pares de unión. Por ejemplo, el marcador puede ser un hapteno o antígeno que se unirá a su pareja de unión. Por ejemplo, los pares de unión adecuados incluyen, pero no se limitan a: antígenos (tales como proteínas (incluyendo péptidos)) y anticuerpos (incluyendo fragmentos de los mismos (FAb, etc.)); proteínas y moléculas pequeñas, incluyendo biotina/estreptavidina; enzimas y sustratos o inhibidores; otros pares que interaccionan de proteína-proteína; receptor-ligandos; e hidratos de carbono y sus parejas de unión. También son útiles los pares de proteínas de unión de ácido nucleico-ácido nucleico.

En algunas realizaciones, la pluralidad de moléculas de ácido nucleico y/o la pluralidad de sondas de oligonucleótidos se inmovilizan sobre un soporte por un par de unión al soporte. En algunas realizaciones, el soporte es una perla magnética.

En algunas realizaciones, los métodos descritos en la presente memoria comprenden modificar las sondas de oligonucleótidos específicamente unidas a la pluralidad de moléculas de ácido nucleico. En algunas realizaciones, los métodos descritos en la presente memoria comprenden fragmentar la pluralidad de moléculas de ácido nucleico. En algunas realizaciones, añadir adaptadores a la pluralidad de moléculas de ácido nucleico. En algunas realizaciones, los adaptadores se añaden a la pluralidad de moléculas de ácido nucleico por amplificación.

Métodos para reducir el daño oxidativo a una molécula de ADN

Algunas realizaciones descritas en la presente memoria proporcionan el uso de métodos para reducir el daño oxidativo a una molécula de ácido nucleico, cuyos métodos comprenden preparar la molécula de ácido nucleico en ausencia de EDTA, más preferiblemente en ausencia de Fe^{2+} (EDTA).

El uso de un agente quelante distinto del EDTA, o sin usar un agente quelante, puede ser conveniente para los métodos y composiciones descritos en la presente memoria. En algunas realizaciones, preparar la molécula de ácido nucleico comprende preparar la molécula de ácido nucleico en presencia de EGTA, DTPA, un sideróforo, o una de sus combinaciones. Preparar la molécula de ácido nucleico comprende preparar la molécula de ácido nucleico en presencia de un sideróforo. El sideróforo es la sal de mesilato de desferrioxamina B (DFO-B).

Los métodos y composiciones descritos en la presente memoria son útiles para reducir las especies de oxígeno reactivas (ROS) y/o radicales libres durante la preparación de la molécula de ácido nucleico. Por ejemplo, en algunas realizaciones, preparar la molécula de ácido nucleico comprende exponer la molécula de ácido nucleico a $Fe(III)$. En algunas realizaciones, preparar la molécula de ácido nucleico comprende exponer la molécula de ácido nucleico a una perla magnética. Al preparar la muestra en ausencia de EDTA y/o incluyendo EGTA, DTPA, un sideróforo, o una de sus combinaciones, se pueden reducir las ROS.

Las ROS se generan a partir del metabolismo normal del oxígeno en las células y son reconocidas como una lista de moléculas activas, tales como superóxido; peróxido de hidrógeno; radical hidroxilo; ion hidroxilo; óxido nítrico, etc. Se sabe que las ROS causan daños oxidativos en moléculas de ácidos nucleicos. Entre las lesiones descubiertas hasta ahora, una de las más abundantes en el ADN y ARN es la 8-oxoguanina. Véase Kasai et al., *Nucleic Acids Res.* (1984) 12:2127-2136; Kasai y Nishimura, *Environ Health Perspect.* (1986) 67:111-116; Kanvah et al., *Acc. Chem. Res.* (2010) 43(2):280-287. Por lo tanto, en algunas realizaciones, el daño oxidativo comprende una mutación puntual en la molécula de ácido nucleico. Como se usa en la presente memoria, la expresión "mutación puntual" se refiere a una sustitución, una eliminación o una inserción de un nucleótido o más. En algunas realizaciones, la mutación puntual es una transversión de C a A. Al reducir las ROS, se puede reducir el daño oxidativo a ácidos nucleicos.

Algunos de los métodos y composiciones proporcionados en la presente memoria incluyen preparar bibliotecas a partir de ácidos nucleicos obtenidos de muestras. Como se usa en la presente memoria "muestra" incluye una variedad de fuentes y composiciones que contienen ácidos nucleicos. La muestra puede ser una muestra biológica, pero el término también incluye otras muestras, p. ej., artificiales, que comprenden ácidos nucleicos tales como productos de la PCR o composiciones que ya comprenden ácidos nucleicos purificados que se pueden concentrar más y/o purificar más. Las muestras biológicas incluyen partículas víricas, células, tejidos, órganos y cualquier parte de un organismo.

En algunas realizaciones de los métodos y composiciones proporcionados en la presente memoria, se reparan los ácidos nucleicos en una población de ácidos nucleicos para preparar una biblioteca de ácidos nucleicos. Por ejemplo, las mellas se pueden rellenar y reparar; los extremos salientes se pueden copiar y formar segmentos bicatenarios de un ácido nucleico. Los métodos para reparar ácidos nucleicos son bien conocidos en la técnica. En algunas realizaciones, la reparación de ácidos nucleicos puede incluir la escisión de bases modificadas o dañadas, eliminación de sitios abásicos, relleno de mellas, ligado de ligado de mellas, eliminación de grupos de bloqueo 3', e inversión de entrecruzamientos tales como dímeros de pirimidina.

Métodos para aumentar la puntuación Q de una reacción de secuenciación

Algunas realizaciones descritas en la presente memoria proporcionan el uso de métodos para aumentar la puntuación de Q (calidad Phred) de una reacción de secuenciación, cuyos métodos comprenden preparar una molécula de ácido nucleico en ausencia de EDTA y en presencia de una sal de mesilato de desferrioxamina B (DFO-B). La puntuación

Q fue desarrollada originalmente por el programa Phred, y se define como una propiedad que está relacionada logarítmicamente con las probabilidades P de error de llamada de una base: $Q = -10\log_{10}P$. Por ejemplo, si Phred asigna una puntuación de calidad de 30 a una base, las probabilidades de que esta base sea llamada de forma incorrecta son 1 en 1000.

5 La puntuación Q de una reacción de secuenciación se puede optimizar usando los métodos y composiciones descritos en la presente memoria. Por ejemplo, la reacción de secuenciación puede tener una puntuación Q que es, es aproximadamente, o es más de 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, o una puntuación Q que es un intervalo entre cualquiera de estos valores, por ejemplo, de 30 a 50, de 30 a 40, de 40 a 50, de 44 a 46, etc. En algunas realizaciones, la puntuación Q es mayor de aproximadamente 34. En algunas realizaciones, la puntuación Q es mayor de aproximadamente 38. En algunas realizaciones, la puntuación Q es mayor de aproximadamente 42.

15 Los métodos y composiciones descritos en la presente memoria son especialmente útiles para reacciones de secuenciación profunda. Como se usa en la presente memoria, la "secuenciación profunda" se refiere a una reacción de secuenciación durante la cual el número total de lecturas es muchas veces mayor que la longitud de la secuencia que se está estudiando. La profundidad (cobrimiento) se usa para referirse a las lecturas medias que representan un nucleótido dado en la secuencia reconstruida. La profundidad se puede calcular a partir de la longitud del genoma original (G), el número de lecturas (N), y la longitud de lectura media (L) como $N \times L/G$. Por consiguiente, la secuenciación profunda significa que la profundidad de la reacción de secuenciación es mayor de 7. Por ejemplo, un nucleótido dado puede estar cubierto por una profundidad que es, es aproximadamente, o es más de 30, 31, 32, 33, 34, 35, 36, 37, 38, 39, 40, 41, 42, 43, 44, 45, 46, 47, 48, 49, 50, o una profundidad que es un intervalo entre cualquiera de estos valores, por ejemplo, de 30 a 50, de 30 a 40, de 40 a 50, de 44 a 46, etc. En algunas realizaciones, los métodos descritos en la presente memoria se pueden usar para la secuenciación ultraprofunda que tiene una profundidad de secuenciación >100 (Ajay et al., *Genome Res.* (2011) 21:1498-505). En algunas realizaciones, la aplicación de secuenciación profunda es una aplicación de secuenciación profunda relacionada con el cáncer. En algunas realizaciones, los métodos comprenden secuenciar la molécula de ácido nucleico en ausencia de EDTA y/o en presencia de EGTA, DTPA, un sideróforo, o una de sus combinaciones, como se describe en la presente memoria.

Métodos de secuenciación

30 Los métodos descritos en la presente memoria se pueden usar junto con una variedad de técnicas de secuenciación de ácidos nucleicos. Las técnicas particularmente aplicables son aquellas en donde los ácidos nucleicos están unidos a ubicaciones fijas en una matriz, de modo que sus posiciones relativas no cambian y en donde se generan imágenes de la matriz de forma repetida. Se pueden aplicar en particular realizaciones en las que se obtienen imágenes en diferentes canales de color, por ejemplo, que coinciden con diferentes marcadores usados para distinguir un tipo de base de nucleótido de otro tipo. En algunas realizaciones, el procedimiento para determinar la secuencia de nucleótidos de un ácido nucleico objetivo puede ser un procedimiento automático. Las realizaciones preferidas incluyen técnicas de secuenciación por síntesis ("SBS").

35 Las "técnicas de secuenciación por síntesis ("SBS")" en general implican la extensión enzimática de una cadena de ácido nucleico nueva por la adición iterativa de nucleótidos contra una cadena molde. En métodos tradicionales de SBS, se puede proporcionar un solo monómero nucleótido a un nucleótido objetivo en presencia de una polimerasa en cada suministro. Sin embargo, en los métodos descritos en la presente memoria, se puede proporcionar más de un tipo de monómero nucleótido a un ácido nucleico objetivo en presencia de una polimerasa, en un suministro.

40 La SBS puede usar monómeros nucleótidos que tienen un resto terminador o aquellos que carecen de cualquier resto terminador. Los métodos que usan monómeros nucleótidos que carecen de terminadores incluyen, por ejemplo, pirosecuenciación y secuenciación usando nucleótidos marcados con γ -fosfato, como se expone con más detalle más adelante. En métodos que usan monómeros nucleótidos que carecen de terminadores, el número de nucleótidos añadido en cada ciclo en general es variable y depende de la secuencia del molde y el modo de suministro de nucleótidos. Para las técnicas de SBS que usan monómeros nucleótidos que tienen un resto terminador, el terminador puede ser efectivamente irreversible en las condiciones de secuenciación usadas como es el caso de la secuenciación de Sanger tradicional que usa dideoxinucleótidos, o el terminador puede ser reversible como es el caso para métodos de secuenciación desarrollados por Solexa (ahora Illumina, Inc.).

45 50 Las técnicas de SBS pueden usar monómeros nucleótidos que tienen un resto marcador o aquellos que carecen de un resto marcador. Por consiguiente, los sucesos de incorporación se pueden detectar basándose en una característica del marcador, tal como la fluorescencia del marcador; una característica del monómero nucleótido tal como el peso molecular o la carga; un subproducto de incorporación del nucleótido, tal como la liberación de pirofosfato; o similares. En realizaciones, donde están presentes dos o más nucleótidos diferentes en un reactivo de secuenciación, los diferentes nucleótidos pueden ser distinguibles entre sí, o alternativamente, los dos o más marcadores diferentes pueden ser indistinguibles con las técnicas de detección que se están usando. Por ejemplo, los diferentes nucleótidos presentes en un reactivo de secuenciación pueden tener diferentes marcadores y pueden ser distinguibles usando técnicas ópticas adecuadas como se ilustra por los métodos de secuenciación desarrollados por Solexa (ahora Illumina, Inc.).

Las realizaciones preferidas incluyen técnicas de pirosecuenciación. La pirosecuenciación detecta la liberación de pirofosfato inorgánico (PPi) cuando se incorporan nucleótidos particulares a la cadena nueva (Ronaghi, M., Karamohamed, S., Pettersson, B., Uhlen, M. y Nyren, P. (1996) "Real-time DNA sequencing using detection of pyrophosphate release." *Analytical Biochemistry* 242(1), 84-9; Ronaghi, M. (2001) "Pyrosequencing sheds light on DNA sequencing." *Genome Res.* 11(1), 3-11; Ronaghi, M., Uhlen, M. y Nyren, P. (1998) "A sequencing method based on real-time pyrophosphate." *Science* 281(5375), 363; patente de EE.UU. N° 6,210,891; patente de EE.UU. N° 6,258,568 y patente de EE.UU. N° 6,274,320). En la pirosecuenciación, el PPi liberado se puede detectar convirtiéndolo inmediatamente en trifosfato de adenosina (ATP) por la ATP sulfurilasa, y el nivel de ATP generado se detecta por fotones producidos por luciferasa. Los ácidos nucleicos que se van a secuenciar pueden estar ligados a posiciones características en una matriz y se pueden generar imágenes de la matriz para capturar señales quimioluminiscentes que son producidas debido a la incorporación de un nucleótido en las posiciones características de la matriz. Se puede obtener una imagen después de que la matriz se trate con un tipo de nucleótido particular (p. ej., A, T, C o G). Las imágenes obtenidas después de añadir cada tipo de nucleótido diferirán con respecto a que posiciones características se detectan en la matriz. Estas diferencias en la imagen reflejan el contenido de secuencia diferente de las posiciones características en la matriz. Sin embargo, las ubicaciones relativas de cada posición característica permanecerán inalteradas en las imágenes. Las imágenes se pueden almacenar, procesar y analizar usando los métodos expuestos en la presente memoria. Por ejemplo, las imágenes obtenidas después de tratamiento de la matriz con cada tipo de nucleótido diferente se pueden manipular de la misma forma que se ilustra en la presente memoria para imágenes obtenidas de diferentes canales de detección para métodos de secuenciación basados en terminadores reversibles.

En otro tipo de SBS de ejemplo, la secuenciación en ciclo se lleva a cabo por la adición por pasos de nucleótidos terminadores reversibles que contienen, por ejemplo, un marcador colorante escindible o fotoblanqueable como se describe, por ejemplo, en la publicación de patente internacional N° WO 04/018497 y patente de EE.UU. 7,057,026. Este procedimiento es comercializado por Solexa (ahora Illumina, Inc.), y se describe también en la publicación de patente internacional N° WO 91/06678 y publicación de patente internacional N° WO 07/123,744. La disponibilidad de terminadores marcados por fluorescencia en los que se puede tanto invertir la terminación como escindir el marcador fluorescente, facilita la secuenciación por terminación reversible cíclica (CRT) eficaz. También se pueden modificar genéticamente conjuntamente polimerasas para la incorporación y extensión eficaces a partir de estos nucleótidos modificados.

Preferiblemente en realizaciones de secuenciación basadas en terminadores reversibles, los marcadores no inhiben sustancialmente la extensión en las condiciones de reacción de SBS. Sin embargo, los marcadores de detección pueden ser eliminables, por ejemplo, por escisión o degradación. Se pueden capturar imágenes después de la incorporación de marcadores en las posiciones características de ácidos nucleicos en la matriz. En realizaciones particulares, cada ciclo implica el suministro simultáneo de cuatro tipos de nucleótidos diferentes a la matriz y cada tipo de nucleótido tiene un marcador espectralmente distinto. Se pueden obtener entonces cuatro imágenes, que usa cada una un canal de detección que es selectivo para uno de los cuatro marcadores diferentes. Alternativamente, se pueden añadir diferentes tipos de nucleótidos de forma secuencial y se puede obtener una imagen de la matriz entre cada etapa de adición. En dichas realizaciones cada imagen mostrará posiciones características de ácidos nucleicos que han incorporado nucleótidos de un tipo particular. Estarán presentes o ausentes diferentes posiciones características en las diferentes imágenes debido al contenido de secuencia diferente de cada posición característica. Sin embargo, la posición relativa de las posiciones características permanecerá inalterada en las imágenes. Las imágenes obtenidas de dichos métodos de SBS-terminador reversible se pueden almacenar, procesar y analizar como se expone en la presente memoria. Después de la etapa de captura de imágenes, los marcadores se pueden eliminar y los restos terminadores reversibles se pueden eliminar para los ciclos posteriores de adición y detección de nucleótidos. La eliminación de los marcadores después de que hayan sido detectados en un ciclo particular y antes de un ciclo posterior, puede proporcionar la ventaja de reducir la señal de fondo y la interferencia entre ciclos. Se exponen más adelante ejemplos de marcadores útiles y métodos de eliminación.

En realizaciones particulares, algunos o todos los monómeros nucleótidos pueden incluir terminadores reversibles. En dichas realizaciones, los terminadores reversibles/flúores escindibles pueden incluir flúor unido al resto de ribosa por un enlace 3' éster (Metzker, *Genome Res.* 15:1767-1776 (2005)). Otros planteamientos han separado la química de terminadores de la escisión del marcador fluorescente (Ruparel et al., *Proc Natl Acad Sci USA* 102: 5932-7 (2005)). Ruparel et al. describieron el desarrollo de terminadores reversibles que usaban un grupo 3'-alilo pequeño para bloquear la extensión, pero se podía desbloquear fácilmente por un tratamiento corto con un catalizador de paladio. El fluoróforo se unía a la base por un conector fotoescindible que se podía escindir fácilmente por una exposición de 30 segundos a la luz UV de longitud de onda larga. Por lo tanto, se puede usar la reducción de disulfuro o la fotoescisión como un conector escindible. Otro planteamiento para la terminación reversible es el uso de la terminación natural que sucede después de la colocación de un colorante voluminoso en un dNTP. La presencia de un colorante voluminoso cargado en el dNTP puede actuar como un terminador eficaz por impedimento estérico y/o electrostático. La presencia de un suceso de incorporación previene incorporaciones adicionales salvo que se elimine el colorante. La escisión del colorante elimina el flúor e invierte eficazmente la terminación. Se describen también ejemplos de nucleótidos modificados en la patente de EE.UU. 7,427,673 y patente de EE.UU. 7,057,026.

Se describen sistemas de SBS de ejemplo adicionales y métodos que se pueden usar con los métodos y sistemas descritos en la presente memoria en la publicación de patente de EE.UU. N° 2007/0166705, publicación de patente

de EE.UU. N° 2006/0188901, patente de EE.UU. 7,057,026, publicación de patente de EE.UU. N° 2006/0240439, publicación de patente de EE.UU. EE.UU. N° 2006/0281109, publicación de patente internacional N° WO 05/065814, publicación de patente de EE.UU. N° 2005/0100900, publicación de patente internacional N° WO 06/064199, publicación de patente internacional N° WO 07/010,251, publicación de patente de EE.UU. EE.UU. N° 2012/0270305 y publicación de patente de EE.UU. N° 2013/0260372.

Algunas realizaciones pueden usar la detección de cuatro nucleótidos diferentes usando menos de cuatro marcadores diferentes. Por ejemplo, la SBS se puede llevar a cabo usando métodos y sistemas descritos en la publicación de patente de EE.UU. N° 2013/0079232. Como un primer ejemplo, un par de tipos de nucleótidos se pueden detectar a la misma longitud de onda, pero distinguirse basándose en una diferencia en la intensidad de un miembro del par en comparación con el otro, o basándose en un cambio para un miembro del par (p. ej., por modificación química, modificación fotoquímica o modificación física) que hace que la señal clara aparezca o desaparezca en comparación con la señal detectada para el otro miembro del par. Como un segundo ejemplo, tres o cuatro tipos de nucleótidos diferentes se pueden detectar en condiciones particulares mientras que un cuarto tipo de nucleótido carece de un marcador que sea detectable en esas condiciones, o se detecta mínimamente en esas condiciones (p. ej., detección mínima debido a la fluorescencia de fondo, etc.). La incorporación de los tres tipos de nucleótidos en un ácido nucleico se puede determinar basándose en la presencia de sus respectivas señales y la incorporación del cuarto tipo de nucleótido en el ácido nucleico se puede determinar basado en la ausencia o detección mínima de cualquier señal. Como un tercer ejemplo, un tipo de nucleótido puede incluir marcador(es) que son detectados en dos canales diferentes, mientras que otros tipos de nucleótidos no son detectados en más de uno de los canales. Las tres configuraciones de ejemplo mencionadas antes no se consideran mutuamente exclusivas y se pueden usar en diferentes combinaciones. Una realización de ejemplo que combina los tres ejemplos, es un método de SBS basado en fluorescencia que usa un primer tipo de nucleótido que se detecta en un primer canal (p. ej., dATP que tiene un marcador que se detecta en el primer canal cuando es excitado por una primera longitud de onda de excitación), un segundo tipo de nucleótido que se detecta en un segundo canal (p. ej., dCTP que tiene un marcador que se detecta en el segundo canal cuando es excitado por una segunda longitud de onda), un tercer tipo de nucleótido que se detecta tanto en el primer como en el segundo canal (p. ej., dTTP que tiene al menos un marcador que se detecta en ambos canales cuando es excitado por la primera y/o segunda longitud de onda de excitación) y un cuarto tipo de nucleótido que carece de un marcador que no se detecta, o se detecta mínimamente, en cualquier canal (p. ej., dGTP que no tiene marcador).

Además, como se describe en la publicación de patente de EE.UU. N° 2013/0079232, se pueden obtener datos de secuenciación usando un solo canal. En dichos planteamientos de secuenciación llamados de un colorante, el primero tipo de nucleótido está marcado pero el marcador se elimina después de generar la primera imagen, y el segundo tipo de nucleótido se marca solo después de generarse una primera imagen. El tercer tipo de nucleótido retiene su marcador tanto en la primera como en la segunda imagen, y el cuarto tipo de nucleótido permanece sin marcar en ambas imágenes.

Algunas realizaciones pueden usar la secuenciación por técnicas de ligado (SBL). Dichas técnicas usan ADN ligasa para incorporar oligonucleótidos e identificar la incorporación de dichos oligonucleótidos. Los oligonucleótidos típicamente tienen diferentes marcadores que se correlacionan con la identidad de un nucleótido particular en una secuencia con la que los oligonucleótidos se hibridan. Como con otros métodos de SBS, se pueden obtener imágenes después del tratamiento de una matriz de posiciones características de ácidos nucleicos con los reactivos de secuenciación marcados. Cada imagen mostrará posiciones características de ácidos nucleicos que han incorporado marcadores de un tipo particular. Estarán presentes o ausentes diferentes posiciones características en las diferentes imágenes debido al contenido de secuencia diferente de cada posición característica, pero la posición relativa de las posiciones características permanecerá inalterada en las imágenes. Las imágenes obtenidas de los métodos de secuenciación basados en ligado se pueden almacenar, procesar y analizar como se expone en la presente memoria. Se describen sistemas de SBL de ejemplo y métodos que se pueden usar con los métodos y sistemas descritos en la presente memoria patente de EE.UU. 6,969,488, patente de EE.UU. 6,172,218 y patente de EE.UU. 6,306,597.

Algunas realizaciones pueden usar la secuenciación por nanoporos (Deamer, D. W. y Akeson, M. "Nanopores and nucleic acids: prospects for ultrarapid sequencing." *Trends Biotechnol.* 18, 147-151 (2000); Deamer, D. y D. Branton, "Characterization of nucleic acids by nanopore analysis". *Acc. Chem. Res.* 35:817-825 (2002); Li, J., M. Gershow, D. Stein, E. Brandin, y J. A. Golovchenko, "DNA molecules and configurations in a solid-state nanopore microscope" *Nat. Mater.* 2:611-615 (2003)). En dichas realizaciones, el ácido nucleico objetivo pasa a través de un nanoporo. El nanoporo puede ser un poro sintético o proteína de membrana biológica, tal como α -hemolisina. Cuando el ácido nucleico objetivo pasa a través del nanoporo, cada par de bases puede ser identificado midiendo fluctuaciones en la conductancia eléctrica del poro. (Patente de EE.UU. 7,001,792; Soni, G. V. y Meller, "A. Progress toward ultrafast DNA sequencing using solid-state nanopores." *Clin. Chem.* 53, 1996-2001 (2007); Healy, K. "Nanopore-based single-molecule DNA analysis." *Nanomed.* 2, 459-481 (2007); Cockroft, S. L., Chu, J., Amorin, M. y Ghadiri, M. R. "A single-molecule nanopore device detects DNA polymerase activity with single-nucleotide resolution." *J. Am. Chem. Soc.* 130, 818-820 (2008)). Los datos obtenidos de la secuenciación por nanoporos se pueden almacenar, procesar y analizar como se expone en la presente memoria. En particular, los datos se pueden tratar como una imagen de acuerdo con el tratamiento de ejemplo de imágenes ópticas y otras imágenes que se expone en la presente memoria.

Algunas realizaciones pueden usar métodos que implican el control en tiempo real de la actividad de la ADN

polimerasa. Las incorporaciones de nucleótidos se pueden detectar a través de interacciones de transferencia de energía por resonancia de fluorescencia (FRET) entre una polimerasa que lleva fluoróforo y nucleótidos marcados con γ -fosfato como se describe, por ejemplo, en la patente de EE.UU. 7,329,492 y patente de EE.UU. 7,211,414, o las incorporaciones de nucleótidos se pueden detectar con guías de onda en modo cero como se describe, por ejemplo, en la patente de EE.UU. 7,315,019 y usando análogos de nucleótidos fluorescentes y polimerasas genéticamente modificadas como se describe, por ejemplo, en la patente de EE.UU. 7,405,281 y publicación de patente de EE.UU. N° 2008/0108082. La iluminación se puede restringir a un volumen a escala de zeptolitro alrededor de una polimerasa unida a una superficie de modo que la incorporación de nucleótidos marcados por fluorescencia se puede observar con un ruido de fondo bajo (Levene, M. J. et al. "Zero-mode waveguides for single-molecule analysis at high concentrations." *Science* 299, 682-686 (2003); Lundquist, P. M. et al. "Parallel confocal detection of single molecules in real time." *Opt. Lett.* 33, 1026-1028 (2008); Korfach, J. et al. "Selective aluminum passivation for targeted immobilization of single DNA polymerase molecules in zero-mode waveguide nano structures." *Proc. Natl. Acad. Sci. USA* 105, 1176-1181 (2008)). Las imágenes obtenidas de dichos métodos se pueden almacenar, procesar y analizar como se expone en la presente memoria.

Algunas realizaciones de SBS incluyen la detección de un protón liberado tras la incorporación de un nucleótido en un producto de extensión. Por ejemplo, la secuenciación basada en la detección de protones liberados puede usar un detector eléctrico y técnicas asociadas que están disponibles en el comercio en Ion Torrent (Guilford, CT, una filial de Life Technologies) o métodos de secuenciación y sistemas descritos en la publicación de patente de EE.UU. N° 2009/0026082; publicación de patente de EE.UU. N° 2009/0127589; publicación de patente de EE.UU. N° 2010/0137143; o publicación de patente de EE.UU. N° 2010/0282617. Los métodos expuestos en la presente memoria para amplificar ácidos nucleicos objetivo usando exclusión cinética se pueden aplicar fácilmente a sustratos usados para detectar protones. Más específicamente, los métodos expuestos en la presente memoria se pueden usar para producir poblaciones clonales de amplicones que se usan para detectar protones.

Los métodos de SBS anteriores se pueden llevar a cabo ventajosamente en formatos de multiplex de modo que se manipulan simultáneamente múltiples ácidos nucleicos objetivo diferentes. En realizaciones particulares, se pueden tratar diferentes ácidos nucleicos objetivo en una recipiente de reacción común o sobre una superficie de un sustrato particular. Esto permite el suministro conveniente de reactivos de secuenciación, eliminación de reactivos sin reaccionar y la detección de sucesos de incorporación en una forma de multiplex. En realizaciones que usan ácidos nucleicos objetivo unidos a superficie, los ácidos nucleicos objetivo pueden estar en un formato de matriz. En un formato de matriz, los ácidos nucleicos objetivo pueden estar típicamente unidos a una superficie de una forma espacialmente distinguible. Los ácidos nucleicos objetivo se pueden unir por unión covalente directa, unión a una perla u otra partícula o unión a una polimerasa u otra molécula que está unida a la superficie. La matriz puede incluir una sola copia de un ácido nucleico objetivo en cada sitio (denominado también una posición característica) o pueden estar presentes múltiples copias que tienen la misma secuencia en cada sitio o posición característica. Se pueden producir múltiples copias por métodos de amplificación tales como, amplificación en puente o PCR en emulsión como se describe con más detalle más adelante.

Los métodos expuestos en la presente memoria pueden usar matrices que tienen posiciones características en una densidad que es, es aproximadamente, es menor de o es mayor de 10 posiciones características/cm², 100 posiciones características/cm², 500 posiciones características/cm², 1,000 posiciones características/cm², 5,000 posiciones características/cm², 10,000 posiciones características/cm², 50 000 posiciones características/cm², 100,000 posiciones características/cm², 1,000,000 posiciones características/cm², 5 000 000 posiciones características/cm², o una densidad que es un intervalo entre cualquiera de estos valores, por ejemplo, de 10 posiciones características/cm² a 5,000,000 posiciones características/cm², de 100 posiciones características/cm² a 1,000,000 posiciones características/cm², de 500 posiciones características/cm² a 100 000 posiciones características/cm², de 1,000 posiciones características/cm² a 50,000 posiciones características/cm², de 5,000 posiciones características/cm² a 10,000 posiciones características/cm², etc.

Una ventaja de los métodos expuestos en la presente memoria es que proporcionan una detección rápida y eficaz de una pluralidad de ácidos nucleicos objetivo en paralelo. Por consiguiente, la presente descripción proporciona sistemas integrados capaces de preparar y detectar ácidos nucleicos usando técnicas conocidas en la materia tales como las ilustradas antes. Por lo tanto, un sistema integrado de la presente descripción puede incluir componentes fluidos capaces de suministrar reactivos de amplificación y/o reactivos de secuenciación a uno o más fragmentos de ADN inmovilizados, comprendiendo el sistema componentes tales como bombas, válvulas, depósitos, conductos de fluidos y similares. Se puede configurar una celda de flujo y/o usar en un sistema integrado para detectar ácidos nucleicos objetivo. Las celdas de flujo de ejemplo se describen, por ejemplo, en la publicación de patente de EE.UU. N° 2010/0111768 A1 y solicitud de patente de EE.UU. N° 13/273,666. Como se ha ilustrado para celdas de flujo, se pueden usar uno o más de los componentes fluidos de un sistema integrado para un método de amplificación y para un método de detección. Considerando una realización de secuenciación de ácido nucleico como un ejemplo, se pueden usar uno o más de los componentes fluidos de un sistema integrado para un método de amplificación expuesto en la presente memoria y para el suministro de reactivos de secuenciación en un método de secuenciación tal como los ilustrados antes. Alternativamente, un sistema integrado puede incluir sistemas fluidos separados para llevar a cabo métodos de amplificación y llevar a cabo métodos de detección. Los ejemplos de sistemas de secuenciación integrados que son capaces de crear ácidos nucleicos amplificados y también determinar la secuencia de los ácidos nucleicos incluyen, sin limitación, la plataforma MiSeq™ (Illumina, Inc., San Diego, CA) y dispositivos descritos en la

solicitud de patente de EE.UU. N° 13/273,666.

Ejemplos

5 Con el fin de facilitar la comprensión, se proporcionan realizaciones específicas para ayudar a interpretar la propuesta técnica, es decir, estas realizaciones tienen fines solo ilustrativos, pero de ninguna forma limitan el alcance de la invención. Salvo que se especifique de otra forma, las realizaciones que no indican las condiciones específicas están de acuerdo con las condiciones convencionales o las condiciones recomendadas por el fabricante.

Ejemplo 1. Comparación de puntuaciones Q para muestras preparadas usando diferentes agentes quelantes.

10 Se llevó a cabo un estudio comparativo usando un tampón Tampón de lavado de estreptavidina SWS para la preparación de muestras de ácidos nucleicos, donde el tampón se modificó para eliminar el EDTA e incluir DFO-B, EGTA o DTPA en concentraciones de 2.0 mM, 1.5 mM, 1.0 mM, 0.5 mM, 0.2 mM, 0.1 mM, 0.05 mM o 0.0 mM (EWS_noEDTA). El tampón no modificado que contenía EDTA 0.2 mM se incluyó como control (EWS_Control). Después de la preparación de muestra, las muestras se secuenciaron usando el protocolo de secuenciación de HiSeq 15 2500, en reactivos convencionales, y se calcularon las puntuaciones Q. Las puntuaciones Q para las muestras individuales muestran un aumento diferente y significativo en la calidad máxima de las llamadas de bases debido a la reducción en la oxidación y a la transversión de C a A (Figura 1). La eliminación de EDTA de la muestra conducía a un aumento significativo en la puntuación Q, dando como resultado la inclusión de DFO-B una mejora adicional. La sustitución de EDTA por EGTA o DTPA también dio como resultado una puntuación Q mejor.

20 El resultado muestra que la sal de mesilato de DFO-B, un sideróforo bacteriano, supera el daño oxidativo al ADN durante el procedimiento de preparación de la biblioteca bioquímica debido a su capacidad para quelar el FeIII. Las ventajas adicionales proporcionadas por DFO-B se basan en la capacidad del agente para reaccionar directamente con radicales libres formados en determinadas condiciones en disolución. Los radicales libres también pueden dañar el ADN conduciendo a la pérdida de integridad de las muestras preparadas para la secuenciación.

La anterior descripción detallada de realizaciones se refiere a los dibujos que acompañan, que ilustran realizaciones específicas de la presente descripción.

25

REIVINDICACIONES

1. Un método para preparar una biblioteca de ácidos nucleicos, que comprende:
proporcionar una pluralidad de moléculas de ácido nucleico de una muestra; y
manipular la pluralidad de moléculas de ácido nucleico en un reactivo que comprende una sal de mesilato de desferrioxamina B (DFO-B) y el reactivo no contiene EDTA.
2. El método de la reivindicación 1, en donde la manipulación de la pluralidad de moléculas de ácido nucleico comprende hibridar la pluralidad de moléculas de ácido nucleico con una pluralidad de sondas de oligonucleótidos.
3. El método de la reivindicación 2, en donde la pluralidad de moléculas de ácido nucleico y/o la pluralidad de sondas de oligonucleótidos están inmovilizadas sobre un soporte.
4. El método de la reivindicación 3, en donde la pluralidad de moléculas de ácido nucleico y/o la pluralidad de sondas de oligonucleótidos están inmovilizadas sobre el soporte por un par de unión.
5. El método de la reivindicación 3 o 4, en donde el soporte es una perla magnética.
6. El método de una cualquiera de las reivindicaciones 1-4, en donde la manipulación de la pluralidad de moléculas de ácido nucleico comprende separar sondas de oligonucleótidos no unidas específicamente a la pluralidad de moléculas de ácido nucleico.
7. El método de una cualquiera de las reivindicaciones 1-6, que comprende modificar las sondas de oligonucleótidos específicamente unidas a la pluralidad de moléculas de ácido nucleico.
8. El método de una cualquiera de las reivindicaciones 1-7, que comprende fragmentar la pluralidad de moléculas de ácido nucleico.
9. El método de una cualquiera de las reivindicaciones 1-8, que comprende añadir adaptadores a la pluralidad de moléculas de ácido nucleico.
10. El método de la reivindicación 9, en donde los adaptadores se añaden a la pluralidad de moléculas de ácido nucleico por amplificación.
11. Método para aumentar la puntuación Q (Phred) de una reacción de secuenciación, en donde el método comprende preparar la molécula de ácido nucleico en ausencia de EDTA y en presencia de una sal de mesilato de desferrioxamina B (DFO-B), opcionalmente en donde la puntuación Q es mayor de aproximadamente 34, opcionalmente en donde la puntuación Q es mayor de aproximadamente 38, opcionalmente en donde la puntuación Q es mayor de aproximadamente 42.
12. El método de la reivindicación 11, en donde la reacción de secuenciación es una aplicación de secuenciación profunda, opcionalmente en donde la aplicación de secuenciación profunda es secuenciación profunda relacionada con el cáncer.
13. Método para reducir el daño oxidativo a una molécula de ácido nucleico, en donde el método comprende preparar la molécula de ácido nucleico en ausencia de EDTA y en presencia de una sal de mesilato de desferrioxamina B (DFO-B).
14. El método de la reivindicación 13, en donde preparar la molécula de ácido nucleico comprende exponer la molécula de ácido nucleico a Fe(III).
15. El método de la reivindicación 13 o 14, en donde el daño oxidativo comprende una mutación puntual en la molécula de ácido nucleico, opcionalmente en donde la mutación puntual es una transversión de C a A.

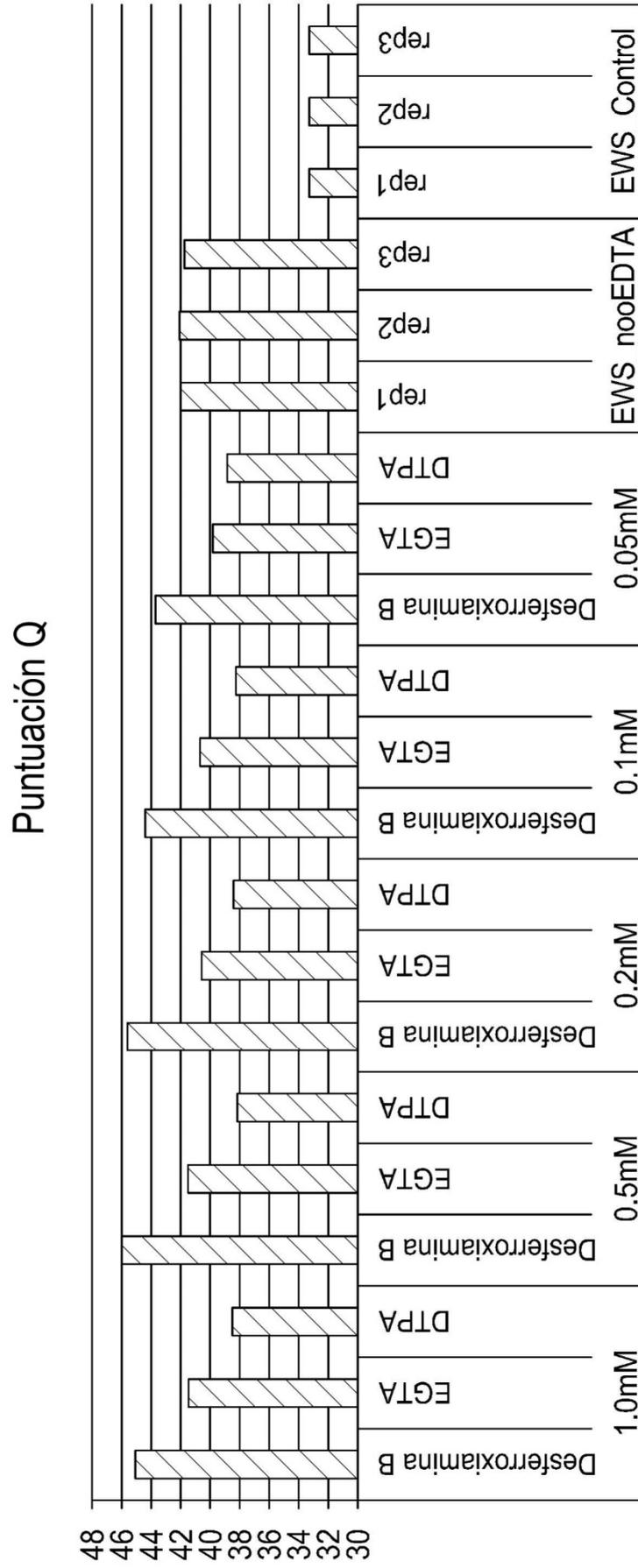


FIG. 1