

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 627 851**

51 Int. Cl.:

**C12Q 1/68** (2006.01)

**C12Q 1/48** (2006.01)

**G01N 33/53** (2006.01)

**C12P 19/34** (2006.01)

**C12N 15/10** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **12.01.2012 PCT/US2012/021040**

87 Fecha y número de publicación internacional: **09.08.2012 WO12106081**

96 Fecha de presentación y número de la solicitud europea: **12.01.2012 E 12741790 (5)**

97 Fecha y número de publicación de la concesión europea: **08.03.2017 EP 2670864**

54 Título: **Métodos para reducir el daño en ácidos nucleicos**

30 Prioridad:

**31.01.2011 US 201113018255**  
**01.02.2011 US 201161438522 P**

45 Fecha de publicación y mención en BOPI de la traducción de la patente:  
**31.07.2017**

73 Titular/es:

**ILLUMINA, INC. (50.0%)**  
**5200 Illumina Way**  
**San Diego, CA 92122, US y**  
**ILLUMINA CAMBRIDGE LIMITED (50.0%)**

72 Inventor/es:

**KLAUSING, KAY;**  
**SHEN, MIN-JUI RICHARD;**  
**MOORE, JOHN;**  
**SMITH, VINCENT;**  
**HALL, KEVIN;**  
**GORMLEY, NIAL ANTHONY;**  
**IOANNOU, AVGOUSTA;**  
**FRITZILAS, EPAMEINONDAS y**  
**RIGATTI, ROBERTO**

74 Agente/Representante:

**ELZABURU, S.L.P**

**Observaciones :**

**Véase nota informativa (Remarks, Remarques o Bemerkungen) en el folleto original publicado por la Oficina Europea de Patentes**

ES 2 627 851 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

Métodos para reducir el daño en ácidos nucleicos

**Antecedentes**

5 Los métodos para procesar muestras de ácido nucleico para y durante la secuenciación y otras técnicas pueden exponer los ácidos nucleicos a una serie de condiciones dañinas que degradan o dañan los ácidos nucleicos dando como resultado muestras de ácido nucleico que no pueden procesarse más o, p.ej., en el caso de secuenciación, pueden dar lugar a altas tasas de error.

10 Por ejemplo, los métodos para preparar muestras de ácidos nucleicos para secuenciación y otras técnicas a menudo implican fragmentación de secuencias de ácido nucleico de mayor tamaño en secuencias de ácido nucleico más pequeñas que pueden conducir a ácidos nucleicos dañados dando como resultado una identificación errónea de los ácidos nucleicos durante, p.ej., técnicas de secuenciación.

15 A modo de otro ejemplo, métodos utilizados para detectar y caracterizar estructuras de ácido nucleico pueden emplear estrategias de etiquetado que se basan en emisión de radiación electromagnética (EM) de un cromóforo que absorbe luz en estado excitado que puede dar lugar a daño del ácido nucleico. Ejemplos de tales procesos fotoluminiscentes incluyen emisión de fosforescencia y fluorescencia. La detección de fluorescencia, por ejemplo, se ha utilizado en la secuenciación de ADN con gran efecto debido, en parte, al alto grado de sensibilidad que permite la detección de una sola molécula.

20 La realización de etapas de detección fluorescente iterativas en un contexto de matriz, tal como secuenciación por síntesis, puede causar pérdida de intensidad de la señal de fluorescencia (véase, por ejemplo, Fedurco et al., documento WO2006/064199). Este problema se abordó, en parte, mediante la adición de ascorbato a una solución de detección para aumentar el número de ciclos de detección útiles de aproximadamente ocho a diez ciclos, en ausencia de ascorbato, a aproximadamente 25 ciclos en presencia de ascorbato. Los posibles mecanismos que subyacen a esta pérdida de señal son numerosos, y pueden incluir la escisión de miembros de ácidos nucleicos individuales desde el soporte. El documento WO2007/064905 también describe composiciones, dispositivos, sistemas y métodos para reducir y/o prevenir el fotodaño de uno o más agentes reaccionantes en reacciones analíticas de luminiscencia.

30 Hay una serie de vías por las que el daño en ácidos nucleicos puede ocurrir durante la irradiación en la detección de fluorescencia. La emisión de fluorescencia se produce normalmente con la emisión de luz de una longitud de onda más larga (menor energía) que la fuente de irradiación original. Sin embargo, bajo condiciones en las que la radiación EM intensa está siendo absorbida por el fluoróforo, tal como en la fluorescencia inducida por láser (LIF), es posible que una molécula absorba dos fotones, que puede conducir a la emisión de radiación de mayor energía de menor longitud de onda que la fuente de excitación original. Esta absorción de más de un fotón puede hacer que el fluoróforo emita radiación EM en la región visible UV pudiendo contribuir a la dimerización de bases del ácido nucleico y/o a la generación de especies reactivas de oxígeno.

35 Por ejemplo, se ha indicado que la exposición de células enteras a la radiación ultravioleta (UV) puede causar daño al ADN a través de la reacción de fotocicloaddición [2+2] fotoquímica directa de timina o citosina para proporcionar dímeros de ciclobutano pirimidina, tales como TT, TC y CC. Tales reacciones de fotocicloaddición directa pueden producirse en las regiones de UV B y UV C que se extienden desde aproximadamente 100 nm hasta aproximadamente 315 nm.

40 En la región UV A a través de una porción de la región visible, que abarca desde aproximadamente 315 nm a aproximadamente 500 nm, una mezcla compleja de mecanismos indirectos también puede causar daño al ADN a través de fotosensibilización de otros componentes celulares. Dichos mecanismos indirectos pueden dar como resultado la formación de dímeros de pirimidina y la modificación oxidativa del ADN a través de especies reactivas tales como oxígeno singlete, anión superóxido y formación de radicales hidroxilo favorecida por hierro. Finalmente, también se ha indicado que el oxígeno singlete reactivo puede ser generado por desactivación de la fluorescencia de un fluoróforo de estado excitado por oxígeno triplete. Cualquier combinación de dimerización de pirimidina directa o indirecta y daño de ácido nucleico debida a diversas especies reactivas de oxígeno observadas en células enteras puede ser la causa subyacente de la pérdida de intensidad de señal de fluorescencia observada en el contexto de matriz.

50 Existe la necesidad de reducir o inhibir el daño del ácido nucleico durante el procesamiento de los ácidos nucleicos para y durante la secuenciación y otras técnicas. Por ejemplo, existe la necesidad de reducir el daño del ácido nucleico durante la fragmentación. Además, existe la necesidad de reducir más la pérdida de intensidad de señal fluorescente para aplicaciones en secuenciación por síntesis para facilitar la secuenciación de secuencias de nucleótidos largas, incluyendo secuencias de 50, 75, 100, 200 y 500 nucleótidos o más. Además, las soluciones para la pérdida de intensidad de la señal fluorescente en el contexto de secuenciación son fácilmente aplicables a otras plataformas de detección de ácidos nucleicos que emplean etapas de irradiación múltiples.

**Compendio**

La presente invención se define en las reivindicaciones adjuntas.

5 Se proporciona en la presente memoria un método para inhibir la degradación de ácidos nucleicos durante una etapa de procesamiento de ácidos nucleicos seleccionada entre fragmentación y detección que comprende poner en contacto los ácidos nucleicos con una solución que comprende ácido gálico, análogos, derivados de los mismos o mezclas de los mismos durante la etapa de procesamiento, en donde el contacto inhibe la degradación de los ácidos nucleicos.

10 Se proporciona también en la presente memoria un método para inhibir la degradación de ácidos nucleicos inducida por la luz. El método incluye irradiar una porción de dichos ácidos nucleicos en presencia de una solución que comprende uno o más compuestos, p.ej., compuestos polifenólicos. También se proporciona un método para detectar un ácido nucleico que tiene una etiqueta fluorescente. El método incluye irradiar al menos una porción de dicho ácido nucleico con luz, en donde dicha luz comprende una longitud de onda idónea para inducir una emisión de fluorescencia, detectar la emisión de fluorescencia. Opcionalmente, el compuesto es ácido gálico, análogos o derivados del mismo, por ejemplo, un éster alquílico inferior del mismo, o mezclas de los mismos.

15 Se proporciona aquí un método para reducir o inhibir el daño de ácidos nucleicos durante la preparación de una muestra de ácidos nucleicos que comprende fragmentar las secuencias de ácidos nucleicos de la muestra en una solución que comprende uno de más compuestos, inhibiendo los compuestos la degradación de las secuencias de ácidos nucleicos de la muestra. También se proporciona un método para eliminar ácidos nucleicos dañados de una muestra que comprende secuencias de ácidos nucleicos, comprendiendo el método proporcionar una muestra de secuencias de ácidos nucleicos, y (i) tratar la muestra con una enzima que escinde secuencias de ácidos nucleicos dañadas o (ii) tratar la muestra con un anticuerpo que se une selectivamente a uno o más tipos de ácidos nucleicos dañados, en donde el tratamiento de la muestra elimina los ácidos nucleicos dañados de la muestra.

20 El detalle de una o más realizaciones se exponen en los dibujos adjuntos y la descripción siguiente. Otras características, objetos y ventajas resultarán evidentes a partir de la descripción y los dibujos, y de las reivindicaciones.

**Breve descripción de los dibujos**

La Figura 1 muestra una representación gráfica de la eficacia relativa de compuestos polifenólicos representativos y estructuras relacionadas, proporcionando protección contra la degradación inducida por la luz. El aumento de eficacia se muestra de arriba (menos) a abajo (más).

30 La Figura 2 muestra una representación gráfica de la eficacia relativa de compuestos polifenólicos representativos y estructuras relacionadas, proporcionando protección contra la degradación inducida por la luz. El aumento de eficacia se muestra de arriba (menos) a abajo (más).

La Figura 3 muestra una representación gráfica que representa el error en función del número de grupos que pasan el filtro (PF).

35 La Figura 4 muestra representaciones gráficas de cobertura en una secuencia control, en presencia de ácido gálico, en presencia de urea y en presencia de ácido gálico y urea, durante 75 ciclos.

La Figura 5 muestra representaciones gráficas de cobertura en una secuencia control, en presencia de ácido gálico, en presencia de urea y en presencia de ácido gálico y urea, durante 100 ciclos.

40 La Figura 6 es una representación gráfica que muestra que la fragmentación del ADN mediante el cizallamiento de Covaris tiene un valor máximo de Q que disminuye al aumentar la duración del cizallamiento.

La Figura 7 es una gráfica que muestra el perfil de mutación del ADN después de la fragmentación por cizallamiento de Covaris en presencia de ácido gálico (L1), en presencia de ascorbato (L2), sin tratamiento (L7) y tratamiento de digestión control HindIII (L8). L4 y L6 muestran el perfil de mutación del ADN después de la fragmentación mediante el cizallamiento de Covaris seguido por el tratamiento con FPG (L4) o PreCR (L6).

45 La Figura 8 es una gráfica que muestra el perfil de mutación del ADN después de la fragmentación por cizallamiento de Covaris en presencia de ácido gálico (L1), en presencia de ascorbato (L2), sin tratamiento (L7) y tratamiento de digestión control HindIII (L8). L4 y L6 muestran el perfil de mutación del ADN después de la fragmentación mediante el cizallamiento de Covaris seguido por el tratamiento con FPG (L4) o PreCR (L6).

**Descripción detallada**

50 La presente solicitud proporciona, en parte, un método para inhibir la degradación de ácidos nucleicos durante una etapa de procesamiento de ácidos nucleicos seleccionada entre la fragmentación y la detección, que comprende poner en contacto los ácidos nucleicos con una solución que comprende ácido gálico, un análogo o derivado del mismo o mezclas de los mismos durante la etapa de procesamiento, en donde el contacto inhibe la degradación de

los ácidos nucleicos. Opcionalmente, los ácidos nucleicos están en una matriz de ácidos nucleicos unidos a un soporte. Opcionalmente, la solución comprende además uno o más compuestos que inhiben adicionalmente la degradación de los ácidos nucleicos seleccionados del grupo que consiste en urea, ácido ascórbico o análogos o derivados de los mismos.

- 5 Opcionalmente, la etapa de procesamiento es una etapa de procesamiento de detección. Por lo tanto, la presente solicitud también proporciona un método para inhibir la degradación inducida por la luz de ácidos nucleicos durante una etapa de detección que incluye irradiar una porción de los ácidos nucleicos en presencia de una solución que tiene uno o más compuestos, por ejemplo, un compuesto polifenólico. La presencia de uno o más compuestos en la solución inhibe la cantidad de degradación inducida por la luz de los ácidos nucleicos.
- 10 La presente solicitud también está dirigida, en parte, a un método para detectar un ácido nucleico que tiene una etiqueta fluorescente, que incluye a) irradiar el ácido nucleico con luz que tiene una longitud de onda adecuada para inducir una emisión de fluorescencia; b) detectar la emisión de fluorescencia; y repetir estas etapas iterativamente. La etapa de irradiación se lleva a cabo en presencia de una solución que incluye un compuesto polifenólico y sirve para reducir la degradación inducida por la luz del ácido nucleico.
- 15 Opcionalmente, la etapa de procesamiento es la fragmentación. Las muestras de ácido nucleico se pueden fragmentar antes de una variedad de técnicas que incluyen, sin limitación, técnicas de secuenciación rápida o de alto rendimiento tales como secuenciación por síntesis y secuenciación por ligación, técnicas de detección de micromatrices de ácidos nucleicos tales como chips de genes y micromatrices de ADN y técnicas de reacción en cadena de la polimerasa cuantitativas (Q-PCR) tales como reacción en cadena de la polimerasa en tiempo real (PCR) y la PCR multiplex. Sin embargo, la fragmentación de muestras de ácido nucleico, por ejemplo, por cizallamiento (p.ej., hidrocizallamiento, nebulización, sonicación o cizallamiento acústico), da como resultado daño en el ácido nucleico. Así, se proporciona en la presente memoria un método para reducir o inhibir el daño en el ácido nucleico durante la preparación de una muestra de ácido nucleico que comprende fragmentar las secuencias de ácido nucleico en la muestra en una solución que comprende uno de más compuestos, inhibiendo los compuestos la degradación de las secuencias de ácido nucleico en la muestra. También se proporciona un método para inhibir la degradación del ácido nucleico durante la preparación de una muestra de ácido nucleico que comprende fragmentar las secuencias de ácido nucleico en la muestra en presencia de uno o más compuestos, inhibiendo los compuestos la degradación del ácido nucleico de las secuencias de ácido nucleico en la muestra.
- 20 Como se ha indicado anteriormente, opcionalmente, la muestra de ácido nucleico se está preparando para la secuenciación. La fidelidad de la información de secuencia en la muestra se mejora en comparación con una muestra de ácido nucleico preparada en ausencia de los compuestos. Así, también se proporciona en la presente memoria un método para mejorar la fidelidad de la información de secuencia de ácido nucleico durante la secuenciación. El método incluye proporcionar una biblioteca de secuencias de ácidos nucleicos a ser secuenciados, fragmentar la biblioteca de secuencias de ácidos nucleicos en presencia de uno o más compuestos, inhibiendo los compuestos la degradación del ácido nucleico, y secuenciar la biblioteca de secuencias de ácidos nucleicos, siendo mejorada la fidelidad de la información de secuencia de ácido nucleico en comparación con un control. Como se emplea a lo largo de la memoria, un control o valor de control incluye la fidelidad de información de secuencia obtenida en una muestra control (p.ej., una muestra preparada en ausencia de los compuestos descritos en la presente memoria) o puede comprender un patrón conocido.
- 30 Opcionalmente, los métodos proporcionados comprenden además tratar la muestra con una enzima que escinde secuencias de ácido nucleico que comprenden ácidos nucleicos dañados para separar cualquier ácido nucleico dañado de la muestra. Opcionalmente, los métodos proporcionados comprenden además tratar la muestra con un anticuerpo que se une selectivamente a uno o más tipos de ácidos nucleicos dañados para separar cualquier ácido nucleico dañado de la muestra. Así, en realizaciones particulares se proporciona un método para eliminar los ácidos nucleicos dañados de una muestra que comprende secuencias de ácidos nucleicos. El método incluye proporcionar una muestra de secuencias de ácidos nucleicos y (i) tratar la muestra con una enzima que escinde las secuencias de ácido nucleico dañadas o (ii) tratar la muestra con un anticuerpo que se une selectivamente a uno o más tipos de ácidos nucleicos dañados, en donde el tratamiento de la muestra elimina los ácidos nucleicos dañados de la muestra. Opcionalmente, antes del tratamiento de la muestra, la muestra se fragmenta en presencia de uno o más compuestos. Opcionalmente, se utiliza un compuesto o se usa más de un compuesto. Opcionalmente, la enzima es formamidopirimidina ADN glicosilasa (FPG). Opcionalmente, el anticuerpo se une a 8-oxo-G. Se conoce la enzima que elimina los ácidos nucleicos dañados. Además, los anticuerpos que se unen específicamente a uno o más tipos de ácidos nucleicos dañados son conocidos y se describen, por ejemplo, en van Loon y Hubscher, PNAS 106(43):18201-6(2009).
- 40 Compuestos ilustrativos, ejemplificados en las Figuras 1 y 2 y Tablas 1 y 2, utilizados en las soluciones proporcionadas protegen a los ácidos nucleicos de la degradación inducida por la luz que puede ocurrir durante las etapas de detección realizadas en diversos ensayos, que incluye, por ejemplo, el llamamiento de bases en secuenciación por síntesis. Los compuestos descritos en la presente memoria han sido identificados como compuestos que mejoran con éxito los efectos de degradación, incluyendo la degradación inducida por la luz, entre otras clases numerosas de compuestos que pueden bloquear vías de degradación mecánica plausibles. Tales clases de compuestos incluyen, sin limitación, inactivadores de radicales hidroxilo, inactivadores de especies
- 55
- 60

reactivas de oxígeno (oxígeno singlete, anión superóxido), eliminadores de oxígeno, inactivadores de estado triplete e inactivadores de agujeros.

En algunas realizaciones, se encontró que la presencia de urea, con o sin ascorbato, en la solución reducía adicionalmente la cantidad de degradación de los ácidos nucleicos cuando se usaba con un compuesto polifenólico. La Tabla 2 y las Figuras 4 y 5 muestran esta sinergia en presencia de urea y ascorbato. Mientras que la urea tiene un impacto significativo en la reducción de la degradación cuando se usa junto con un compuesto polifenólico, el uso de urea en solitario tiene un impacto menor con respecto a cualquiera de tales beneficios protectores.

Métodos de la invención que emplean una solución que tiene un compuesto polifenólico protegen la integridad de los ácidos nucleicos cuando están expuestos a irradiación repetida y/o intensa, como se podría emplear en una variedad de contextos, que incluyen sin limitación, técnicas de alto rendimiento o secuenciación rápida como, secuenciación por síntesis y secuenciación por ligación, técnicas de detección de micromatrices de ácido nucleico tales como chips de genes y micromatrices de ADN y técnicas de reacción en cadena de polimerasa cuantitativa (Q-PCR) tales como reacción en cadena de polimerasa (PCR) en tiempo real y la PCR multiplex. Como se ha descrito anteriormente, la presencia de urea puede potenciar el papel protector del compuesto polifenólico en cualquiera de estos contextos. En los ejemplos descritos en la presente memoria, el uso de la solución en la secuenciación mediante síntesis en un formato de matriz, en particular, redujo el decaimiento de la señal fluorescente y proporcionó tasas de error menores de 50 a 120 ciclos de etapas de detección repetidas.

Las soluciones, que se pueden proporcionar en forma de kit, también se pueden utilizar en técnicas de detección de fluorescencia que se han empleado en bioquímica mecanicista. Por ejemplo, la polarización de fluorescencia ha sido indicada como una potente técnica en el estudio de interacciones moleculares, que incluye sin limitación, interacciones receptor-ligando, tales como interacciones hormona-receptor, interacciones proteína-péptido e interacciones ADN-proteína. Por ejemplo, Singleton et al. *Tetrahedron* 63(17):(2007) incorporó un análogo de guanina fluorescente en oligonucleótidos en el estudio de interacciones de la proteína RecA con el ADN. La solución de detección empleada en los métodos descritos en la presente memoria puede utilizarse en medidas cinéticas en tiempo real en donde se emplean fluoróforos. Los efectos protectores contra la degradación inducida por la luz proporcionados por la solución de detección no necesitan estar limitados a la protección de ácidos nucleicos. Por lo tanto, por ejemplo, la solución de detección también se puede utilizar para proteger la integridad de proteínas, péptidos, carbohidratos y moléculas pequeñas, cualquiera de las cuales puede ser susceptible a especies reactivas de oxígeno, o similares, generadas bajo condiciones para medir la emisión de fluorescencia.

Como se utiliza en la presente memoria, el término "ácido nucleico" pretende significar al menos dos nucleótidos enlazados covalentemente. El ácido nucleico comprende el término oligonucleótido, polinucleótido y sus equivalentes gramaticales. Un ácido nucleico de la presente invención contendrá generalmente uniones fosfodiéster, aunque en algunos casos los análogos de ácido nucleico pueden tener estructuras alternativas, que comprenden, por ejemplo, fosforamida (Beaucage et al., *Tetrahedron* 49(10):1925 (1993) y sus referencias; Letsinger, *J. Org. Chem.* 35:3800 (1970); Sprinzl et al., *Eur. J. Biochem.* 81:579 (1977). Letsinger et al., *Nucl. Acids Res.* 14:3487 (1986); Sawai et al, *Chem. Lett.* 805 (1984), Letsinger et al., *J. Am. Chem. Soc.* 11:4470 (1988); y Pauwels et al., *Chemica Scripta* 26:141 9(1986)), fosforotioato (Mag et al., *Nucleic Acids Res.* 19:1437 (1991); y Patente de EE.UU. N° 5.644.048), fosforoditioato (Briu et al., *J. Am. Chem. Soc.* 111:2321 (1989), enlaces O-metilfosforoamidito (véase Eckstein, *Oligonucleotides and Analogues: A Practical Approach*, Oxford University Press), y estructuras y enlaces de ácidos nucleicos y peptidos (véase Egholm, *J. Am. Chem. Soc.* 114:1895 (1992); Meier et al., *Chem. Ont. Ed. Otros ácidos nucleicos análogos incluyen aquellos con estructuras positivas* (Denpcy et al., *Proc. Natl. Acad. Sci. USA* 92:6097 (1995); estructuras no iónicas (Patente de EE.UU. N° 5.386.023, 5.637.684, 5.602.240, 5.216.141 y 4.469.863; Kiedrowski et al., *Angew. Chem. Intl. Ed English* 30:423 (1991); Letsinger et al., *J. Am. Chem. Soc.* 110:4470 (1988); Letsinger et al., *Nucleoside & Nucleotide* 13:1597 (1994); Capítulos 2 y 3, *ASC Symposium Series* 580, "Carbohydrate Modifications in Antisense Research," Ed. Y.S. Sanghui and P. Dan Cook; Mesmaeker et al., *Bioorganic & Medicinal Chem. Lt.* 4:395 (1994); Jeffs et al., *J. Biomolecular NMR* 34:17 (1994); *Tetrahedron Lett.* 37:743 (1996)) y cadenas sin ribosa, que incluyen aquellas descritas en las Patentes de EE.UU. núms 5.235.033 y 5.034.506, y capítulos 6 y 7, *ASC Symposium Series* 580, "Carbohydrate Modifications in Antisense Research", Ed. Y.S. Sangui y P. Dan Cook. Los ácidos nucleicos que contienen uno o más azúcares carbocíclicos también se incluyen dentro de la definición de ácidos nucleicos (véase Jenkins et al., *Chem. Soc. Rev.* (1995) pag. 169-176). Varios análogos de ácidos nucleicos se describen en Rawls, *C & E News* June 2, 1997 página 35. Estas modificaciones de la cadena de ribosa-fosfato se pueden hacer para facilitar la adición de etiquetas, o para aumentar la estabilidad y la semivida de tales moléculas en entornos fisiológicos.

Un ácido nucleico contendrá generalmente una secuencia específica de cuatro bases de nucleótidos: adenina (A); citosina (C); guanina (G); y timina (T). El uracilo (U) también puede estar presente, por ejemplo, como sustituto natural de la timina cuando el ácido nucleico es ARN. El uracilo también se puede utilizar en el ADN. Un ácido nucleico utilizado en la invención puede incluir también bases nativas o no nativas. A este respecto, un ácido desoxirribonucleico nativo puede tener una o más bases seleccionadas del grupo que consiste en adenina, timina, citosina o guanina y un ácido ribonucleico puede tener una o más bases seleccionadas del grupo que consiste en uracilo, adenina, citosina o guanina. Se entenderá que un ácido desoxirribonucleico utilizado en los métodos o composiciones aquí expuestos puede incluir bases de uracilo y un ácido ribonucleico puede incluir una base de timina. Ejemplos de bases no nativas que pueden incluirse en un ácido nucleico, ya sea que tengan una estructura

nativa o una estructura análoga, incluyen, sin limitación, inosina, xantantina, hipoxantantina, isocitosina, isoguanina, 2-aminopurina, 5-metilcitosina, 5-hidroximetilcitosina, 2-aminoadenina, 6-metil adenina, 6-metil guanina, 2-propil guanina, 2-propiladenina, 2-tiouracilo, 2-tiotimina, 2-tiocitosina, 15-halouracilo, 15-halocitosina, 5-propinil uracilo, 5-propinil citosina, 6-azo uracilo, 6-azo citosina, 6-azo timina, 5-uracilo, 4-tiouracilo, 8-halo adenina o guanina, 8-amino adenina o guanina, 8-tiol adenina o guanina, 8-tioalquil adenina o guanina, 8-hidroxil adenina o guanina, 5-halo uracilo o citosina, 7-metilguanina, 7-metiladenina, 8-azaguanina, 8-azaadenina, 7-desazaguanina, 7-desazaadenina, 7-deazaguanina, 3-deazaadenina o similares. Una realización particular puede utilizar isocitosina e isoguanina en un ácido nucleico con el fin de reducir la hibridación no específica, como se describe generalmente en la Patente de EE.UU. N° 5.681.702.

10 Una base no nativa utilizada en un ácido nucleico puede tener una actividad de apareamiento de bases universal, en donde es capaz de aparearse con cualquier otra base de forma natural. Ejemplos de bases que tienen actividad de apareamiento de bases universal incluyen 3-nitropirrol y 5-nitroindol. Otras bases que se pueden utilizar incluyen aquellas que tienen actividad de apareamiento de bases con un subconjunto de las bases naturales, tal como inosina, que aparee base con citosina, adenina o uracilo.

15 Tal como se usa en la presente memoria, el término "matriz de ácidos nucleicos" significa un soporte sólido que tiene una pluralidad de ácidos nucleicos distinguibles en el espacio dispuestos sobre el mismo o en el mismo. Los ácidos nucleicos se pueden disponer en un patrón ordenado o aleatorio de características. Una característica individual puede ser, por ejemplo, una molécula de ácido nucleico espacialmente aislada, o un conjunto de moléculas de ácido nucleico tales como un agrupamiento. Una matriz puede ser una matriz compuesta que comprende una pluralidad de matrices individuales configuradas para permitir el procesamiento de múltiples muestras. Las matrices individuales, denominadas en la presente memoria como "submatrices", incluyen grupos de características de ácido nucleico. Sub-matrices aparecen en distintas regiones dentro de una matriz más grande. Las propias sub-matrices pueden estar ordenadas o desordenadas. Tales submatrices opcionalmente pueden ser espacialmente direccionables. Las sub-matrices pueden incluir agrupamientos de ácidos nucleicos idénticos. Un ejemplo de una matriz compuesta, compuesta por submatrices individuales, es una placa de microtitulación que tiene pocillos en los que la placa en su conjunto es una matriz de ácidos nucleicos (o matriz compuesta) mientras que cada pocillo individual representa una submatriz dentro de la matriz compuesta más grande.

Tal como se utiliza en la presente memoria, el término "ácido nucleico miembro" significa un único ácido nucleico unido a un soporte que es parte de una matriz y/o parte de una submatriz dentro de una matriz compuesta.

30 Tal como se utiliza en la presente memoria, el término "soporte" se refiere a un sustrato para inmovilizar una matriz de ácidos nucleicos. Un "soporte" es un material que tiene una superficie rígida o semirrígida a la que se puede unir una matriz de ácido nucleico o sobre la cual se pueden sintetizar y/o modificar ácidos nucleicos. Los soportes pueden incluir cualquier resina, microperlas, vidrio, vidrio de poro controlado (CPG), soporte de polímero, membrana, papel, plástico, tubo o placa de plástico, perla de plástico, perla de vidrio, lámina, cerámica, chip de silicio, placa de varios pocillos, membrana de nylon, fibra óptica y membrana de PVDF.

35 Un soporte puede incluir cualquier sustrato plano similar a un sello y sustratos planos que tengan pocillos, tales como una placa de microtitulación, incluyendo placas de 96 pocillos. Los sustratos planos modelos incluyen chips, placas, sustratos grabados, placas de microtitulación y reactores de células de flujo, que incluyen reactores de células de flujo de múltiples carriles que tienen múltiples canales microfluídicos, tales como las células de flujo de ocho canales usadas en la estación de trabajo de secuenciación cBot (Illumina, Inc., San Diego, CA).

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

45 La composición de un soporte puede variar, dependiendo, por ejemplo, del formato, la química y/o el método de unión y/o el método de síntesis de ácidos nucleicos. Los materiales de soporte que se pueden utilizar según la presente descripción incluyen, pero no se limitan a, polipropileno, polietileno, polibutileno, poliuretanos, nailon, metales y otros materiales adecuados. Las composiciones ilustrativas incluyen soportes, y funcionalidades químicas impartidas a ellos, utilizadas en síntesis de polipéptidos, polinucleótidos y/o fracción orgánica. Tales composiciones incluyen, por ejemplo, plásticos, cerámicas, vidrio, poliestireno, melamina, metilestireno, polímeros acrílicos, material paramagnético, sólido de torio, grafito, dióxido de titanio, látex o dextranos reticulados tales como Sepharose™, celulosa, nylon, micelas reticuladas y Teflon™, así como cualquier otro material que se pueda encontrar descrito en, por ejemplo, "Microsphere Detection Guide" de Bangs Laboratories, Fishers IN. Una partícula de soporte puede estar hecha de almidón reticulado, dextranos, celulosa, proteínas, polímeros orgánicos que incluyen polímeros de

estireno que incluyen poliestireno y metilestireno así como otros copolímeros de estireno, plásticos, vidrio, cerámica, polímeros acrílicos, materiales magnéticamente sensibles, coloides, sólido de torio, grafito, dióxido de titanio, nylon, látex o TEFION®. "Microsphere Detection Guide" de Bangs Laboratories, Fishers Inc. es una guía útil. Otros soportes modelo dentro del alcance de la presente descripción incluyen, por ejemplo, los descritos en la publicación de la solicitud de EE.UU. N° 2002/0102578 y en la Patente de EE.UU. N° 6.429.027.

Como se utiliza en la presente memoria, el daño o degradación de ácido nucleico se refiere a la modificación de una base de ácido nucleico, ya sea nativa o no nativa, desde su forma originalmente pretendida. A modo de ejemplo, el daño o degradación de ácido nucleico se refiere a una situación en donde la guanina se convierte en 8-oxo-guanina, por ejemplo por especies reactivas de oxígeno generadas durante la manipulación de muestras de ácido nucleico. Otros tipos de daño de ácido nucleico incluyen, pero no se limitan a, sitios abásicos y homo y heterodimerización de timina y citosina.

Un sitio abásico se define como una posición de nucleósido en una cadena polinucleótida a partir de la cual se ha eliminado el componente base. Los sitios abásicos pueden ocurrir naturalmente en el ADN bajo condiciones fisiológicas por hidrólisis de residuos de nucleósidos, pero también se pueden formar bajo condiciones artificiales. Una vez formados, los sitios abásicos se pueden escindir (por ejemplo, mediante tratamiento con una endonucleasa u otra enzima de escisión de hebra única, exposición a calor o álcali), proporcionando un medio para la escisión específica del sitio de una hebra de polinucleótido.

Como se utiliza en la presente memoria, la expresión "degradación inducida por la luz" significa el daño inducido por la luz a uno o más ácidos nucleicos en una matriz de ácidos nucleicos por exposición a la iluminación. Tal degradación incluye la eliminación completa o parcial de ácidos nucleicos individuales del soporte al que está unida la matriz. Por ejemplo, la degradación inducida por la luz puede incluir división de la estructura de fosfodiéster en cualquiera de los nucleótidos dentro de un ácido nucleico individual. Dicha degradación puede incluir también eliminación o reacción de una base de ácido nucleico o una etiqueta fluorescente que causa una pérdida en la función de hibridación o fluorescencia. La degradación inducida por la luz también puede incluir reticulación foto-inducida de nucleótidos. El resultado de la degradación inducida por la luz puede manifestarse como una disminución de la sensibilidad de detección de fluorescencia en una o más regiones o sub-matrices de una matriz de ácidos nucleicos cuando se realiza un ciclo a través de etapas de detección repetidas, como se puede observar, por ejemplo, cuando se realiza secuenciación por síntesis, secuenciación por ligación y escaneo de micro-matrices. Cuando se utiliza en conjunto con el término "inhibidor", esto se refiere a un bloqueo completo o parcial en el alcance del daño, por ejemplo, como se puede cuantificar por la resistencia observada de la emisión fluorescente. El daño por la luz se puede medir, por ejemplo, como una función de la intensidad de la señal de fluorescencia frente al número de etapas de irradiación repetidas (detección) realizadas sobre la matriz de ácidos nucleicos. Este proceso se denomina a veces decadencia de intensidad T. Otra evaluación del daño por la luz se puede medir como una función de tasa de error frente al número de etapas de irradiación repetidas (detección) realizadas sobre la matriz de ácidos nucleicos.

Como se utiliza en la presente memoria, el término "tasa de error de detección" se refiere a una medida de la frecuencia de error en la identificación de uno o más ácidos nucleicos etiquetados con fluorescencia en una matriz y/o submatriz de ácidos nucleicos. Por ejemplo, cuando se mide la fluorescencia en un esquema que emplea etiquetas fluorescentes multicolor, se puede producir un error por identificación errónea de una submatriz etiquetada cuando, por ejemplo, la relación señal/ruido se deteriora debido a la degradación inducida por la luz de una pluralidad de miembros de ácido nucleico en la submatriz. Así, la tasa de error de detección se incrementa con la pérdida continua de miembros de ácido nucleico individuales de una matriz o submatriz durante numerosos ciclos repetidos de irradiación.

Como se utiliza en la presente memoria, el término "irradiación" se refiere a la exposición de una serie de ácidos nucleicos a la iluminación. La exposición puede ser para el propósito de detección de fluorescencia, por ejemplo. La irradiación se puede realizar con un láser o una fuente de luz similar. La irradiación se puede realizar sobre una sección seleccionada del espectro UV-visible y puede emplear uno o más filtros de banda de longitud de onda. La irradiación se puede realizar durante un periodo de tiempo para recoger datos de emisión fluorescentes suficientes. El término "intenso" cuando se utiliza en referencia a la iluminación se refiere a la cantidad de potencia que se emplea durante la irradiación de una porción de una matriz de ácidos nucleicos. La irradiación con láser intenso incluye una cantidad entre aproximadamente 5 milivatios a aproximadamente 500 milivatios y una densidad de potencia en una cantidad de entre aproximadamente 1 a aproximadamente 200 W/mm<sup>2</sup>.

Como se utiliza en la presente memoria, el término "solución" se refiere a una solución que contiene el compuesto o compuestos descritos en la presente memoria que reducen la degradación o daño de ácido nucleico inducido por fragmentación. La solución se utiliza durante la preparación de una muestra de ácido nucleico para manipulación adicional en, por ejemplo, técnicas de secuenciación de ácidos nucleicos. Los compuestos o compuestos para uso según los métodos proporcionados se pueden suministrar como un líquido o en forma de un sólido a disolver en un disolvente adecuado antes del uso.

Una solución tal como se utiliza en el presente documento también se puede referir como una solución de detección cuando la etapa de procesamiento es una etapa de procesamiento de detección. Tal como se utiliza en la presente

memoria, el término "solución de detección" significa una solución que contiene compuestos de la presente invención que reducen la degradación inducida por la luz tras la exposición de una matriz de ácidos nucleicos a la iluminación. La solución de detección es la solución que se utiliza durante una etapa de detección que emplea irradiación.

- 5 Como se utiliza en la presente memoria, el término "tampón", cuando se usa solo, se refiere a cualquier otra solución tampón no utilizada como una solución de detección. Las soluciones tampón incluyen las utilizadas en reacción de polimerasa, hibridaciones, lavado o cualquier otra operación realizada antes del uso de la solución de detección empleada en la invención.

10 Como se utiliza en la presente memoria, el término "compuesto polifenólico" se refiere a un compuesto aromático que tiene múltiples grupos hidroxilo (es decir, grupos fenólicos) sobre un benceno u otro anillo aromático. El benceno, u otro anillo aromático, puede estar opcionalmente sustituido con otros sustituyentes y/o anillos fusionados. Los ejemplos de compuestos polifenólicos incluyen, sin limitación, ácido gálico y ésteres alquílicos inferiores del mismo, éteres monometílicos del mismo y combinaciones de ésteres alquílicos inferiores y éteres monometílicos del mismo, pirogalol e hidroquinonas tales como t-butilhidroquinona (TBHQ), 2,4,5-trihidroxi-butirofenona (THBP).

15 Como se utiliza en la presente memoria, el término "éster de alquilo inferior" se refiere a un éster de cadena alquilo C1-C6 de un ácido carboxílico. En algunas realizaciones, un "éster de alquilo inferior" se refiere a un éster de cadena alquilo C1-C4 de un ácido carboxílico. Los ésteres representativos incluyen ésteres de metilo, etilo, propilo, butilo, pentilo y hexilo. Cualquiera de los ésteres anteriores puede estar opcionalmente ramificado. Tales ésteres ramificados incluyen ésteres de isopropilo, ésteres de sec-butilo, ésteres de isobutilo y ésteres de terc-butilo, por ejemplo.

20 En algunas realizaciones, la presente invención proporciona un método para inhibir la degradación inducida por la luz de ácidos nucleicos durante una etapa de detección que incluye irradiar los ácidos nucleicos en presencia de una solución que tiene un compuesto polifenólico. La solución inhibe la cantidad de degradación inducida por la luz de los ácidos nucleicos.

25 En algunos aspectos, la presente descripción proporciona un método para detectar un ácido nucleico que tiene una etiqueta fluorescente que incluye a) irradiar el ácido nucleico con luz que tiene una longitud de onda adecuada para inducir una emisión de fluorescencia; b) detectar dicha emisión de fluorescencia; y c) repetir las etapas a) y b). La etapa de irradiación se lleva a cabo en presencia de una solución de detección que incluye un compuesto polifenólico. La solución de detección inhibe la degradación inducida por la luz del ácido nucleico. En algunas realizaciones, la solución de detección incluye ácido gálico, un éster alquílico inferior del mismo, o mezclas de los mismos. En otras realizaciones, la solución de detección incluye una mezcla de 1) ácido gálico, un éster alquílico inferior del mismo, o mezclas de los mismos y 2) uno o más compuestos seleccionados entre urea, ácido ascórbico o una sal del mismo y ácido isoascórbico o sal del mismo.

30 Los métodos de la presente solicitud incluyen una solución para uso durante una etapa de irradiación. La solución incluye un compuesto polifenólico que puede ser cualquier sistema aromático que tenga dos o más grupos hidroxilo fenólicos, cualquiera de los cuales puede ser también un éter alquílico inferior. Los compuestos polifenólicos pueden pertenecer a cualquier número de clases estructurales tales como lignanos, taninos, galocatequinas y flavonoides, tales como flavonoles, flavonas, catequinas, flavanonas, antocianidinas e isoflavonoides. Los compuestos polifenólicos de la presente invención pueden existir en una forma glicosilada con uno o más restos de azúcar unidos al compuesto polifenólico. Tales formas glicosiladas pueden conferir propiedades útiles de solubilidad en agua, por ejemplo. Compuestos polifenólicos modelo incluyen, sin limitación, apigenina, astragalina, aurantinidina, azaleatina, butina, ácido cafeico, catequina, cianidina, epicatequina, epigalocatequina, galocatequina, ácido gálico y sus ésteres alquílicos inferiores, pirogalol, delphinidina, ácido elágico, eriodictilol, homoeriodictilol, europinidina, fisetina, ácido ferúlico, galangina, genisteína, gossipetina, hesperidina, hidroquinonas, tales como p-hidroquinona, t-butilhidroquinona (TBHQ), 2,4,5-trihidroxi-butirofenona (THBP), hidroxitiroso, isorhamnetina, isosacuranetina, kaempferol, kaempferide, luteolina, luteolinidina, malvidina, miricetina, morina, naringenina, naringina, natsudaidina, paquipodol, pelargonidina, peonidina, petunidina, floroglucinol, pinocembrina, poncirin, pterocarpans, pirocatecol, quercetina, resorcinol, rhamnacina, rhamnetina, rosinidina, rutina, sinapil, alcohol, sakuranetina, sakuranina, esterubina y ácido tánico. En algunas realizaciones, el compuesto polifenólico puede incluir ácido gálico, un éster alquílico inferior del mismo, tal como galato de metilo, galato de etilo, galato de propilo, éteres monometílicos de los mismos y combinaciones de ésteres alquílicos inferiores y éteres monometílicos de los mismos, o mezclas de los mismos.

35 Opcionalmente, en las soluciones de los métodos proporcionados, al menos uno de uno o más compuestos es ácido ascórbico o una sal, análogo o derivado del mismo. El uso de ascorbato en una solución de detección para secuenciación se ha descrito en el documento WO2006/064199 de Fedurco et al.,. A menos que se indique lo contrario, los términos ascorbato y ácido ascórbico, tal como se utilizan en la presente memoria, se refieren tanto al isómero L como al isómero D, y mezclas, que incluyen mezclas racémicas de las mismas. Ambos isómeros ópticos son capaces de funcionar como compuestos en los métodos proporcionados. Una sal modelo incluye ascorbato sódico, p.ej. L-ascorbato sódico. Existen numerosos análogos y derivados de ascorbato conocidos que tienen

actividad que se pueden utilizar en los métodos proporcionados. Derivados y análogos adecuados incluyen aquellos en los que el grupo 5- y/o 6-hidroxi está esterificado o modificado de otro modo. Alternativamente, el grupo 5- y/o 6-hidroxi se puede reemplazar por un grupo funcional alternativo, tal como halo o amino. Otros derivados son aquellos en los que el grupo 5- y/o 6-hidroxi está ausente y (es decir, con un átomo de hidrógeno en lugar del grupo hidroxilo). Ejemplos representativos de tales derivados incluyen, pero no se limitan a, 6-O-tosil-L-ascorbato, 5-desoxi-L-ascorbato, 6-bromo-6-desoxi-L-ascorbato, 6-amino-6-desoxi-L-ascorbato, 6-carboxilato de ácido L-ascórbico y alcanosatos de 6-O-ascorbilo tales como palmitato de 6-ascorbilo (ascorbato de palmitoilo).

Como se ha mencionado anteriormente, opcionalmente, al menos uno de uno o más compuestos es ácido gálico, un derivado o análogo del mismo o combinaciones del mismo. El uso de ácido gálico en una solución de detección para secuenciación se ha descrito en la Patente de EE.UU. N° 9.115.353. Los derivados incluyen, pero no se limitan a, ésteres de alquilo inferiores o monometil éteres de ácido gálico. Como se utiliza en la presente memoria, el término éster alquílico inferior se refiere a un éster de cadena alquilo C1-C6 de un ácido carboxílico. En algunas realizaciones, un éster alquílico inferior se refiere a un éster de cadena alquilo C1-C4 de un ácido carboxílico. Los ésteres representativos incluyen ésteres de metilo, etilo, propilo, butilo, pentilo y hexilo. Cualquiera de los ésteres anteriores puede estar opcionalmente ramificado. Tales ésteres ramificados incluyen ésteres de isopropilo, ésteres de sec-butilo, ésteres de isobutilo y ésteres de terc-butilo, por ejemplo.

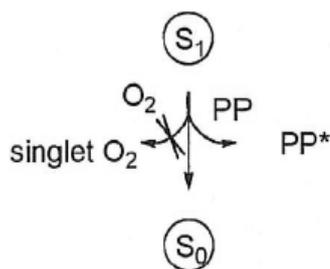
También se reconocerá que otros derivados éster de ácido gálico pueden ser útiles en la invención donde se confiere suficiente hidrosolubilidad al éster de galato. En algunas realizaciones, la solubilidad adicional se puede realizar mediante el uso de una pequeña cantidad de un codisolvente tal como dimetilsulfóxido (DMSO). Mediante el empleo de tales codisolventes en pequeñas cantidades, el uso de cualquier éster de galato puede ser eficaz para reducir la degradación inducida por la luz de ácidos nucleicos. En algunas realizaciones, se puede emplear un grupo éster hidrosoluble, tal como un éster de PEG de ácido gálico. En tales realizaciones, el éster de galato puede beneficiarse de la actividad antioxidante mejorada de un éster sobre un ácido carboxílico, sin sacrificar la hidrosolubilidad útil del grupo funcional de ácido carboxílico.

Opcionalmente, se puede utilizar una combinación de dos o más compuestos en la solución. Preferiblemente, al menos uno de los compuestos en tales combinaciones es ácido gálico o un análogo o derivado del mismo.

El uno o más compuestos o compuestos estarán presentes en la solución en un intervalo de concentración de entre aproximadamente 0,1 mM a aproximadamente 5 M, inclusive. Por ejemplo, uno o más compuestos o compuestos estarán presentes en la solución en un intervalo de concentración que incluye 0,1 mM, 0,5 mM, 1 mM, 5 mM, 10 mM, 20 mM, 30 mM, 40 mM, 50 mM, 60 mM, 70 mM, 80 mM, 90 mM, 100 mM, 200 mM, 300 mM, 400 mM, 500 mM, 1M, 2M, 3M, 4M y 5M, y cualquier concentración entre aproximadamente 0,1 mM y aproximadamente 5 M, inclusive. Por ejemplo, cuando se emplea ácido gálico, sus análogos, derivados o mezclas de los mismos, este componente de la solución puede estar presente en una concentración comprendida entre aproximadamente 0,1 mM y aproximadamente 200 mM.

En algunas realizaciones, la degradación puede incluir daño oxidativo, tal como se describe en Cooke et al. FASEB J. 17(10):1195-1214(2003). Por ejemplo, en el caso de degradación inducida por la luz, puede incluir la escisión de la estructura fosfodiéster en cualquiera de los nucleótidos de un ácido nucleico. Otras vías de degradación pueden incluir la formación de 8-oxo-2'-desoxiguanosina, la eliminación o reacción de una base de ácido nucleico o etiqueta fluorescente que causa una pérdida en la función de hibridación o fluorescencia y la reticulación fotoinducida de nucleótidos. La degradación por especies de radicales de oxígeno reactivos tales como el radical hidroxilo pueden dañar los ácidos nucleicos mediante procesos de abstracción de átomos de hidrógeno, tales como del grupo metilo de timina o de los enlaces C-H de la 2'-desoxirribosa. Las especies de radicales de oxígeno también pueden oxidar nucleótidos por reacción con insaturaciones en bases de ácidos nucleicos, por ejemplo.

La solución que tiene un compuesto polifenólico puede inhibir el daño inducido por la luz de ácidos nucleicos mediante diversos mecanismos. Sin estar limitado por la teoría, el compuesto polifenólico puede actuar como un desactivador de fluorescencia más rápido del fluoróforo excitado que el oxígeno, reduciendo o impidiendo así la formación de oxígeno singlete como se indica en la Ecuación 1:



Ecuación 1

Como se indica en la ecuación 1, el fluoróforo  $S_1$  excitado de estado singlete puede reaccionar preferentemente con un compuesto polifenólico (PP) en lugar de la reacción con oxígeno, para proporcionar un retorno del fluoróforo al estado fundamental singlete  $S_0$  mientras se forma un estado excitado del compuesto polifenólico (PP\*). Un experto en la técnica reconocerá que el retorno al estado fundamental  $S_0$  puede alternativamente proceder a través del cruce entre sistemas a un estado excitado triplete del fluoróforo  $T_1$  (no mostrado). Esta vía no radiativa al estado triplete del fluoróforo puede ser promovida por interacción con el compuesto polifenólico. En última instancia, la energía transferida al compuesto polifenólico se puede disipar eficazmente.

Alternativamente, cualquier oxígeno singlete generado mediante desactivación de fluorescencia con oxígeno puede reaccionar con el compuesto polifenólico para preservar la integridad de los ácidos nucleicos. De forma similar, cualquier especie de oxígeno reactivo tal como formación de anión superóxido o radical hidroxilo formada mediante absorción de fotón múltiple, como se describe anteriormente, también se puede extinguir mediante el compuesto polifenólico presente en la solución de detección. Otro posible mecanismo mediante el cual el compuesto polifenólico puede reducir o prevenir la degradación inducida por la luz es interceptando y disipando cualquier radiación UV visible de alta energía generada por absorción de fotón múltiple por el fluoróforo. En tal mecanismo, se reduce o se evita la formación de especies reactivas de oxígeno y/o la dimerización indirecta de pirimidina causada por una emisión de alta energía. Finalmente, un experto en la técnica apreciará que estos procesos fotoquímicos de alta energía pueden operar en cualquier combinación.

Como se describe anteriormente, la inhibición de estos procesos de degradación inducidos por la luz se refiere a un bloqueo total o parcial del grado de daño, por ejemplo, tal como se puede cuantificar mediante la intensidad observada de la emisión fluorescente con exposición repetida a iluminación intensa. La degradación inducida por la luz se puede medir, por ejemplo, como una función de la intensidad de la señal de fluorescencia frente al número de etapas de irradiación repetidas (detección) realizadas en la matriz de ácidos nucleicos tal como en la secuenciación mediante ejemplos de síntesis descritos en la presente memoria más adelante. Este proceso a veces se conoce como decaimiento de intensidad. Otra evaluación de daño ácido nucleico inducido por la luz se puede medir como una función de la tasa de error frente al número de etapas de irradiación repetidas (detección) realizadas en una matriz de ácidos nucleicos.

Otra evaluación adicional de daño inducido por la luz puede ser para evaluar productos de degradación de ácidos nucleicos. Por ejemplo, la medición de 8-oxo-2 'desoxiguanosina se utiliza comúnmente para evaluar el daño oxidativo a los ácidos nucleicos. El daño también se puede evaluar cuantificando el grado de homo y heterodimerización de timina y citosina. Cuando se emplean métodos de detección en soporte sólido, la inhibición de daño se puede evaluar fácilmente comparando matrices irradiadas en presencia y ausencia de la solución de detección y midiendo una pérdida de intensidad de señal de fluorescencia con ciclos repetidos de irradiación.

En algunos aspectos, los métodos de la presente descripción se pueden aplicar a ácidos nucleicos en solución, mientras que en otros aspectos, los métodos de la descripción se pueden aplicar a ácidos nucleicos unidos a un soporte, tal como en un formato de matriz. Cuando se emplean ácidos nucleicos sobre un soporte, los ácidos nucleicos pueden estar presentes en matrices ordenadas o no ordenadas. Cuando se utiliza la solución que tiene uno o más de los compuestos proporcionados, se puede reducir la cantidad de ácidos nucleicos escindida de un soporte. Así, en algunos aspectos, la degradación inducida por la luz incluye eliminación de un miembro de ácido nucleico de una matriz de ácidos nucleicos.

Opcionalmente, la etapa de procesamiento es una etapa de procesamiento de detección y la etapa de procesamiento de detección puede incluir irradiación de los ácidos nucleicos. Cuando se realiza una etapa de detección, la irradiación de los ácidos nucleicos incluye el uso de una fuente de luz apropiada para excitar una etiqueta, tal como una etiqueta de fluorescencia. Un experto en la técnica reconocerá que hay numerosas etiquetas disponibles y que las condiciones de la irradiación, tales como elección de longitud de onda para irradiación y detección, se guiarán mediante la elección de etiquetas que se empleen en el proceso de detección de ácido nucleico. Las etiquetas pueden ser las mismas para cada tipo de nucleótido, o cada tipo de nucleótido puede tener una etiqueta única. La etiqueta se utiliza en la identificación de una incorporación particular de nucleótido dentro del ácido nucleico. Así, por ejemplo, la adenina modificada, la guanina, la citosina y la timina pueden haber unido un fluoróforo diferente para permitir que sean discriminados entre sí fácilmente. Cuando se secuencia sobre matrices, se puede utilizar una mezcla de nucleótidos etiquetados y sin etiquetar. Las etiquetas adecuadas incluyen, pero no se limitan a, etiquetas fluorescentes, etiquetas de masa, etiquetas magnéticas y similares. A modo de ejemplo, las etiquetas incluyen biotina, dinitrofenol y fluoresceína. Los nucleótidos etiquetados modelo para su uso en los presentes métodos se describen en el documento WO 04/018497 y en las patentes de EE.UU. N° 7.541.444 y 7.057.026.

Etiquetas detectables tales como fluoróforos pueden estar enlazados a nucleótidos a través de la base del ácido nucleico que utiliza un enlazador adecuado. Los fluoróforos adecuados incluyen los descritos en el documento WO 2007/135368. El enlazador puede ser lábil a los ácidos, fotolábil o contener un enlace disulfuro. Etiquetas y enlaces modelo incluyen los descritos en el documento WO03/048387. En la invención se pueden emplear otros enlaces, en particular enlazadores que contienen azida escindible con fosfina, como se describe en el documento WO2004/018493.

La irradiación se lleva a cabo con una fuente de irradiación que puede incluir, por ejemplo, un láser u otra fuente para generar suficiente intensidad de luz de excitación para producir emisión detectable. La fuente de irradiación se puede utilizar junto con cualquier combinación de filtros de longitud de onda. La etapa de irradiación puede ser pulsada en algunas realizaciones, o continua en otras realizaciones. La intensidad de una fuente de luz puede estar en un intervalo entre aproximadamente 5 milivatios y aproximadamente 500 milivatios. La fuente de luz puede ser una lámpara de descarga de arco corto de alta energía, por ejemplo. Las lámparas de descarga modelo incluyen quemadores de mercurio, que varían en vatios de 50 a 200 vatios, y los quemadores de xenón que varían de 75 a 150 vatios, y LEDs. La elección exacta de intensidad de la fuente de luz puede depender del coeficiente de desactivación del fluoróforo. Los mayores coeficientes de desactivación indican que la absorción de un fotón, también denominada cuanto, en una región de longitud de onda dada es más probable. El rendimiento cuántico indica la relación entre fotones emitidos y fotones absorbidos, y es usualmente un valor entre 0,1 y 1,0. El rendimiento cuántico es una medida de la eficiencia de las emisiones. Los valores de rendimiento cuántico inferiores a 1 resultan de la pérdida de energía a través de vías no radiactivas, como calor o una reacción fotoquímica, en lugar de emisión de fluorescencia. El coeficiente de desactivación, el rendimiento cuántico, la intensidad de la fuente de luz y la vida útil de la fluorescencia son factores importantes que contribuyen a la intensidad y utilidad de la emisión de fluorescencia.

Métodos de la invención pueden incluir una etapa de irradiación llevada a cabo en un intervalo de aproximadamente 360 nm a aproximadamente 800 nm, con una fuente de luz que tiene potencia en un intervalo entre aproximadamente 5 a aproximadamente 500 m vatios. Como se ha descrito anteriormente, entre aproximadamente 360 y aproximadamente 800 nm, la degradación inducida por la luz se puede mejorar por la presencia de la solución de detección. Esto facilita el uso de múltiples fluoróforos para la detección simultánea de emisiones fluorescentes. Así, el método de la invención se puede llevar a cabo mientras se incorporan las cuatro bases comunes con etiquetas diferenciales. La etapa de irradiación se puede llevar a cabo durante un periodo de tiempo de aproximadamente 0,1 segundos a aproximadamente 10 minutos. 175-500 ms sobre GA, velocidad de escaneo de 2 mm/s. Tales tiempos de irradiación pueden incluir irradiación pulsada o continua.

Cuando se realizan múltiples etapas que incorporan nucleótidos etiquetados diferenciadamente, los métodos de la invención pueden incluir reemplazar una solución con la solución proporcionada (p.ej., solución de detección) antes de la etapa de irradiación. Por ejemplo, en la secuenciación por síntesis, la solución utilizada para añadir el nucleótido etiquetado se puede reemplazar con la solución de detección antes de la irradiación. En otras realizaciones, la solución de detección se puede utilizar a lo largo de cada uno de las etapas en la secuenciación por síntesis, incluyendo etapas de lavado, además de la etapa de detección.

La luz fluorescente emitida desde el fluoróforo se puede detectar a la longitud de onda apropiada que utiliza un sistema de detección adecuado tal como, por ejemplo, una cámara de dispositivo de carga acoplada (CCD), que puede estar opcionalmente acoplada a un dispositivo de aumento, a una cámara fluorescente o a un microscopio confocal. Un sistema de detección alternativo adecuado emplea un detector semiconductor complementario de óxido metálico (CMOS). Si se lleva a cabo la secuenciación sobre una matriz, la detección de una base incorporada se puede llevar a cabo utilizando un microscopio de exploración confocal para escanear la superficie de la matriz con un láser, para captar la imagen de etiquetas fluorescentes unidas a los nucleótidos incorporados. Alternativamente, se puede utilizar un detector sensible 2-D, tal como un detector de carga acoplada (CCD), para visualizar las señales generadas. Esta técnica es particularmente útil con matrices de una sola molécula. Otras técnicas como la microscopía óptica de barrido de campo cercano (SNOM) están disponibles y se pueden utilizar cuando se captan imágenes de matrices densas. Para una descripción de la microscopía óptica de barrido de campo cercano, véase Moyer et al., *Laser Focus World* 29:10, 1993. Una técnica adicional que se puede utilizar es la microscopía de fluorescencia de reflexión interna total (TIRFM) específica de superficie; Véase, por ejemplo, Vale et al., *Nature*, (1996) 380: 451 - 453).

Cuando se emplean soluciones de la invención para la secuenciación por síntesis, por ejemplo, el método puede incluir añadir un nucleótido adicional etiquetado con fluorescencia a la matriz y repetir las etapas de detección de cada ciclo. Los métodos de la invención que emplean etapas de adición y detección de nucleótidos repetidas pueden incluir al menos 25 ciclos en algunas realizaciones, al menos 75 ciclos en otras realizaciones y al menos 100 ciclos en otras realizaciones. Los métodos de la invención incluyen repetir las etapas de adición y detección para un número de ciclos en un intervalo de entre aproximadamente 100 ciclos a aproximadamente 1.000 ciclos, en algunas realizaciones, entre aproximadamente 100 ciclos y aproximadamente 500 ciclos, en otras realizaciones y entre aproximadamente 100 ciclos hasta aproximadamente 300 ciclos, incluso en otras realizaciones.

La capacidad de secuenciar con precisión 25 o más, 50 o más, 75 o más, o 100 o más nucleótidos consecutivos en una reacción de secuenciación es una ventaja significativa en aplicaciones tales como la realineación del genoma.

En algunas realizaciones, una matriz de ácidos nucleicos incluye un molde de cebador. La "secuenciación por síntesis" de ácidos nucleicos implica la adición secuencial de uno o más nucleótidos a una cadena polinucleotídica creciente en la dirección 5' a 3' usando una polimerasa para formar una cadena polinucleotídica extendida complementaria a un ácido nucleico molde a secuenciar. Se describen métodos de secuenciación modelo, por ejemplo, en Bentley et al., en *Nature* 456:53-59 (2008), en los documentos WO 04/018497; US 7057026; WO 91/06678; WO 07/123744; US 7 329 492; US 7.211.414; US 7.315.019; US 7.405.281 y US 2008/010802. La

identidad de la base presente en uno o más de los nucleótidos añadidos se determina en la etapa de detección. La identidad de la base añadida se puede determinar después de cada etapa de incorporación de nucleótidos. La secuencia del molde se puede inferir utilizando las reglas convencionales de emparejamiento de bases de Watson-Crick. Para evitar dudas, la "secuenciación" también puede abarcar la incorporación e identificación de un único nucleótido. La determinación de la identidad de una base única puede ser útil, por ejemplo, en la puntuación de polimorfismos de un solo nucleótido.

El molde de ácido nucleico para una reacción de secuenciación puede incluir una región bicatenaria que tiene un grupo hidroxilo 3' libre que sirve como un cebador o punto de iniciación para la adición de más nucleótidos en la reacción de secuenciación. La región del molde a secuenciar sobrepasará este grupo hidroxilo 3' libre en la cadena complementaria. El cebador que lleva el grupo hidroxilo 3' libre se puede añadir como un componente separado (p.ej., un cebador de secuenciación de oligonucleótidos convencional) que hibrida con una región del molde a secuenciar. Alternativamente, el cebador y la cadena de molde a secuenciar pueden formar cada uno parte de una hebra de ácido nucleico parcialmente complementaria capaz de formar un dúplex intramolecular, tal como por ejemplo una estructura de bucle en horquilla. Los nucleótidos se añaden sucesivamente al grupo hidroxilo 3' libre, que resulta en la síntesis de una cadena polinucleotídica en la dirección 5' a 3'. Después de cada adición de nucleótidos se puede determinar la naturaleza de la base que se ha añadido, proporcionando así información de secuencia para el molde de ácido nucleico.

La incorporación de un nucleótido en una cadena de ácido nucleico (o polinucleótido) se refiere a la unión del nucleótido al grupo hidroxilo 3' libre de la cadena de ácido nucleico a través de la formación de un enlace fosfodiéster con el grupo 5' fosfato del nucleótido. El molde de ácido nucleico a secuenciar puede ser ADN o ARN, o incluso una molécula híbrida que incluye tanto desoxinucleótidos como ribonucleótidos. El ácido nucleico puede incluir nucleótidos de origen natural y/o no natural y enlaces de la estructura naturales o no naturales.

Los moldes de ácido nucleico a secuenciar se pueden unir a un soporte sólido a través de cualquier método de unión adecuado conocido en la técnica. La unión se puede realizar a través de un enlace covalente, por ejemplo. Si los moldes están "agrupados" en un soporte sólido, entonces la matriz puede tomar cualquier forma conveniente. Así, el método de la invención es aplicable a todos los tipos de matrices de "alta densidad", que incluye matrices de una sola molécula y matrices agrupadas.

En algunas realizaciones, la solución reduce una tasa de error de detección, cuando se secuencia, superior al 20% con respecto a un control que carece del compuesto. La tasa de error de detección se reduce adicionalmente en presencia de uno o más compuestos seleccionados entre urea, ácido ascórbico o una sal del mismo, y ácido isoascórbico o sal del mismo en la solución. En otras realizaciones, la solución reduce la tasa de error de detección superior al 40% con respecto a un control que carece del compuesto polifenólico, tasa de error que también se incrementa por la presencia de uno o más compuestos seleccionados entre urea, ácido ascórbico o sal, y ácido isoascórbico o sal del mismo. En otras realizaciones adicionales, la solución reduce una tasa de error de detección superior al 50% con respecto a un control que carece del compuesto polifenólico.

Las tasas de error se pueden determinar, por ejemplo, como se describe en Bentley, et al., "Accurate whole human genome sequencing using reversible terminator chemistry" *Nature* 456:53-59 (2008). La tasa de error es el error de secuenciación por ciclo como se determina mediante alineación de la secuencia phiX frente a un genoma estándar phiX que utiliza el algoritmo ELAND que es parte del análisis de tubería estándar como se describe en la Pipeline User Guide (Illumina, Inc., San Diego, CA) y Bentley et al., *Nature* 456:53-59 (2008).

En algunas realizaciones, la presente descripción proporciona un método para inhibir la degradación inducida por la luz de ácidos nucleicos durante una etapa de detección que incluye irradiar una porción de los ácidos nucleicos en presencia de una solución que incluye ácido gálico, análogos, derivados, un éster alquílico inferior de los mismos o mezclas de los mismos. La solución reduce la cantidad de degradación inducida por la luz de los ácidos nucleicos. El uso de ácido gálico y sus derivados se muestran a continuación en los Ejemplos I y II. El ácido gálico en su forma libre es capaz de ionizarse a carboxilato y su eficacia como inhibidor de degradación inducida por la luz puede, por tanto, ser dependiente del pH del sistema. Por comparación, los ésteres de galato inferiores pueden demostrar una eficacia relativamente constante sobre un intervalo de pH.

Sin estar limitado por la teoría, un experto en la técnica reconocerá que un grupo de ácido carboxílico es modestamente más electrón atrayente que un éster. Así, si el ácido gálico y compuestos relacionados están proporcionando un papel antioxidante, se podría esperar que un éster funcione ligeramente mejor que un ácido carboxílico reduciendo la degradación inducida por la luz. Sin embargo, este beneficio de ésteres de galato se puede mejorar por la necesidad de hidrosolubilidad del éster. A este respecto, el ácido carboxílico ionizable es útil para conferir hidrosolubilidad. Como tal, en algunas realizaciones, las soluciones de detección de la presente invención proporcionan ácido gálico, ésteres de alquilo inferior de ácido gálico o mezclas de los mismos.

Un experto en la técnica también reconocerá que otros derivados de éster de ácido gálico pueden ser útiles en la invención donde se confiere suficiente hidrosolubilidad al éster de galato. En algunas realizaciones, la solubilidad adicional se puede realizar mediante el uso de una pequeña cantidad de un codisolvente tal como sulfóxido de dimetilo (DMSO). Mediante el empleo de tales codisolventes en pequeñas cantidades, el uso de cualquier éster de

galato puede ser eficaz en reducir la degradación inducida por la luz de ácidos nucleicos. En algunas realizaciones, se puede emplear un grupo éster hidrosoluble, tal como un éster de PEG de ácido gálico. En tales realizaciones, el éster de galato se puede beneficiar de la actividad antioxidante mejorada de un éster sobre un ácido carboxílico, sin sacrificar la hidrosolubilidad útil del grupo funcional de ácido carboxílico.

5 En otros aspectos más, la presente descripción proporciona un método para inhibir la degradación inducida por la luz de ácidos nucleicos durante una etapa de detección que incluye los ácidos nucleicos en presencia de una solución de detección que incluye 1) ácido gálico, un éster alquílico inferior del mismo o mezclas de los mismos, y 2) uno o más compuestos seleccionados entre urea, ácido ascórbico o sal del mismo, y ácido isoascórbico o sal del mismo. Esta solución de detección reduce la cantidad de degradación inducida por la luz de los ácidos nucleicos. Como se demuestra en los Ejemplos I y II, la urea ha mostrado un sinergismo en la reducción de la degradación inducida por la luz de ácidos nucleicos. La urea no sustituida, (NH<sub>2</sub>)<sub>2</sub>CO, no se clasifica generalmente como un antioxidante, y así el papel de la urea proporcionando un efecto sinérgico reduciendo la degradación inducida por la luz no es fácilmente atribuible a la actividad antioxidante. Por ejemplo, se ha indicado que la urea no proporciona protección aparente contra especies reactivas de oxígeno a concentraciones fisiológicas (Glazer, FASEB, J. 2: 2487 - 2491 (1988)). Además, se ha indicado que la urea no proporciona protección aparente contra la reacción con oxígeno singlete (Dahl et al., Photochem., Photobiol., 47 (3): 357-362 (1988)). La urea absorbe en la región UV-visible sustancialmente sólo por debajo de 245 nm. En esta región del espectro UV, la dimerización de pirimidina es una vía significativa con respecto al daño de ácido nucleico inducido por la luz. Sin embargo, el papel de la urea proporcionando protección adicional contra el daño inducido por la luz bajo las condiciones de detección de fluorescencia, como se proporciona en los Ejemplos, aún no se entiende completamente.

En algunos aspectos, la presente descripción se dirige a un kit para uso de acuerdo con los métodos antes mencionados. Así, se proporciona en la presente memoria un kit que puede incluir uno o más de los compuestos y (i) anticuerpos, y/o (ii) enzimas descritas en la presente memoria y combinaciones de las mismas. Los compuestos pueden incluir ácido gálico, análogos, derivados, un éster alquílico inferior del mismo, o mezclas de los mismos. En algunas realizaciones, el kit puede incluir además compuestos adicionales, tampones u otros agentes y/o un conjunto de instrucciones para llevar a cabo la preparación de una solución para utilizar en la preparación de muestras de ácido nucleico. Adicionalmente o alternativamente, el kit puede incluir uno o más nucleótidos, una enzima capaz de catalizar la incorporación de los nucleótidos en una cadena de ácido nucleico complementaria a un molde de ácido nucleico a secuenciar y un compuesto polifenólico adecuado para preparar una solución. El compuesto polifenólico puede incluir ácido gálico, análogos, derivados, un éster alquílico inferior del mismo, o mezclas de los mismos. En algunos aspectos, el kit incluye además un compuesto secundario seleccionado entre urea, ácido ascórbico o una sal del mismo, y ácido isoascórbico o sal del mismo. El kit puede incluir adicionalmente sales tampón y un conjunto de instrucciones para llevar a cabo la preparación de una solución para uso en experimentos de fluorescencia.

Los aspectos descritos en la presente memoria se pueden utilizar en una variedad de métodos y/o composiciones conocidos para detectar ácidos nucleicos, en donde ocurre o se sospecha que ocurra la degradación inducida por la luz de ácidos nucleicos. Por ejemplo, los métodos y composiciones descritos en la presente memoria son particularmente útiles cuando un método de detección requiere exposición repetida o prolongada de ácidos nucleicos a la luz. En particular, la detección de ácidos nucleicos etiquetados con fluorescencia se utiliza a menudo irradiando una muestra con luz que tiene una longitud de onda adecuada para inducir una emisión fluorescente desde unos ácidos nucleicos que contienen una etiqueta fluorescente. Durante las etapas de irradiación de estos métodos, como se describe en la presente memoria, se puede producir daño a los ácidos nucleicos. La exposición repetida o prolongada puede resultar en una disminución en la sensibilidad de detección de fluorescencia, que se puede manifestar, por ejemplo, como un aumento en las tasas de error de detección y una reducción en las relaciones señal-ruido. Ejemplos no limitantes de métodos en donde los métodos y/o composiciones para detectar ácidos nucleicos que tienen exposición repetida o prolongada a luz incluyen técnicas de alto rendimiento o secuenciación rápida tales como secuenciación por síntesis y secuenciación por ligación, técnicas de detección de micromatriz de ácido nucleico tales como chips de genes y micro-matrices de ADN y técnicas cuantitativas de reacción en cadena de la polimerasa (Q-PCR) tales como reacción en cadena de polimerasa en tiempo real (PCR) y la PCR multiplex.

Un método útil para secuenciación rápida o de alto rendimiento es secuenciación por síntesis (SBS). Las técnicas de SBS que requieren irradiación repetida o prolongada de ácidos nucleicos con luz incluyen, pero no se limitan a, los sistemas Genome Analyzer (Illumina Inc., San Diego, CA) y los sistemas True Single Molecule Sequencing (tSMS)<sup>TM</sup> (Helicos BioSciences Corporation, Cambridge, MA). Brevemente, se utilizan varias reacciones de secuenciación por síntesis para dilucidar la identidad de una pluralidad de bases en posiciones diana dentro de una secuencia diana. Todas estas reacciones dependen del uso de una secuencia de ácido nucleico diana que tiene al menos dos dominios; un primer dominio al que se hibridará un cebador de secuenciación, y un segundo dominio adyacente, para el que se desea información de secuencia. Tras la formación de un complejo de ensayo, se utilizan enzimas de extensión para añadir dNTPs a un cebador de secuenciación que se hibrida con el primer dominio y cada adición de dNTPs se lee para determinar la identidad del dNTP añadido. Esto puede continuar durante muchos ciclos. Técnicas SBS tales como, los sistemas Genome Analyzer (Illumina Inc., San Diego, CA) y los sistemas True Single Molecule Sequencing (tSMS)<sup>TM</sup> (Helicos BioSciences Corporation, Cambridge, MA), utilizan nucleótidos etiquetados para determinar la secuencia de una molécula de ácido nucleico diana. Una molécula de ácido nucleico diana se puede

hibridar con un cebador e incubar en presencia de una polimerasa y un nucleótido etiquetado que contiene un grupo bloqueante. El cebador se extiende de manera que se incorpora el nucleótido. La presencia del grupo bloqueante permite solamente una ronda de incorporación, es decir, la incorporación de un solo nucleótido. La presencia de la etiqueta permite la identificación del nucleótido incorporado. Se puede añadir una pluralidad de bases de un solo nucleótido homogéneas durante cada ciclo, tal como se utilizan en los sistemas True Single Molecule Sequencing (tSMS)<sup>TM</sup> (Helicos BioSciences Corporation, Cambridge, MA) o, alternativamente, se pueden añadir las cuatro bases de nucleótidos durante cada ciclo simultáneamente, tal como se utiliza en los sistemas Genome Analyzer (Illumina Inc., San Diego, CA), particularmente cuando cada base está asociada con una etiqueta distinguible. Después de identificar el nucleótido incorporado mediante su etiqueta correspondiente, tanto la etiqueta como el grupo bloqueante pueden ser eliminados, permitiendo de este modo una ronda subsiguiente de incorporación e identificación. La determinación de la identidad de la base de nucleótido añadida incluye exposición repetida de las bases etiquetadas recién añadidas a una fuente de luz que puede inducir una emisión detectable debida a la adición de una base de nucleótido específica, es decir, dATP, dCTP, dGTP o dTTP. Los métodos y composiciones descritos en la presente memoria son particularmente útiles para tales técnicas SBS.

Otro método útil para la técnica de secuenciación rápida o de alto rendimiento es la secuenciación por ligación. La secuenciación por ligación es un método bien conocido para la secuenciación que requiere la irradiación repetida o prolongada de las sondas di-base con luz. Los sistemas ilustrativos que utilizan la secuenciación por síntesis incluyen el sistema SOLiD<sup>TM</sup> de Applied Biosystems (Life Technologies, Carlsbad, CA). Brevemente, los métodos para la secuenciación por ligación incluyen la hibridación de cebadores de secuenciación con secuencias adaptadoras inmovilizadas a perlas templadas. Un conjunto de cuatro sondas di-base etiquetadas con fluorescencia compiten por la ligación al cebador de secuenciación. La especificidad de la sonda di-base se logra interrogando cada primera y segunda base en cada reacción de ligación. Después de una serie de ciclos de ligación, el producto de extensión se retira y el molde se reinicia con un cebador de secuenciación complementario a la posición n-1 para una segunda ronda de ciclos de ligación. Se realizan múltiples ciclos de ligación, detección y escisión con el número de ciclos que determinan la longitud de lectura final. La secuenciación por métodos de ligación se han desarrollado por Applied Biosystems en su plataforma Agencourt (véase Ronaghi et al., *Science* 281: 363 (1998), Dressman et al., *Proc. Nat. Acad. Sci. USA* 100: 8817-8822 (2003). Mitra et al., *Proc. Nat. Acad. Sci. USA* 100: 55926 - 5931 (2003)).

Además, los métodos y soluciones descritos en la presente memoria pueden ser particularmente útiles para la secuenciación a partir de una matriz de ácidos nucleicos, donde múltiples secuencias se pueden leer simultáneamente desde múltiples posiciones en la matriz ya que cada nucleótido en cada posición se puede identificar en base a su etiqueta identificable. Se describen métodos modelo en el documento US 2009/0088327; US 2010/0028885; Y US 2009/0325 172.

Otros métodos conocidos en la técnica donde la irradiación repetida o prolongada de ácidos nucleicos con luz incluye técnicas de detección de micromatriz de ácido nucleico, tales como chips de genes y micromatrices de ADN. Es bien conocido en la técnica que existen problemas de fiabilidad y consistencia cuando se escanean micromatrices de ácido nucleico, particularmente cuando se escanea la misma micromatriz más de una vez. Sin embargo, se pueden requerir múltiples escaneos para obtener el intervalo dinámico completo de los ácidos nucleicos etiquetados, por ejemplo, cuando se usa una micromatriz para determinar los niveles de expresión génica. Un solo escaneo intenta capturar todo el intervalo de expresión en las muestras dadas. Esto puede no dar la verdadera imagen de la expresión de todo el conjunto de genes cuando un amplio intervalo de expresión está presente. Por ejemplo, un gen en la muestra puede expresar tan sólo 200 copias, mientras que un gen separado en la misma muestra puede expresar 50.000 copias. En este aspecto, los métodos y composiciones descritos en la presente memoria son particularmente útiles para mantener la integridad de los ácidos nucleicos en múltiples escaneos.

Ejemplos de técnicas de detección de micromatrices de ácidos nucleicos conocidas en la técnica incluyen, pero no se limitan a, LabCard (ACLARA Bio Sciences Inc., Santa Clara, CA); GeneChip (Affymetrix, Inc, Santa Clara, CA); LabChip (Caliper Technologies Corp., Hopkinton, MA); micromatrices producidas por la tecnología SurePrint (Agilent Technologies, Santa Clara, CA); una matriz de baja densidad con detección electroquímica (Clinical Micro Sensors Inc., Pasadena, CA); LabCD System (Tecan Trading AG, Zurich, Suiza); Omni Grid (Gene Machines, Stillwater, OK); Q Array (Genetix Ltd., Boston, MA); Sistemas de espectrometría de masas automatizados de alto rendimiento con tecnología de expresión en fase líquida (GeneTrace Systems, Inc., Menlo Park, CA); un sistema de puntos por inyección térmica (Hewlett Packard Company, Palo Alto, CA); HySeq HyChip (Hyseq, Inc., Sunnyvale, CA); BeadArray (Illumina, Inc., San Diego, CA); GEM (Incyte Microarray Systems, Fremont, CA); un sistema de micromatriz de alto rendimiento que puede dispensar de 12 a 64 puntos en múltiples láminas de vidrio (Intelligent Bio-Instruments, Waltham, MA); Molecular Biology Workstation y NanoChip (Nanogen, Inc., San Diego, CA); un chip de vidrio microfluídico (Orchid Cellmark, Inc., Dayton, OH); un robot manipulador de biochips (*Biochip Arrayer*) con cuatro puntas piezoeléctricas PiezoTip de goteo a demanda (Packard Instruments, Inc., Meriden, CT); Flex Jet (Rosetta Inpharmatic, Inc., Kirkland, WA); un espectrómetro de masas MALDI-TOF (Sequenome, San Diego, CA); ChipMaker 2 y ChipMaker 3 (Arrayit Corporation, Sunnyvale, CA); y GenoSensor (Abbot Molecular, Des Plaines, Illinois) según se identificaron y describieron en Heller, *Annu. Rev. Biomed. Eng.* 4:129-153 (2002). Ejemplos de chips de genes o micromatrices también se describen en las Patentes de EE.UU. núms de publicación: 2007-0111322, 2007-0099198, 2007-0084997, 2007-0059769 y 2007-0059765 y en las Patentes de EE.UU. Nº: 7.138.506, 7.070.740 y 6.989.267.

Las técnicas de reacción en cadena de la polimerasa cuantitativa (Q-PCR) tales como la reacción en cadena de la polimerasa en tiempo real (PCR) y la PCR multiplex son métodos bien conocidos de caracterización y cuantificación de ácidos nucleicos. Tales técnicas requieren exposición repetida o prolongada de ácidos nucleicos a la luz, en donde los métodos y composiciones descritos en la presente memoria son útiles para inhibir la degradación inducida por la luz. Cinco de las industrias más populares para realizar la PCR en tiempo real y/o la PCR multiplex incluyen TaqMan® (Life Technologies, Carlsbad, CA), sondas fluorescibles, sondas FRET, Scorpions® (Sigma-Aldrich, Inc., St. Louis, MO ) y SYBR® Green (Life Technologies, Carlsbad, CA). Todas estas industrias permiten la detección de productos de la PCR mediante la generación de una señal fluorescente. Las sondas TaqMan®, sondas fluorescibles, las sondas FRET y Scorpions® dependen de la transferencia de energía de resonancia de Förster (FRET) para generar la señal de fluorescencia mediante el acoplamiento de una molécula de colorante fluorogénico y una fracción de desactivador a los mismos o diferentes sustratos de oligonucleótidos. SYBR Green es un colorante fluorogénico que presenta poca fluorescencia cuando está en solución, pero emite una señal fluorescente fuerte al unirse al ADN de doble cadena.

Las sondas TaqMan® dependen de la actividad 5'-nucleasa de la ADN polimerasa utilizada para la PCR para hidrolizar un oligonucleótido que se hibrida con el amplicón diana. Las sondas TaqMan® son oligonucleótidos etiquetados doblemente que tienen un colorante indicador fluorescente unido al extremo 5' y una fracción de desactivación acoplada al extremo 3'. Típicamente, las sondas TaqMan® consisten en una sonda de oligonucleótidos de 18-22 pb. Estas sondas están diseñadas para hibridarse con una región interna de un producto de la PCR. En el estado no hibridado, la proximidad del colorante indicador fluorescente y las moléculas de desactivación previene la detección de la señal fluorescente de la sonda. Cuando se lleva a cabo un experimento de la PCR en tiempo real, se añade a la mezcla de reacción de la PCR una sonda TaqMan®, complementaria a la secuencia diana. Durante la PCR, la sonda se hibrida específicamente entre el cebador directo e inverso a una región interna del producto de la PCR. Cuando la polimerasa replica un molde sobre el cual se une una sonda TaqMan®, la actividad nucleasa 5' de la polimerasa divide la sonda. Esto desacopla los tintes fluorescentes y desactivadores y FRET ya no ocurre. Así, la fluorescencia aumenta en cada ciclo, proporcional a la cantidad de escisión de sonda.

Al igual que las sondas TaqMan®, las sondas fluorescibles también utilizan FRET para detectar y cuantificar el producto de la PCR sintetizado a través de un fluoróforo acoplado al extremo 5' y un desactivador unido al extremo 3' de un sustrato de oligonucleótido. Una sonda fluorescible consta de 4 partes, específicamente un bucle, un tallo, un fluoróforo 5' y un desactivador 3'. El bucle es típicamente una región de 18-30 pares de bases de la sonda fluorescible que es complementaria a la secuencia diana. La secuencia del tallo se sitúa en ambos extremos del bucle y es típicamente 5-7 pb de largo. Ambas secuencias del tallo son complementarias entre sí. El extremo 5' de la sonda fluorescible contiene un fluoróforo y el extremo 3' de la sonda fluorescible contiene un colorante de desactivación que cuando la sonda está en una forma de bucle cerrado impide que el fluoróforo emita luz. Cuando una sonda fluorescible hibrida con una diana, el colorante fluorescente y el desactivador se separan, FRET no ocurre, y el colorante fluorescente emite luz tras la irradiación. Sin embargo, a diferencia de las sondas TaqMan, las sondas fluorescibles están diseñadas para permanecer intactas durante la reacción de amplificación y deben volver a conectarse a la diana en cada ciclo para la medición de la señal. Las sondas fluorescibles pueden informar de la presencia de ácidos nucleicos específicos de una solución homogénea. Para la PCR cuantitativa, las sondas fluorescibles se unen a la diana amplificada después de cada ciclo de amplificación y la señal resultante es proporcional a la cantidad de molde.

Las sondas FRET son un par de sondas fluorescentes diseñadas para hibridar con regiones adyacentes en el ADN diana, tal como se describe en Didenko, *Biotechniques* 31(5):1106-1121 (2001). Los fluoróforos son elegidos de tal manera que el espectro de emisión de uno se superpone significativamente con el espectro de excitación del otro. Durante la PCR, los dos oligonucleótidos diferentes hibridan con regiones adyacentes del ADN diana de tal manera que los fluoróforos, que están acoplados a los oligonucleótidos, están en estrecha proximidad en la estructura híbrida. El fluoróforo donante es excitado mediante una fuente de luz externa, luego pasa parte de su energía de excitación al fluoróforo aceptor adyacente. El fluoróforo aceptor excitado emite luz a una longitud de onda diferente que puede ser detectada en canales específicos y medida. La fuente de luz no puede excitar el colorante aceptor.

Con las sondas Scorpion®, se logra el cebado específico de la secuencia y la detección del producto de la PCR que utiliza un solo oligonucleótido como se describe en Bates et al., *Molecular Plant Pathology* 2(5):275-280 (2001), Hart et al., *J. Clin. Microbiol.* 39(9):3204-12 (2001), y Thelwell et al., *Nucleic Acids Research* 28(19):3752-61 (2000). Los cebadores de Scorpion® son moléculas bifuncionales en las que un cebador está unido covalentemente a la sonda.

La sonda Scorpion® mantiene una configuración de horquilla en el estado no hibridado. Un fluoróforo se une al extremo 5' y se desactiva por un resto acoplado al extremo 3'. La porción 3' del tallo también contiene una secuencia que es complementaria al producto de extensión del cebador. Esta secuencia está unida al extremo 5' de un cebador específico a través de un monómero no amplificable. En ausencia de la diana, el desactivador casi absorbe la

fluorescencia emitida por el fluoróforo. En los ciclos iniciales de la PCR, el cebador hibrida con la diana y la extensión se produce debido a la acción de la polimerasa. Después de la extensión del cebador Scorpion®, la secuencia de sonda específica es capaz de unirse a su complemento dentro del amplicón amplificado abriendo así la horquilla y separando el fluoróforo y el desactivador, lo que conduce a un aumento de la fluorescencia emitida. La fluorescencia se puede detectar y medir en el tubo de reacción durante cada ciclo sucesivo de amplificación.

SYBR® Green proporciona uno de los formatos más simple y económico para detectar y cuantificar productos de la PCR en reacciones en tiempo real. SYBR® Green es un colorante de cianina asimétrico como se describe por Zipper et al., *Nucleic Acids Res.* 32(12):e103 (2004). SYBR® Green se une preferentemente al ADN bicatenario, pero teñirá el ADN monocatenario con un rendimiento más bajo. El complejo de colorante de ADN resultante absorbe luz azul (máx = 488 nm) y emite luz verde (máx = 522 nm). Ya que el colorante se une preferentemente al ADN bicatenario, no hay necesidad de diseñar una sonda para cualquier diana particular que se esté analizando. Sin embargo, la detección mediante SYBR Green requiere una optimización extensiva. Ya que el colorante no puede distinguir entre productos específicos y no específicos acumulados durante la PCR, a medida que se acumula cualquier producto de la PCR, aumenta la fluorescencia. Las ventajas de SYBR® Green son que es barato, fácil de usar y sensible. La desventaja es que SYBR® Green se unirá a cualquier ADN bicatenario en la reacción, que incluye dímeros de cebadores y otros productos de reacción no específicos, que resulta en una sobrestimación de la concentración diana. Para las reacciones de un solo producto de la PCR con cebadores bien diseñados, SYBR® Green puede funcionar extremadamente bien, con fondo no específico espurio apareciendo solamente en ciclos muy tardíos. Se conocen colorantes de cianina similares en la técnica e incluyen SYBR® Green II, SYBR Gold, YO (Oxazole Yellow), TO (Naranja de Tiazol) y PG (PicoGreen).

Sondas TaqMan®, sondas fluorescibles y Scorpions® permiten medir múltiples especies de ADN en la misma muestra, también conocida como la PCR multiplex, ya que los colorantes fluorescentes con diferentes espectros de emisión se pueden unir a las diferentes sondas. La PCR multiplex permite que los controles internos se co-amplifiquen y permite la discriminación de alelos en ensayos homogéneos de un solo tubo. Estas sondas de hibridación proporcionan un nivel de discriminación imposible de obtener con SYBR Green, ya que solo se hibridarán con dianas verdaderas en una PCR y no con dímeros de cebadores u otros productos espurios. Sin embargo, la PCR multiplex requerirá también etapas de irradiación repetidas y prolongadas adicionales para cuantificar las múltiples emisiones fluorescentes de una sola muestra. Por consiguiente, los métodos y composiciones descritos en la presente memoria son útiles para inhibir la degradación del ácido nucleico inducida por la luz.

Se describen materiales, composiciones y componentes que se pueden utilizar para, se pueden utilizar junto con, se pueden utilizar en la preparación para, o son productos de los métodos y composiciones descritos. Estos y otros materiales se describen en la presente memoria y se entiende que cuando se describen combinaciones, subconjuntos, interacciones, grupos, etc. de estos materiales, mientras que la referencia específica de cada varias combinaciones y permutaciones individuales y colectivas puede no ser explícitamente descrita, cada una se contempla y describe específicamente en la presente memoria descriptiva. Por ejemplo, si se describe y discute un animal quimérico y se discuten una serie de modificaciones que se pueden hacer al animal quimérico, se contemplan específicamente cada combinación y permutación del animal quimérico y las modificaciones que son posibles, a menos que se indique específicamente de lo contrario. Del mismo modo, cualquier subconjunto o combinación de éstos también se contempla y describe específicamente. Este concepto se aplica a todos los aspectos de esta descripción que incluye, pero no se limita a, etapas en métodos de uso de las composiciones o animales descritos. Así, si hay una variedad de etapas adicionales que se pueden realizar, se entiende que cada una de estas etapas adicionales se puede realizar con cualquier etapa de procedimiento específico o combinación de etapas de método de los métodos descritos, y que cada tal combinación o subconjunto de combinaciones se contempla específicamente y se debería considerar descrita.

Se han descrito varias realizaciones. Sin embargo, se comprenderá que se pueden realizar diversas modificaciones. Por consiguiente, otras realizaciones están dentro del alcance de las reivindicaciones siguientes.

### Ejemplos

Ejemplo I. Tasa de error e intensidad de 20 ciclos

Este Ejemplo muestra tasas de error con ciclos de detección repetidos en una secuencia mediante el formato de síntesis que utiliza una célula de flujo de ocho carriles. Los datos de este ejemplo se generaron utilizando el procedimiento estándar de configuración de una tanda de secuenciación en instrumentos HiSeq (Illumina Inc., San Diego, CA) según HiSeq UserGuide y utilizando reactivos estándar Illumina Sequencing, excepto para la solución de detección. Los métodos para los protocolos estándar están disponibles en Illumina, Inc. y se enumeran en Bentley et al., *Nature* 456:53-59 (2008). Brevemente, se utiliza una pluralidad de nucleótidos modificados etiquetados con fluorescencia para secuenciar grupos de ADN amplificado presentes en la superficie de una célula de flujo. Para

5 iniciar el primer ciclo de secuenciación, uno o más nucleótidos etiquetados de forma diferente, y los reactivos apropiados, p. ej., ADN polimerasa, etc., se fluyen en/a través de la célula de flujo mediante un sistema de flujo de fluido. Después de la incorporación, los nucleótidos no incorporados se eliminan fluyendo una solución de lavado a través de la célula de flujo. Una solución de detección se fluye entonces a través de la célula de flujo mientras se utilizan láseres para excitar los ácidos nucleicos e inducir fluorescencia. A continuación, se añade un reactivo de desbloqueo a la célula de flujo para eliminar los grupos terminadores reversibles de las cadenas de ADN que se extendieron y se detectaron. A continuación, el reactivo de desbloqueo se elimina fluyendo una solución de lavado a través de la célula de flujo. La célula de flujo está entonces lista para un ciclo adicional de secuenciación comenzando con la introducción de un nucleótido etiquetado como se ha expuesto anteriormente. Los etapas de flujo y detección se repiten varias veces para completar una secuencia de secuenciación.

10 Para generar la solución de detección utilizada en los experimentos, se añadieron los compuestos de ensayo de polvo a la solución de detección a las soluciones finales indicadas en la Tabla 1 y se comprobó el pH y se ajustó si era necesario. La herramienta "Image Cycle Pump" se utilizó para ensayar 3 compuestos por tanda tal que los carriles 1 y 2 se escanearon en solución de detección control, carriles 3 y 4 en detección de solución que contiene un primer compuesto de ensayo, carriles 5 y 6 en solución de detección que contenía un segundo compuesto de ensayo y carriles 7 y 8 en solución de detección que contiene un tercer compuesto de ensayo. Se compararon el decaimiento de intensidad y las tasas de error entre los carriles que recibieron la solución de detección control y los que recibieron solución de detección que contenía los compuestos de ensayo.

Ejemplo II. Error por exceso de grupos PF

20 Este Ejemplo muestra una respuesta de dosis de ácido gálico que utiliza 20 (carriles 3 y 4) 40 (carriles 5 y 6) y 80 mM (carriles 7 y 8) con carriles 1 y 2 que son los controles sin ácido gálico.

25 Este experimento se llevó a cabo exactamente como se ha descrito anteriormente en el Ejemplo I. La Figura 3 muestra un gráfico que traza el error frente al número de grupos que pasan el filtro (PF). Las gráficas también indican una mejora en grupos que pasan los filtros con un óptimo de 40 mM – a 80 mM el número de grupos que pasan los filtros es menor si se compara con 40 mM.

Ejemplo III. Gráficas de cobertura

Este ejemplo muestra los errores por posición en una secuencia de referencia.

30 Este experimento se llevó a cabo exactamente como se describe en el Ejemplo I. La Figura 4 muestra gráficas de cobertura para una secuencia control, en presencia de ácido gálico, en presencia de urea y en presencia de ácido gálico y urea, durante 75 ciclos. La Figura 5 muestra gráficas de cobertura para una secuencia control, en presencia de ácido gálico, en presencia de urea, y en presencia de ácido gálico y urea, durante 100 ciclos. Las gráficas de cobertura analizan el error (eje Y) por posición en un genoma dado (eje X, que utiliza el genoma phiX como estándar). Los datos muestran que el genoma phiX contiene puntos calientes con una tasa de error significativamente mayor que el resto del genoma. Ácido gálico en la solución de detección reduce la tasa de error de algún punto caliente significativamente, mientras que no afecta a otro punto caliente tanto. La urea por sí misma tiene muy poco efecto reduciendo las tasas de error pero muestra una notable sinergia con ácido gálico que casi elimina el error en la mayoría de los puntos calientes e incluso reduce el error del punto caliente dominante (alrededor de la posición 4300) mientras que cada uno de los compuestos por sí solos no tiene efecto.

Tabla 1

Número de entrada	Compuesto	concentración mM	Tasa de Error	ciclo 20 Int
1	1,4-dihidroxi-2,6-dimetoxybenceno	10	0,81	73,91
2	Ácido 1,4-dihidroxi-2-naftoico	10	5,22	64,32
3	2,2,6,6-tetrametil-4-piperidinol	10	3,11	50,47
4	Ácido 2,3-dihidroxibenzoico	50	0,63	78,86
5	Ácido 3,5-dihidroxibenzoico	10	0,62	75,66

## ES 2 627 851 T3

Número de entrada	Compuesto	concentración mM	Tasa de Error	ciclo 20 Int
6	Ácido 3,4-dihidroxibenzoico	10	0,73	78,60
7	Ácido-2-fenil-5-benzimidazolsulfónico	30	0,97	79,28
8	3,4,5-trihidroxibenzamida	10	2,08	76,71
9	Ácido 3,4-diaminobenzoico	10	2,01	80,47
10	Ácido 3-aminobenzoico	10	0,33	75,59
11	Ácido 3-amino-4-hidroxibenzoico	10	1,66	80,72
12	Ácido 3-hidroxibenzoico	10	2,23	70,79
13	Ácido 3-hidroxi-4-nitrobenzoico	10	3,72	73,50
14	Ácido 3-hidroxiantranílico	10	0,85	81,59
15	Ácido 3-hidroxicinámico	10	1,49	76,84
16	Ácido 4-aminosalicílico	10	1,75	80,08
17	Ácido 4-aminobenzoico	10	0,57	74,85
18	Ácido 4-amino-3-hidroxibenzoico	10	1,61	79,21
19	Ácido 4-amino-3-metoxibenzoico	10	2,36	74,98
20	Ácido 4-hidroxibenzoico	10	2,01	74,76
21	Ácido 4-hidroxi-3-nitrobenzoico	10	3,09	75,91
22	Ácido 4-hidroxicinámico	10	1,92	68,78
23	Ácido 5-benzoil-4-hidroxi-2-metoxibencenosulfónico	10	2,58	75,53
24	1-aza-3,7-dioxabuciclo[3,3,0]octano-5-metanol	30	0,38	78,19
25	Ácido benzoico	100	2,24	71,98
26	Sebacato de bis[2,2,6,6-tetrametil-4piperidona]	20	0,92	76,17

## ES 2 627 851 T3

Número de entrada	Compuesto	concentración mM	Tasa de Error	ciclo 20 Int
27	Ácido cafeico	10	0,41	76,60
28	Catecol	10	1,58	78,54
29	Ácido clorogénico	10	0,86	79,65
30	1,4-diabicciclo[2,2,2]octano	10	1,02	76,32
31	Sulfato de diaminotolueno	10	0,56	76,08
32	Dietilditiocarbamato	25	1,39	76,28
33	N,N-dietilhidroxilamina	30	0,43	78,94
34	dietilditiocarbamato	10	1,46	79,30
35	N,N-dietilhidroxilamina	10	1,41	76,66
36	1,4-dimetilpiperazina	10	1,03	76,17
37	Ácido dietilentriaminopentaacético	10	1,28	68,53
38	Ácido 5-etoxysalicílico	10	0,50	75,31
39	etil galato	10	0,36	85,89
40	Ácido ferúlico	10	0,77	72,58
41	Ácido gálico	100	1,39	87,37
42	Ácido gálico	200	0,82	78,83
43	Ácido gálico	20	0,58	84,30
44	Ácido gálico	50	0,39	84,98
45	Glutación	10	0,75	74,70
46	Histidina	10	0,92	76,18
47	Ácido 5-metoxisalicílico	10	0,76	73,57

## ES 2 627 851 T3

Número de entrada	Compuesto	concentración mM	Tasa de Error	ciclo 20 Int
48	Hesperidina metil chalcona	30	0,26	79,84
49	Hidrazina	10	0,63	75,32
50	Hidroquinona	10	0,92	79,48
51	Lauril galato	10	1,69	65,94
52	Ácido $\alpha$ -lipoico	10	1,01	76,32
53	Hidrocloreuro de $\beta$ -mercaptoetilamina	10	1,75	72,29
54	Melatonina	1	0,43	72,42
55	3,4,5-trihidroxibenzoato de metilo	20	0,97	82,34
56	3,4,5-trihidroxibenzoato de metilo	10	0,47	82,60
57	Metoxihidroquinona	10	0,90	77,17
58	N-acetilcisteína	10	2,97	76,56
59	Ácido nordihidroguaiarético	10	1,25	68,60
60	Ácido 2-fenilbenzimidazol sulfónico	30	0,30	78,37
61	Ácido p-cumárico	10	1,63	77,05
62	Dihidrocloreuro de 1-4-fenilendiamina	10	1,58	79,04
63	Propil galato	30	1,96	73,40
64	Ácido ftálico	10	0,73	73,97
65	$\alpha$ -(4piridil-N-óxido)N-terc-butilnitrona	10	1,72	70,99
66	Propil galato	10	0,38	84,45
67	Piridoxina	100	1,95	72,53
68	Ácido quínico	100	1,79	68,36

## ES 2 627 851 T3

Número de entrada	Compuesto	concentración mM	Tasa de Error	ciclo 20 Int
69	Quercetina	1	0,50	77,43
70	Rutina	1	0,42	74,18
71	Salicilato	10	0,64	76,49
72	Selenometionina	10	3,34	62,84
73	Selenita de sodio	30	1,46	78,66
74	sulfito de sodio	10	0,63	75,84
75	Espermina	100	0,29	54,69
76	Ácido sulfanílico	10	1,57	76,47
77	Ácido siríngico	10	1,12	82,32
78	Tert-butilhidroquinona	10	1,33	84,93
79	4-hidroxi-2,2,6,6-tetrametilpiperidina 1-oxilo	10	1,97	71,31
80	Ácido tereftálico	10	0,86	74,54
81	Tetracloro-1,4-benzoquinona	10	4,03	54,01
82	2,4,5-trihidroxi-3,6-dimetil-1,4-benzoquinona	10	1,02	84,56
83	1,3,5-tris(2-hidroxietil)isocianurato	10	1,75	76,86
84	Tiourea	100	15,31	51,62
85	Hidrocloreto de 2,2,6,6-tetrametil-4-piperidona	10	1,89	69,38
86	Ácido 6-hidroxi-2,5,7,8-tetrametilcroman-2-carboxílico	10	0,84	74,97
87	L-triptófano	10	1,69	76,44

Tabla 2

Número de entrada	Compuesto	Tasa de Error	ciclo 20 Int
1	Etil galato 10 mM + 10 mM Sebacato de bis[2,2,6,6-tetrametil-4piperidona]	0,57	88,89
2	Etil galato 10 mM + 10 mM Dietiltiocarbamato	0,50	90,48
3	Etil galato 10mM + 10mM N,N-Dietilhidroxilamina	0,49	69,01
4	Etil galato 10 mM + 10mM Ácido 2-fenilbenzimidazol sulfónico	0,35	70,58
5	Etil galato 10 mM +10mM pirogalol	0,31	72,78
6	Etil galato 10 mM + 25mM Ascorbato de sodio	0,34	85,64
7	Etil galato 10 mM + 30mM Hidroquinona	0,45	90,58
8	Etil galato 10 mM + 50mM Ácido gálico	0,34	87,25
9	Ácido gálico 100mM + 250mM Manitol	0,31	77,08
10	Ácido gálico 100mM + 2M Urea	0,17	80,90

Ejemplo IV. El daño de ácido nucleico se puede prevenir mediante inclusión de un compuesto durante la fragmentación de ácido nucleico.

5 La fidelidad con la que una biblioteca secuenciada coincide con la secuencia del genoma original puede variar dependiendo de la frecuencia de mutación de base que ocurre en cualquier etapa desde la extracción del ácido nucleico hasta su secuenciación en una plataforma de secuenciación. Esta frecuencia coloca un límite superior en la probabilidad de que una base secuenciada sea correcta. Por ejemplo, si una mutación ocurre una en mil veces, entonces la confianza máxima (probabilidad) de que cualquier base es correcta es uno en un  $10^3$ , es decir, un máximo de Q30. Se ha observado que el ADN fragmentado mediante cizallamiento de Covaris tiene un valor

10 máximo de Q que disminuye con el aumento de la duración del cizallamiento. Véase la Figura 6. Para determinar si se puede utilizar un compuesto para aumentar la fidelidad con la que una biblioteca de secuencia coincide con la secuencia del genoma original, se generaron datos que utilizan pools de ADN plasmídico. Se agruparon 12 plásmidos, teniendo cada plásmido un inserto de 50 pares de bases (pb) diferente dentro del casete del molde. El pool se cizalló en un Covaris S2, entonces, se sometió a agrupación (clustering) mediante medios estándar seguido de lecturas de 75 pares de pares en un secuenciador GA (Illumina, Inc., San Diego, CA). Se calcularon entonces puntuaciones de calidad máxima a partir de los datos. Específicamente, 1  $\mu$ g de ADN de pool de plásmido se cizalló utilizando Covaris en ausencia o presencia de compuestos como sigue: a) 1  $\mu$ g de ADN de entrada se fragmentó en presencia de ácido gálico 50 mM; o b) 1  $\mu$ g de ADN de entrada se fragmentó en presencia de ascorbato 50 mM. Las soluciones se hacen mediante la adición de 60  $\mu$ l de una solución madre de 100 mM de ácido gálico o ascorbato a 10  $\mu$ l de ADN (1  $\mu$ g) y el volumen se completó hasta 120  $\mu$ l con tampón TE (Tris, EDTA). Los parámetros de covaris utilizados fueron los siguientes: ciclo de trabajo: 20%, intensidad: 5, ciclos/ráfaga: 200 y tiempo: 240 segundos. Estas muestras se pusieron entonces en la célula de flujo para secuenciación sin tratamiento adicional. Como se muestra en la Tabla 1 a continuación y en las Figuras 7 y 8, el daño de ácido nucleico se puede prevenir mediante

15 inclusión de ácido gálico o ascorbato durante la reacción de cizallamiento.

25

Tabla 3. Efecto del ácido gálico y del ascorbato durante la fragmentación del ácido nucleico.

	Muestra Q Máx	Detección Q Máx	Sistema Q Máx
Ácido Gálico	43	63	43
Ascorbato	43	65	44
Sin tratamiento	38	67	38
Control HindIII ADN Digerido	44	60	43

Ejemplo V. Los ácidos nucleicos dañados pueden eliminarse mediante tratamiento enzimático.

5 Para determinar si los ácidos nucleicos dañados se pueden escindir mediante escisión enzimática, 3 µg del pool de ADN de precisión se cizallaron utilizando Covaris. Los parámetros de Covaris utilizados fueron los siguientes: ciclo de trabajo: 20%, intensidad: 5, ciclos/ráfaga: 200 y tiempo: 240 segundos. Después de esto, la muestra se dividió en tubos de 1 µg cada uno. Se trató un tubo (1 µg) con enzima FPG como sigue: Se añadieron 10 µl de enzima FPG (80 unidades) a 1 µg de ADN en tampón 1X NEB 1, suplementado con albúmina de suero bovino (BSA) hasta un volumen final de 200 µl. Alternativamente, se trató un tubo con PreCR™ Repair Mix (NEB, Ipswich, MA). Las reacciones se incubaron durante 30 minutos a 37°C. Las reacciones se limpiaron a continuación utilizando el kit de purificación de PCR Qiagen y se eluyeron en 30 µl de tampón EB (Tris-Cl 10 mM, pH 8,5). Como se muestra en la Tabla 2 a continuación y en las Figuras 2 y 3, el tratamiento de ADN cizallado de Covaris con FPG (una enzima que escinde ADN en las lesiones 8-oxo-G) restaura el valor de la muestra Q Max de 38 a 43 y, así, elimina los ácidos nucleicos dañados de la muestra.

Tabla 4. Efecto de FPG después de la fragmentación de ácido nucleico.

	Muestra Q Máx	Detección Q Máx	Sistema Q Máx
Tratamiento FPG	43	65	43
Tratamiento PreCR	38	64	37
Sin tratamiento	38	67	38
Control HindIII ADN Digerido	44	60	43

15

**REIVINDICACIONES**

1. Un método para inhibir la degradación de ácidos nucleicos durante una etapa de procesamiento de detección de ácidos nucleicos, que comprende
- 5 introducir mediante un sistema de flujo fluido uno o más nucleótidos etiquetados con fluorescencia de manera diferente y una polimerasa en una célula de flujo, comprendiendo dicha célula de flujo una matriz de ácidos nucleicos unida a un soporte;
- reemplazar dicho uno o más nucleótidos etiquetados con fluorescencia de manera diferente y una polimerasa con una solución de detección que comprende ácido gálico, un éster alquílico inferior del mismo, o mezclas de los mismos, e
- 10 irradiar una porción de dichos ácidos nucleicos en presencia de dicha solución de detección para inducir fluorescencia, en donde dicha solución de detección reduce la cantidad de degradación inducida por la luz de los ácidos nucleicos.
2. El método de la reivindicación 1, en donde el ácido gálico, un éster alquílico inferior del mismo, o mezclas de los mismos está presente en un intervalo de concentración que varía de aproximadamente 10 mM a aproximadamente
- 15 200 mM.
3. El método de una cualquiera de las reivindicaciones 1 ó 2, en donde la solución de detección comprende además:
- a) un compuesto seleccionado del grupo que consiste en urea, ácido ascórbico o sal del mismo, y ácido isoascórbico o sal del mismo.
4. El método de la reivindicación 1, en donde:
- 20 a) la irradiación se lleva a cabo en un intervalo de aproximadamente 360 nm a aproximadamente 800 nm,
- b) la irradiación se lleva a cabo con una fuente de luz que tiene potencia en un intervalo de aproximadamente 5 a aproximadamente 500 milivatios; o
- c) la irradiación se lleva a cabo durante un periodo de tiempo de aproximadamente 0,1 segundos a aproximadamente 10 minutos.
- 25 5. El método de la reivindicación 1, en donde la degradación inducida por la luz opcionalmente comprende la eliminación de un miembro de ácido nucleico de la matriz de ácidos nucleicos.
6. El método de la reivindicación 1, que comprende adicionalmente añadir un nucleótido etiquetado con fluorescencia a la matriz y repetir la etapa de procesamiento de detección.
- 30 7. El método de la reivindicación 6, que comprende además repetir las etapas de adición de nucleótidos, reemplazar los dichos uno o más nucleótidos etiquetados con fluorescencia de manera diferente y polimerasa con una solución de detección que comprende ácido gálico, un éster alquílico inferior del mismo, o mezclas de los mismos, y etapas de detección durante al menos 20 o 25 ciclos.
8. El método de la reivindicación 6:
- a) en donde la adición del nucleótido etiquetado con fluorescencia comprende utilizar una polimerasa para añadir un
- 35 solo nucleótido etiquetado con fluorescencia; o
- b) comprende además repetir las etapas de procesamiento de detección y adición durante al menos 50, 75 ó 100 ciclos.
9. El método de la reivindicación 1, en donde el método comprende además secuenciar los ácidos nucleicos.
- 40 10. El método de la reivindicación 1, que comprende además tratar los ácidos nucleicos con una enzima que escinde secuencias de ácido nucleico que comprenden ácidos nucleicos dañados para eliminar cualquier ácido nucleico dañado.
11. El método de la reivindicación 1, en donde la enzima es formamidopirimidina ADN-glicosilasa (FPG).
12. El método de la reivindicación 1, que comprende además añadir un anticuerpo que se une selectivamente a uno o más tipos de ácidos nucleicos dañados para eliminar cualquier ácido nucleico dañado.
- 45 13. El método de la reivindicación 12, en donde el anticuerpo se une a 8-oxo-G.

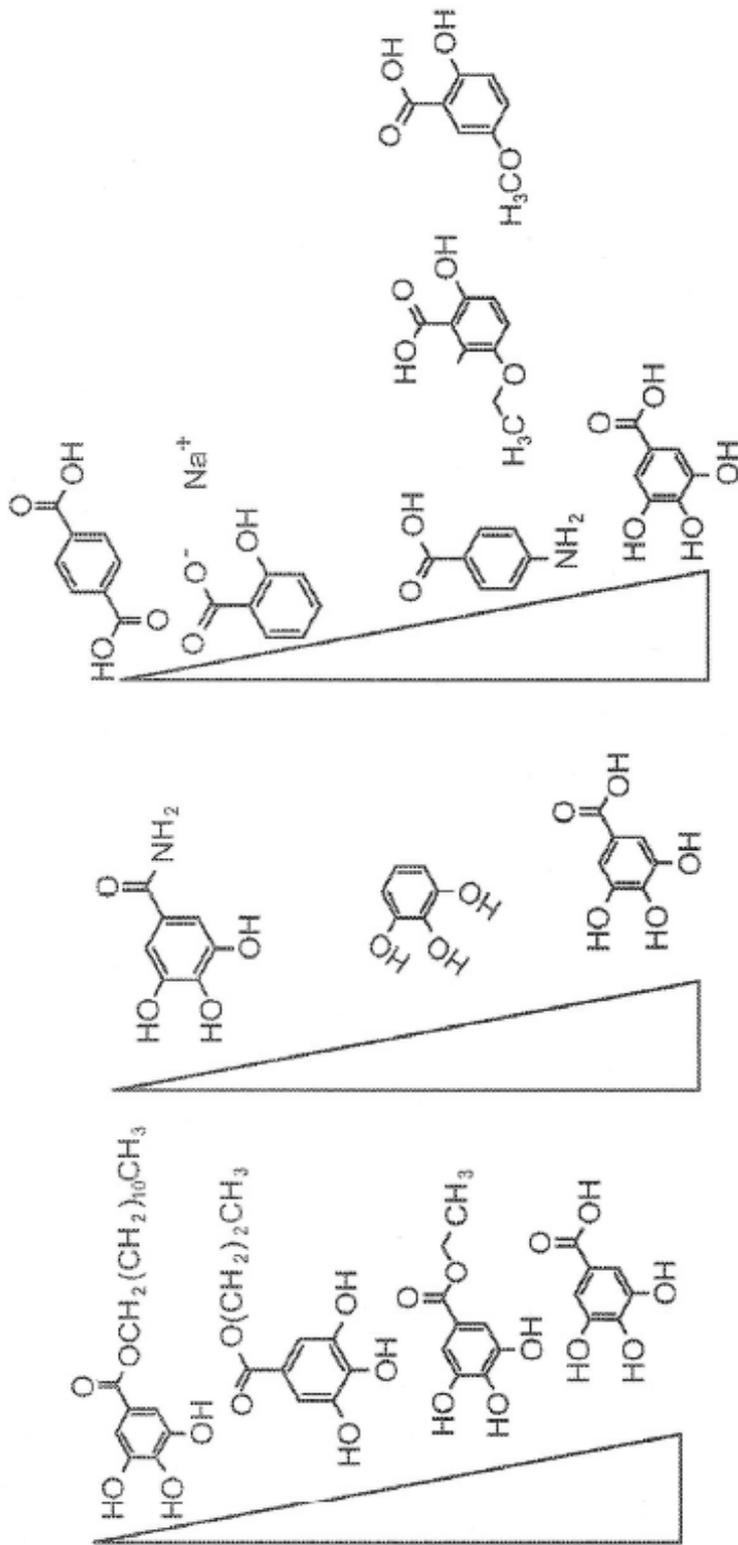


Fig. 1

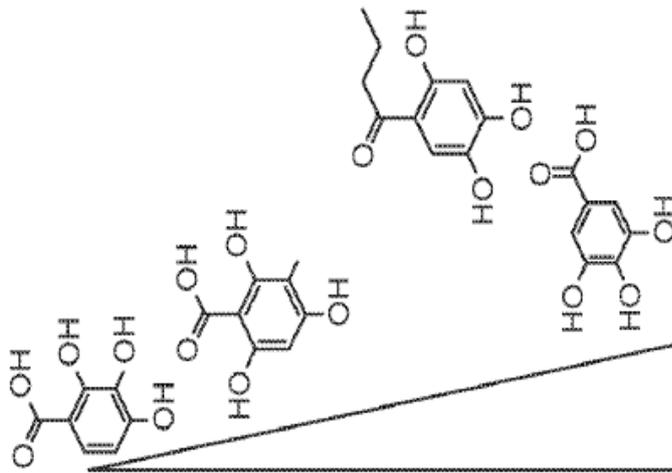
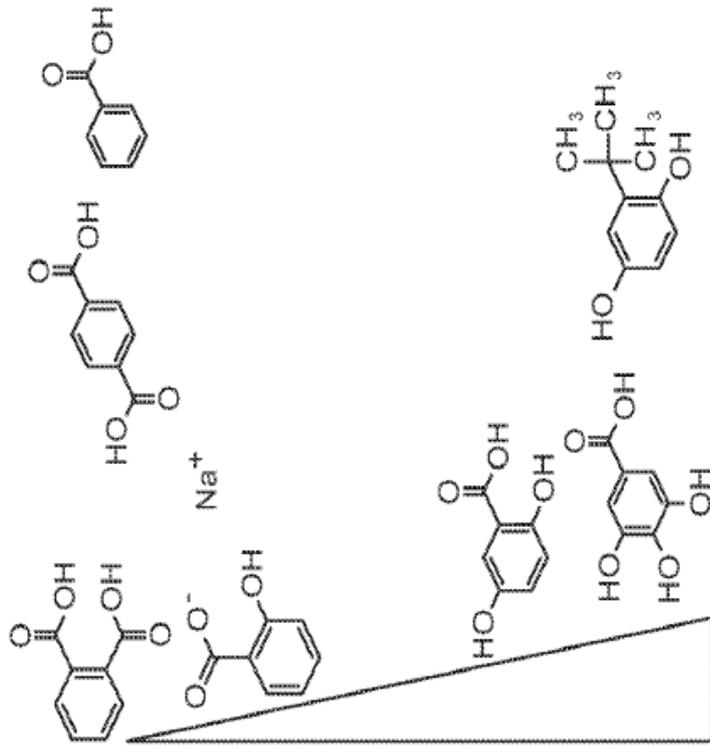


Fig. 2

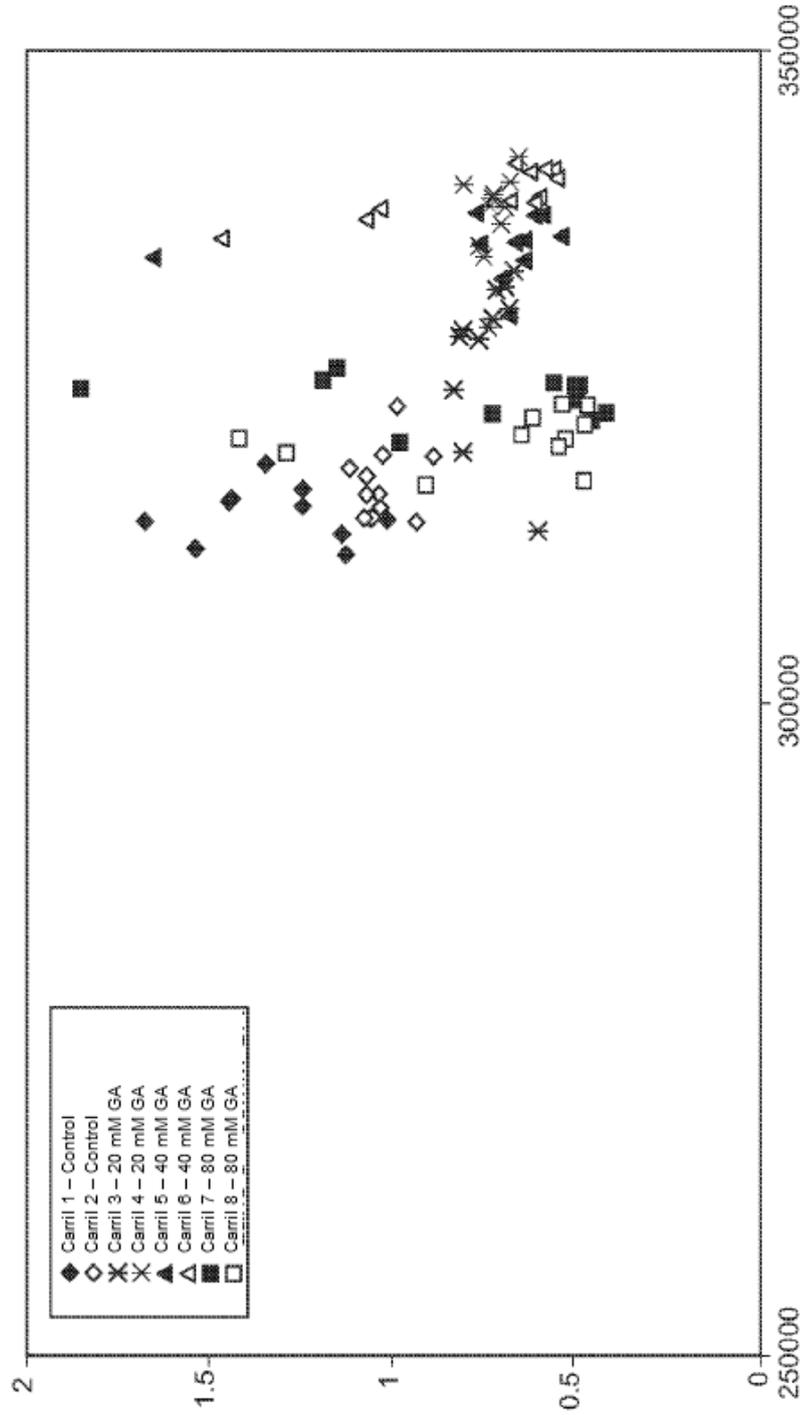


Fig. 3

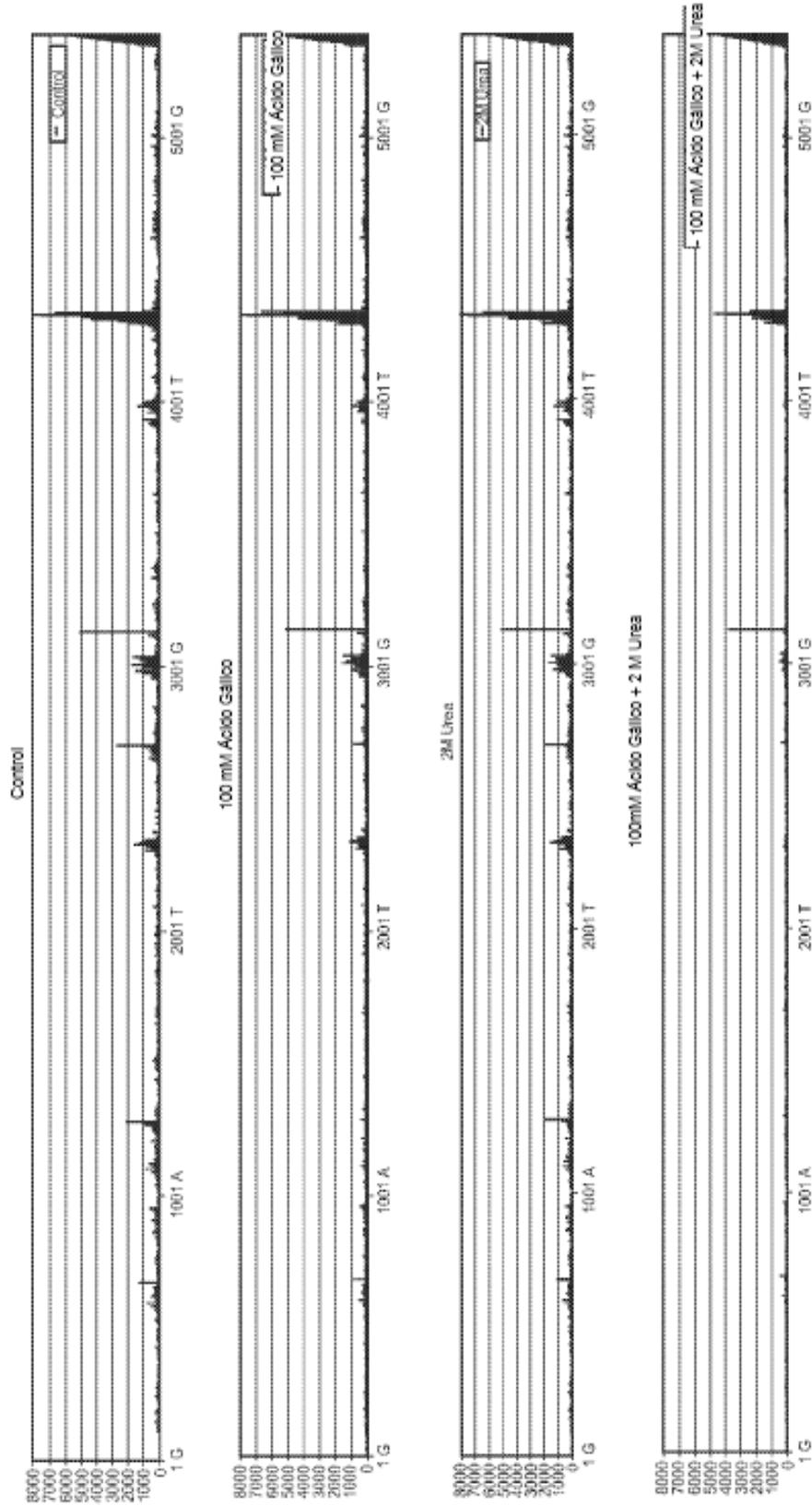


Fig. 4

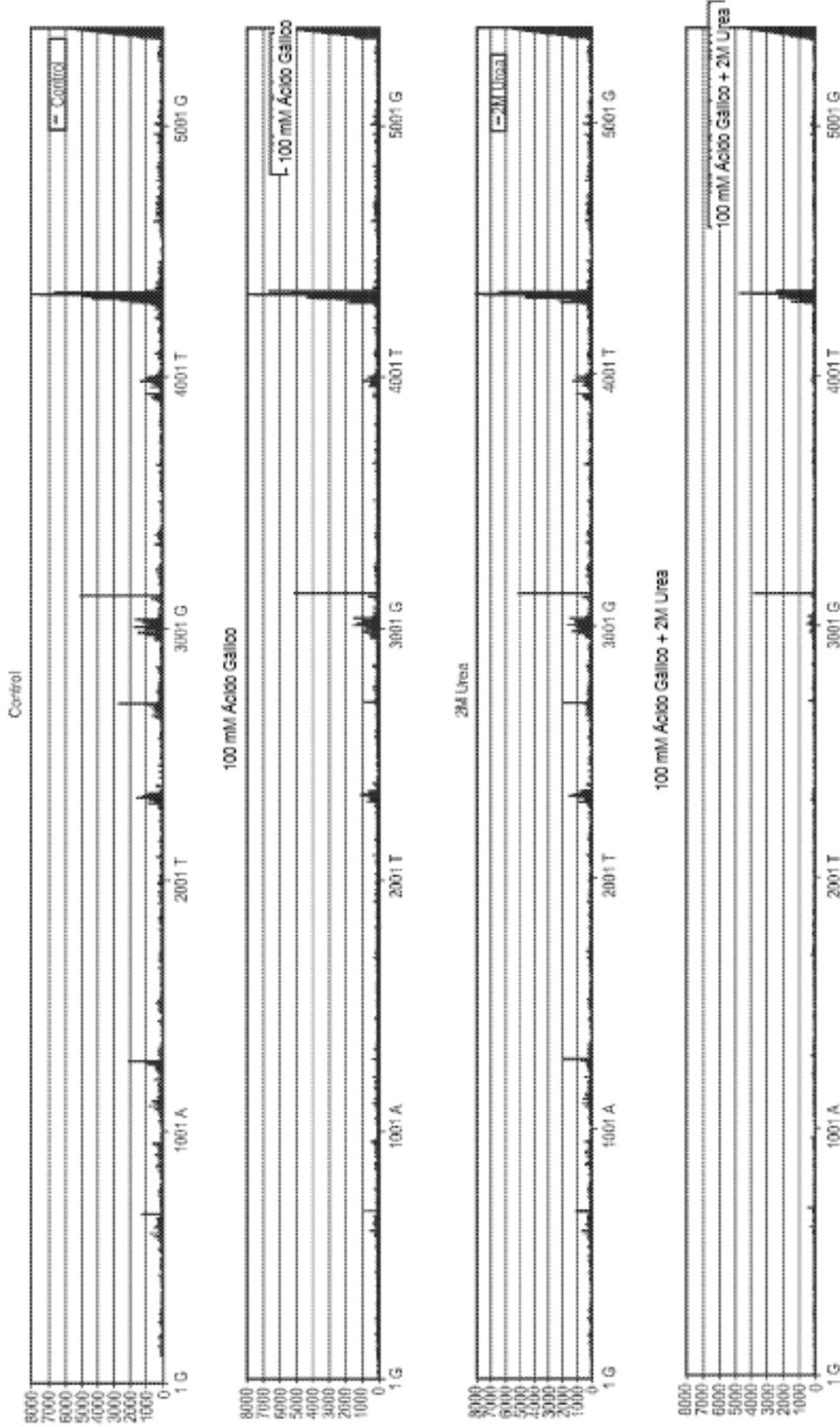


Fig. 5

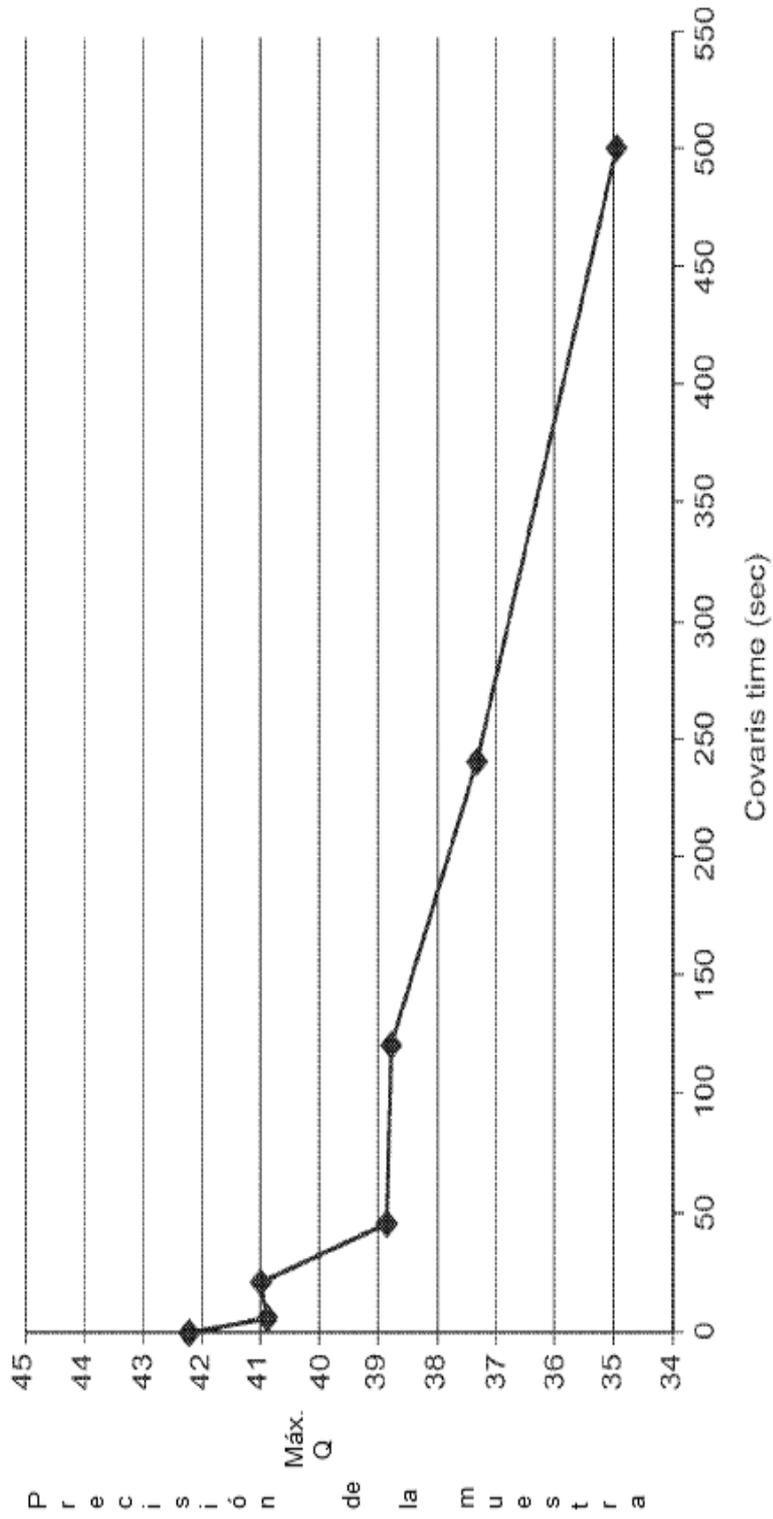


Fig. 6

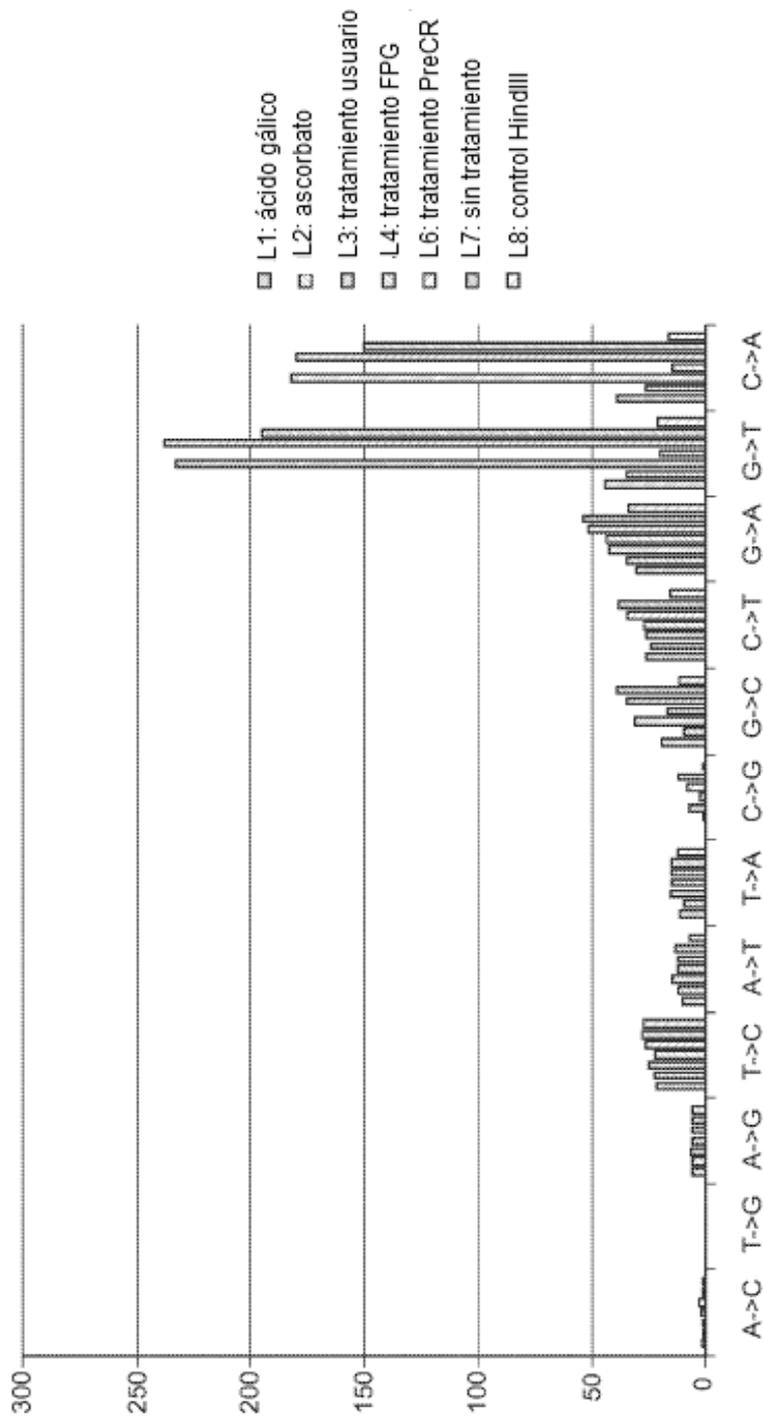


Fig. 7

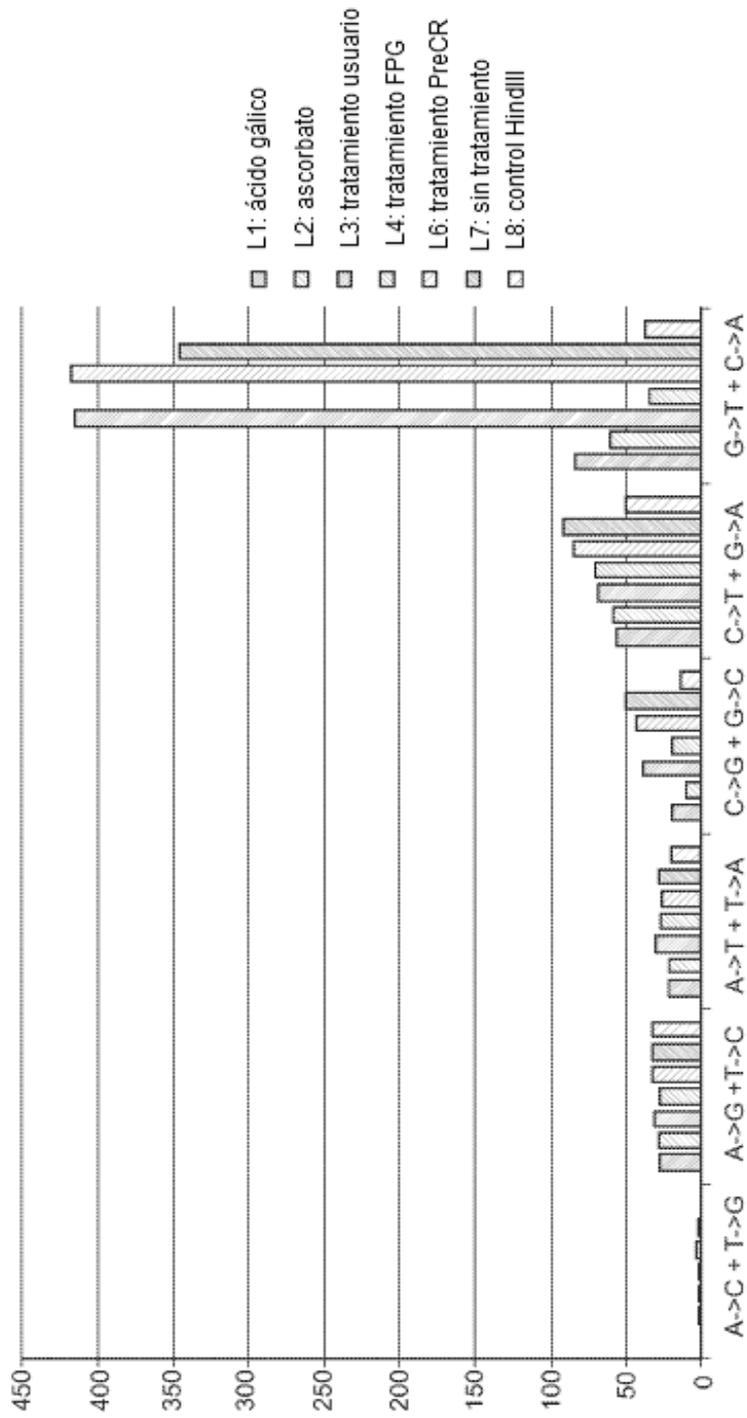


Fig. 8