

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 628 485**

51 Int. Cl.:

**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **25.06.2014 PCT/US2014/044158**

87 Fecha y número de publicación internacional: **08.01.2015 WO15002789**

96 Fecha de presentación y número de la solicitud europea: **25.06.2014 E 14739017 (3)**

97 Fecha y número de publicación de la concesión europea: **05.04.2017 EP 3017063**

54 Título: **Secuenciación mediante síntesis ortogonal**

30 Prioridad:

**03.07.2013 US 201361842501 P**

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

**03.08.2017**

73 Titular/es:

**ILLUMINA, INC. (100.0%)  
5200 Illumina Way  
San Diego, CA 92122, US**

72 Inventor/es:

**FABANI, MARTIN MARIA;  
ROBERT BACIGALUPO, MARIA CANDELARIA y  
MOON, JOHN A,**

74 Agente/Representante:

**ELZABURU, S.L.P**

ES 2 628 485 T3

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

## DESCRIPCIÓN

Secuenciación mediante síntesis ortogonal

### Antecedentes de la invención

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

10 Las plataformas comerciales actualmente disponibles para secuenciar el ADN son relativamente costosas. Estas plataformas utilizan un enfoque de "secuenciación por síntesis", llamado así porque los polímeros de ADN se sintetizan mientras se detecta la adición de cada monómero (es decir, nucleótido) a la estructura polimérica en crecimiento. Debido a que una cadena de ADN de molde dirige estrictamente la síntesis de un nuevo polímero de ADN, se puede inferir la secuencia del ADN molde de la serie de monómeros de nucleótidos que se añadieron a la hebra de crecimiento durante la síntesis. La capacidad para detectar adiciones de monómeros se facilita mediante variantes especialmente diseñadas de los componentes bioquímicos que normalmente llevan a cabo la síntesis de ADN en sistemas biológicos. Estos componentes de ingeniería son relativamente caros de fabricar y se consumen en cantidades relativamente grandes durante la secuenciación por síntesis. Además, el monitoreo de la reacción utiliza un equipo relativamente caro, tal como láser, óptica de detección y sistemas complejos de administración de fluidos. Las plataformas comerciales más exitosas hasta la fecha también requieren costosos reactivos y hardware para amplificar las plantillas de ADN antes de que la secuenciación por síntesis pueda comenzar. La complejidad y el costo de estas plataformas han dificultado su uso en algunos contextos clínicos y de investigación donde hay una clara necesidad de la tecnología.

20 Por lo tanto, existe la necesidad de mejorar las plataformas de secuenciación por síntesis para hacerlas más rentables, rápidas y convenientes. La presente descripción trata esta necesidad y proporciona también otras ventajas.

### Breve resumen

25 La presente descripción proporciona un método para secuenciar plantillas de ácido nucleico. El método puede incluir las etapas de (a) proporcionar un conjunto de sitios, en los que cada sitio incluye un primer molde de ácido nucleico y un segundo molde de ácido nucleico, en el que el primer molde de ácido nucleico tiene una secuencia que es diferente de la secuencia del segundo molde de ácido nucleico; (b) extender un primer cebador unido al primer molde usando una primera especie de polimerasa y un primer conjunto de análogos de nucleótidos, produciendo de este modo un primer producto de extensión de cebador que tiene un primer análogo de nucleótido en cada uno de los sitios; (c) extender un segundo cebador unido a la segunda plantilla usando una segunda especie de polimerasa y un segundo conjunto de análogos de nucleótidos, produciendo de este modo un segundo producto de extensión de cebador que tiene un segundo análogo de nucleótido en cada uno de los sitios, en donde la primera especie de polimerasa es diferente a partir de la segunda especie de polimerasa y en donde el primer conjunto de análogos de nucleótidos es diferente del segundo conjunto de análogos de nucleótidos; (d) detectar el primer producto de extensión del cebador y el segundo producto de extensión del cebador en cada uno de los sitios; y (e) repetir las etapas (b) a (d), determinando así las diferentes secuencias del primer molde y el segundo molde en cada uno de los sitios.

40 También se proporciona en la presente memoria un método para secuenciar moldes de ácido nucleico que incluye las etapas de (a) proporcionar un primer molde de ácido nucleico y un segundo molde de ácido nucleico, en el que el primer molde de ácido nucleico tiene una secuencia que es diferente de la secuencia del segundo molde de ácido nucleico; (b) extender un primer cebador unido al primer molde usando una primera especie de polimerasa y un primer conjunto de análogos de nucleótidos, produciendo de este modo un primer producto de extensión de cebador que tiene un primer análogo de nucleótido; (c) extender un segundo cebador unido al segundo molde usando una segunda especie de polimerasa y un segundo conjunto de análogos de nucleótidos, produciendo de este modo un segundo producto de extensión del cebador que tiene un segundo análogo de nucleótido, en el que la primera especie de polimerasa es diferente de la segunda especie de polimerasa y en el que el primer conjunto de análogos de nucleótidos es diferente del segundo conjunto de análogos de nucleótidos; (d) detectar el primer producto de extensión de cebador y el segundo producto de extensión de cebador usando un detector que tiene una resolución que es menor que la separación espacial entre el primer producto de extensión del cebador y el segundo producto de extensión del cebador; y (e) repetir las etapas (b) a (d), determinando con ello las diferentes secuencias del primer molde y el segundo molde.

### Breve descripción de los dibujos

La figura 1 muestra una representación esquemática de una reacción de síntesis por secuencia ortogonal que incluye una etapa de detección (Fig. 1A) y una etapa de extensión de polimerasa (Figura 1B).

55 La figura 2A muestra conjuntos de ejemplos de desoxirribonucleótidos terminadores reversibles (d-FFN) y ribonucleótidos terminadores reversibles (r-FFN) útiles para una reacción de síntesis por secuenciación ortogonal de 1-tinte.

La figura 2B muestra una representación esquemática de un ciclo de reacción utilizado para una reacción de síntesis por ortogonal de 1 tinte que incluye datos simulados para señales detectadas en 2 canales de emisión diferentes.

La figura 3A muestra conjuntos de ejemplos de desoxirribonucleótidos terminadores reversibles (d-FFN) y ribonucleótidos terminadores reversibles (r-FFN) útiles para una reacción de síntesis por ortogonal de 2 tintes.

- 5 La figura 3B muestra una representación esquemática de un ciclo de reacción usado para una reacción de síntesis por secuencia ortogonal de 2 tintes que incluye datos simulados para señales detectadas en 4 canales de emisión diferentes.

La figura 4 muestra una representación esquemática de un método para preparar moldes para una reacción de síntesis por secuenciación ortogonal.

- 10 La figura 5 muestra una construcción para la síntesis por secuenciación ortogonal en un formato de extremo emparejado.

#### Descripción detallada

- 15 Esta descripción proporciona un método para la detección de ácidos nucleicos de alta densidad. Realizaciones particulares de los métodos de la presente descripción explotan técnicas conocidas para manipular y detectar ácidos nucleicos. Sin embargo, las mejoras expuestas a continuación proporcionan un procesamiento ortogonal de tal manera que se incrementa la densidad de información obtenida por el uso de estas técnicas.

- 20 El ejemplo de la técnica de detección basada en la extensión del cebador es ilustrativo del aumento de la densidad de información que se puede obtener. Específicamente, una secuencia diana de un ácido nucleico se puede hibridar con un cebador y el cebador se extiende mediante una ADN polimerasa para añadir un nucleótido marcado. Se puede usar un formato de matriz con múltiples sitios, teniendo cada sitio una única secuencia diana que difiere de las secuencias diana presentes en otros sitios. Opcionalmente se utilizan también varias especies de nucleótidos diferentes, cada una de las cuales con un marcador distinguible. La extensión del cebador da como resultado el reclutamiento del nucleótido marcado en el ácido nucleico que tiene la secuencia diana. En un formato de matriz, donde se utilizan diferentes nucleótidos marcados, se puede distinguir el marcador que se recluta en cada sitio, y  
25 usar esta información para identificar el ácido nucleico diana en ese sitio. La densidad de la información obtenida a partir de este formato de matriz es una secuencia diana identificada por sitio.

- 30 En un formato ortogonal de la presente descripción, cada sitio de la matriz puede contener dos o más secuencias diana diferentes que son simultáneamente detectables y distinguibles entre sí. En este caso, la información derivada de la matriz puede ser al menos duplicada. Por ejemplo, dos cebadores diferentes se pueden hibridar con las dos secuencias diana diferentes en cada sitio individual. El primer cebador puede ser un cebador de ADN que sea capaz de extenderse mediante la adición de un desoxirribonucleótido marcado usando una ADN polimerasa y el segundo cebador puede ser un cebador de ARN que sea capaz de extenderse por la adición de un ribonucleótido marcado utilizando una ADN polimerasa. Los reactivos utilizados para extender el cebador de ADN son ortogonales a los reactivos utilizados para extender el cebador de ARN en el sentido de que los reactivos de extensión del cebador de ADN no reaccionan inespecíficamente con el cebador de ARN y los reactivos de extensión del cebador de ARN no reaccionan inespecíficamente con el cebador de ADN. Los desoxirribonucleótidos pueden tener marcas que sean distinguibles de los marcadores en los ribonucleótidos. La ortogonalidad resultante en la reactividad bioquímica y el manejo del marcador permite que el evento de extensión del cebador de ADN se distinga del evento de extensión del cebador de ARN en cada sitio. De este modo, las dos secuencias diana pueden detectarse de manera  
40 distinguible.

- La ortogonalidad puede resultar de cualquiera de una variedad de componentes bioquímicos o condiciones de reacción que confieran selectividad a dos eventos de detección diferentes. Como se ejemplifica por las reacciones que usan la ADN polimerasa y la ARN polimerasa, la ortogonalidad puede derivarse de la especificidad de dos polimerasas diferentes para diferentes especies de cebadores y para diferentes clases de nucleótidos. En algunas realizaciones, la ortogonalidad puede derivarse en lugar de la selectividad de diferentes polimerasas para una especie particular de molde (p.ej. ADN vs. ARN) independientemente de si las polimerasas son o no selectivas para una especie particular de cebador o clase de nucleótidos. Por lo tanto, puede ser posible utilizar diferentes polimerasas que extienden los cebadores de ADN con desoxirribonucleótidos, pero que son selectivas diferencialmente para un molde de ADN y un molde de ARN, respectivamente. Por ejemplo, las ADN polimerasas son generalmente selectivas para moldes de ADN y las transcriptasas inversas son generalmente selectivas para moldes de ARN; sin embargo, ambas enzimas pueden usar un cebador de ADN y desoxirribonucleótidos. Se contemplan combinaciones de polimerasas nativas y/o de ingeniería para su uso en sistemas de reacción ortogonales.

- 55 Los conceptos de ortogonalidad ejemplificados anteriormente para una técnica de detección basada en la extensión del cebador pueden aplicarse fácilmente a una técnica de secuenciación por síntesis (SBS). Como se muestra en la Fig. 1A y 1B, y como se expone con más detalle a continuación, cada ciclo SBS se puede llevar a cabo utilizando cebadores ortogonales, polimerasas y nucleótidos para proporcionar una mayor adquisición de la información a partir de una célula de flujo u otro sustrato usado en la técnica SBS. A efectos de ilustración, la ortogonalidad se

ejemplifica para un enfoque de secuenciación, denominado "secuenciación por síntesis ortogonal" (SBOS); sin embargo, otros métodos también pueden beneficiarse de la manipulación y detección ortogonales como se expone con más detalle a continuación. Sin embargo, las composiciones, aparatos y métodos expuestos en este documento no necesitan estar limitados a aplicaciones de secuenciación.

5 La ortogonalidad puede explotarse para incrementar la densidad de adquisición de la información por 2 o más veces. Por ejemplo, se puede obtener un aumento mayor de 2 veces en la densidad de información usando más de dos conjuntos de reactivos ortogonales. Como ejemplo, se pueden utilizar 3 conjuntos de reactivos, incluyendo (1) reactivos de extensión basados en ADN polimerasa, (2) reactivos de extensión basados en ARN polimerasa y (3) una polimerasa modificada acoplada con ANH (ácido nucleicos de 1,5 anhidrohexitol).

10 Como se ha demostrado anteriormente y como se expondrá con más detalle a continuación, la presente descripción proporciona la ventaja de una formación de imágenes de super-resolución de una matriz, por lo que el número de secuencias diana que se resuelven simultáneamente en un sitio dado es mayor que uno. La formación de imágenes de super-resolución puede proporcionar el beneficio de distinguir simultáneamente un número de ácidos nucleicos diana diferentes que sea mayor que el número de sitios en la matriz. De forma similar, se proporciona una super-resolución porque se pueden distinguir dos secuencias diana diferentes en un sustrato de fase sólida usando un detector que tiene una resolución que es inferior a la resolución espacial que de otro modo sería necesaria para distinguir las dos secuencias diana sobre el sustrato.

En realizaciones particulares, esta descripción proporciona reactivos y configuraciones de hardware para la detección eficiente de ácidos nucleicos. Una configuración ejemplar usa menos marcadores que el número de especies de nucleótidos que debe distinguirse en una etapa de extensión de cebador. Por ejemplo, se pueden distinguir cuatro especies de desoxirribonucleótidos basándose en la detección de una única especie de marcador. Como se expone con más detalle a continuación, esto puede conseguirse utilizando un primer conjunto de nucleótidos que incluye las cuatro especies siguientes: (1) una especie que tiene un primer marcador, (2) una especie que tiene un ligando, (3) una especie que tiene un enlace escindible con el primer marcador, y (4) una especie que carece de cualquier marcador o ligando usado en una etapa subsiguiente. Un conjunto ortogonal de nucleótidos (por ejemplo, ribonucleótidos que son ortogonales a desoxirribonucleótidos) puede incluir las siguientes cuatro especies (5) una especie que tiene un segundo marcador, (6) una especie que tiene una mezcla del primer y segundo marcadores, (7) una especie que tiene un enlace escindible con el segundo marcador, y (8) una especie que carece de cualquier marcador o ligando usado en una etapa subsiguiente. Las especies dentro de cada conjunto pueden distinguirse entre sí basándose en una contabilidad adecuada de qué marcadores aparecen o desaparecen después de las etapas fluidicas específicas y los dos conjuntos ortogonales de nucleótidos pueden distinguirse basándose en los dos marcadores diferentes. Volviendo al ejemplo de las 8 especies anteriores, las especies (1) y (5) se pueden distinguir entre sí basándose en los diferentes marcadores y de todas las otras especies debido a su aparición después de una etapa inicial de marcaje y su resistencia al agente de escisión; la especie (2) se puede distinguir basándose en la apariencia del marcador después de la incubación con un receptor marcado; las especies (3) y (7) se pueden distinguir entre sí basándose en los diferentes marcadores y se distinguen de todas las otras especies basándose en la apariencia inicial del marcador y después en la desaparición después del tratamiento con un reactivo de escisión; la especie (6) se puede distinguir de todas las demás especies basándose en la presencia de ambos marcadores con una intensidad que es la mitad de la intensidad para especies totalmente marcadas; y las especies (4) y (8) se pueden distinguir basándose en la inferencia de la ausencia de cualquier otra especie en los conjuntos respectivos que se hayan detectado. Muchas otras configuraciones son posibles para alterar el número de marcadores, el número de manipulaciones de fluidos durante una fase de detección y/o la complejidad del dispositivo de detección para distinguir un cierto número de marcadores. Como tal, la configuración puede adaptarse para adaptarse a un enfoque o aplicación particular.

45 Los términos utilizados en el presente documento se entenderán que adoptan su significado ordinario a menos que se especifique lo contrario. Ejemplos de varios términos usados en este documento y sus definiciones se exponen a continuación.

Tal como se utiliza en este documento, el término "matriz" se refiere a una población de sitios que pueden ser diferenciados entre sí según su ubicación relativa. Diferentes moléculas que están en diferentes sitios de una matriz pueden diferenciarse entre sí de acuerdo a las ubicaciones de los sitios en la matriz. Un sitio individual de una matriz puede incluir una o más moléculas de un tipo particular. Por ejemplo, un sitio puede incluir una única molécula de ácido nucleico diana que tiene una secuencia particular o un sitio puede incluir varias moléculas de ácido nucleico que tienen la misma secuencia (y/o su secuencia complementaria). Los sitios de una matriz pueden ser características diferentes situadas en el mismo sustrato. Las características ejemplares incluyen, sin limitación, pocillos en un sustrato, perlas (u otras partículas) en o sobre un sustrato, proyecciones desde un sustrato, crestas sobre un sustrato o canales en un sustrato. Los sitios de una matriz pueden ser sustratos separados, cada uno de los cuales lleva una molécula diferente. Diferentes moléculas unidas a sustratos separados pueden identificarse de acuerdo con las ubicaciones de los sustratos sobre una superficie a la que están asociados los sustratos o según las ubicaciones de los sustratos en un líquido o gel. Las matrices ejemplares en las que están situados los sustratos separados sobre una superficie incluyen, sin limitación, las que tienen perlas en los pocillos.

Tal como se utiliza en este documento, el término "grupo", cuando se usa en referencia a ácidos nucleicos, se refiere a una población de ácidos nucleicos que está unida a una fase sólida para formar una característica o sitio. Los ácidos nucleicos son generalmente de una sola especie, formando así un grupo homogéneo. Sin embargo, en algunas realizaciones los ácidos nucleicos pueden ser heterogéneos, de manera que las moléculas individuales que tienen secuencias diferentes están presentes en el sitio o característica. Los ácidos nucleicos están generalmente unidos covalentemente, por ejemplo, a través de sus extremos 5', pero en algunos casos son posibles otros medios de unión. Los ácidos nucleicos en un grupo pueden ser de cadena sencilla o de cadena doble. En algunas realizaciones, pero no en todas, los grupos se fabrican mediante un método de amplificación en fase sólida conocido como amplificación de puente. Ejemplos de configuraciones para grupos y métodos para su producción se exponen, por ejemplo, en la patente de EE.UU. N° 5.641.658; publ. de patente de EE.UU. N° 2002/0055100; patente de EE.UU. N° 7.115.400; Publ. de patente de EE.UU. N° 2004/0096853; Publ. de patente de EE.UU. N° 2004/0002090; Publ. de patente de EE.UU. N° 2007/0128624; y Publ. de patente de EE.UU. N° 2008/0009420.

Como se usa en la presente memoria, el término "diferente", cuando se usa en referencia a ácidos nucleicos, significa que los ácidos nucleicos tienen secuencias de nucleótidos que no son iguales entre sí. Dos o más ácidos nucleicos pueden tener secuencias de nucleótidos que sean diferentes a lo largo de toda su longitud. Alternativamente, dos o más ácidos nucleicos pueden tener secuencias de nucleótidos que sean diferentes a lo largo de una porción sustancial de su longitud. Por ejemplo, dos o más ácidos nucleicos pueden tener porciones de secuencia de nucleótidos diana que sean diferentes entre sí mientras tienen también una región de secuencia universal que son iguales entre sí. Generalmente, cuando dos especies se denominan en este documento "diferentes", una de las especies tendrá una propiedad estructural que no es la misma que las propiedades estructurales de la segunda especie. Por ejemplo, dos especies poliméricas diferentes (tales como dos proteínas) pueden tener diferentes secuencias de subunidades monoméricas (tales como diferentes secuencias de aminoácidos para dos proteínas diferentes).

Tal como se utiliza en este documento, la expresión "cada uno", cuando se usa en referencia a una colección de elementos, tiene por objeto identificar un elemento individual en la colección, pero no se refiere necesariamente a cada elemento de la colección. Pueden ocurrir excepciones si la descripción explícita o el contexto dictan claramente lo contrario.

Tal como se utiliza en la presente memoria, el término "sitio" significa un lugar en una matriz donde está presente una especie particular de molécula. Un sitio puede contener solamente una sola molécula o puede contener una población de varias moléculas de la misma especie. Los sitios de una matriz son típicamente discretos. Los sitios discretos pueden ser contiguos o pueden tener espacios entre sí.

Tal como se utiliza en la presente memoria, la expresión "ácido nucleico" se pretende que sea consistente con su uso en la técnica e incluye ácidos nucleicos naturales o análogos funcionales de los mismos. Los análogos funcionales particularmente útiles son capaces de hibridarse con un ácido nucleico de una manera específica de la secuencia o pueden utilizarse como molde para la replicación de una secuencia de nucleótidos particular. Los ácidos nucleicos naturales tienen generalmente una cadena principal que contiene enlaces fosfodiéster. Una estructura análoga puede tener un enlace de cadena principal alternativo que incluya cualquiera de una variedad de las conocidas en la técnica. Los ácidos nucleicos naturales tienen generalmente un azúcar desoxirribosa (por ejemplo, encontrado en el ácido desoxirribonucleico (ADN)) o un azúcar ribosa (por ejemplo, encontrado en el ácido ribonucleico (ARN)). Un ácido nucleico puede contener cualquiera de una variedad de análogos de estos restos de azúcar que se conocen en la técnica. Un ácido nucleico puede incluir 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 conocen en la técnica bases no nativas útiles que se pueden incluir en un ácido nucleico. El término "diana", cuando se usa en referencia a un ácido nucleico, está destinado a ser un identificador semántico para el ácido nucleico en el contexto de un método o composición establecido en la presente memoria y no limita necesariamente la estructura o función del ácido nucleico más allá de lo que de otra manera se indicaría explícitamente.

Tal como se utiliza en la presente memoria, la expresión "molde de ácido nucleico" se refiere a un ácido nucleico o porción del mismo que es capaz de usarse como una guía para la replicación catalizada por polimerasa. Una molécula de ácido nucleico puede incluir múltiples moldes a lo largo de su longitud o, alternativamente, sólo se puede usar un único molde en una realización particular de la presente invención. Un molde de ácido nucleico también puede funcionar como guía para la extensión de cebadores catalizada por ligasa.

Tal como se utiliza en la presente memoria, el término "nucleótido" o la expresión "análogo de nucleótido" pretende incluir nucleótidos naturales, nucleótidos no naturales, ribonucleótidos, desoxirribonucleótidos, didesoxirribonucleótidos y otras moléculas conocidas como nucleótidos. Por ejemplo, los términos se usan en este documento para referirse en general a un resto nucleósido (ya sea ribosa, desoxirribosa, o análogo del mismo) que incluye un resto base y opcionalmente unido a uno o más restos fosfato. El término puede usarse para referirse a una unidad de monómero que está presente en un polímero, por ejemplo, para identificar una subunidad presente en una cadena de ADN o ARN. El término también se puede usar para referirse a una molécula monomérica que no

está necesariamente presente en un polímero, por ejemplo, una molécula que sea capaz de ser incorporada en un polinucleótido de una manera dependiente del molde por una polimerasa.

Ejemplos de nucleótidos incluyen, pero no se limitan a, ribonucleótido monofosfato (a veces denominado ribonucleósido monofosfato), ribonucleótido difosfato (a veces denominado ribonucleósido difosfato), ribonucleótido trifosfato (a veces denominado ribonucleósido trifosfato), desoxinucleótido monofosfato (a veces denominado desoxinucleósido monofosfato), desoxinucleótido difosfato (a veces denominado desoxinucleósido difosfato) y desoxinucleótido trifosfato (a veces denominado desoxinucleósido trifosfato). Para mayor claridad cuando se desea distinguir componentes de ARN de componentes de ADN, el término "ribonucleótido" se puede usar para especificar nucleótidos de ARN, tales como trifosfato de ribouridina, trifosfato de riboguanidina, trifosfato de ribocitadina o trifosfato de riboadenosina; y el término "desoxinucleótido" puede usarse para especificar nucleótidos de ADN, tales como desoxitimidina trifosfato, desoxiguanidina trifosfato, desoxicitidina trifosfato y desoxiadenosina trifosfato. En realizaciones particulares, los nucleótidos son "extensibles", por ejemplo, carecen de un resto de bloqueo de extensión en el hidroxilo 3' o en cualquier otra posición sobre el nucleótido. En otras realizaciones, los nucleótidos están 'bloqueados', teniendo un resto que impide que la posición 3' participe en la extensión mediante la adición de otro nucleótido u oligonucleótido.

Como se usa en la presente memoria, el término "cebador" significa un ácido nucleico que tiene una secuencia que se une a un sitio de unión de cebador en o cerca de una secuencia de molde. Generalmente, el cebador se une en una configuración que permite la replicación del molde, por ejemplo, a través de la extensión de polimerasa del cebador. El cebador puede ser una primera porción de una molécula de ácido nucleico que se une a una segunda porción de la molécula de ácido nucleico, siendo la primera porción una secuencia de cebador y la segunda porción una secuencia de unión de cebador (por ejemplo, un cebador de horquilla). Alternativamente, el cebador puede ser una primera molécula de ácido nucleico que se une a una segunda molécula de ácido nucleico que tiene la secuencia molde. Un cebador puede consistir en ADN, ARN o análogos de los mismos.

Tal como se utiliza en la presente memoria, la expresión "producto de extensión de cebador" significa un cebador que ha sido modificado por la adición de al menos un análogo de nucleótido. Por ejemplo, un cebador puede modificarse por la adición de uno o más análogos de nucleótido a su extremo 3' (por ejemplo, por catálisis de polimerasa), formando de este modo un producto de extensión de cebador. Un producto de extensión de cebador puede producirse alternativamente por ligación de un oligonucleótido al extremo 3' o 5' de un cebador. En este caso, el producto de extensión del cebador se extiende por una longitud equivalente a la longitud del oligonucleótido. Un producto de extensión de cebador puede ser al menos 1, 2, 5, 10, 500, 1000 o más nucleótidos más largo que el cebador. Alternativamente o adicionalmente, un producto de extensión de cebador puede ser no mayor que 1, 2, 5, 10, 500 o 1000 nucleótidos más largo que el cebador. Por ejemplo, el uso de un nucleótido bloqueado proporciona un producto de extensión que es al menos 1 nucleótido más largo que el cebador y también no más de 1 nucleótido más largo que el cebador.

Tal como se utiliza en la presente memoria, la referencia a la manipulación de forma "selectiva" (o manipulación "selectiva") de una primera cosa comparada con una segunda cosa pretende significar que la manipulación tiene un efecto mayor sobre la primera cosa comparado con el efecto sobre la segunda cosa. La manipulación no necesita tener ningún efecto sobre la segunda cosa. La manipulación puede tener un efecto sobre la primera cosa que es al menos 1%, 5%, 10%, 25%, 50%, 75%, 90%, 95% o 99% mayor que el efecto sobre la segunda cosa. La manipulación puede tener un efecto sobre la primera cosa que es al menos 2 veces, 3 veces, 4 veces, 5 veces, 10 veces, 100 veces,  $1 \times 10^3$  veces,  $1 \times 10^4$  veces o  $1 \times 10^6$  veces que el efecto sobre la segunda cosa. La manipulación puede incluir, por ejemplo, la modificación, el contacto, el tratamiento, el cambio, la escisión (por ejemplo, de un enlace químico), la escisión fotoquímica (por ejemplo, de un enlace químico), la formación (por ejemplo, de un enlace químico), la formación fotoquímica (por ejemplo, de un enlace químico), la modificación covalente, modificación no covalente, destrucción, foto-ablación, eliminación, síntesis, polimerización, fotopolimerización, amplificación (por ejemplo, de un ácido nucleico), el copiado (por ejemplo, de un ácido nucleico), la extensión (por ejemplo, de un ácido nucleico), ligación (por ejemplo, de un ácido nucleico), u otra manipulación establecida en la presente memoria o conocida de otro modo en la técnica.

Como se usa en la presente memoria, el término "secuencia", cuando se usa en referencia a un ácido nucleico, se refiere al orden de los nucleótidos (o bases) en los ácidos nucleicos. En los casos en que están presentes diferentes especies de nucleótidos en el ácido nucleico, la secuencia incluye una identificación de la especie de nucleótido (o base) en las respectivas posiciones en el ácido nucleico. Una secuencia es una propiedad de todo o parte de una molécula de ácido nucleico. El término puede usarse de manera similar para describir el orden y la identidad posicional de unidades monómeras en otros polímeros tales como unidades monómeras de aminoácidos de polímeros de proteína.

Como se usa en este documento, el término "especie" se usa para identificar moléculas que comparten la misma estructura química. Por ejemplo, una mezcla de nucleótidos puede incluir varias moléculas de dCTP. Se entenderá que las moléculas de dCTP son de la misma especie que la otra, pero una especie diferente en comparación con dATP, dGTP, dTTP, etc. De forma similar, las moléculas de ADN individuales que tienen la misma secuencia de nucleótidos son de la misma especie, mientras que las moléculas de ADN con diferentes secuencias son especies

diferentes. Como otro ejemplo, una polimerasa de ADN es una especie de polimerasa diferente en comparación con una ARN polimerasa (incluso si las dos polimerasas derivan del mismo organismo).

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

5 La presente descripción proporciona un método para secuenciar moldes de ácido nucleico. El método puede incluir las etapas de (a) proporcionar un conjunto de sitios, en el que cada sitio incluye un primer molde de ácido nucleico y un segundo molde de ácido nucleico, en el que el primer molde de ácido nucleico tiene una secuencia que es diferente de la secuencia del segundo molde de ácido nucleico; (b) extender un primer cebador unido al primer molde usando una primera especie de polimerasa y un primer conjunto de análogos de nucleótidos, produciendo de este modo un primer producto de extensión de cebador que tiene un primer análogo de nucleótido en cada uno de los sitios; (c) extender un segundo cebador unido al segundo molde usando una segunda especie de polimerasa y un segundo conjunto de análogos de nucleótidos, produciendo de este modo un segundo producto de extensión de cebador que tiene un segundo análogo de nucleótido en cada uno de los sitios, en donde la primera especie de polimerasa es diferente de la segunda especie de polimerasa y en donde el primer conjunto de análogos de nucleótidos es diferente del segundo conjunto de análogos de nucleótidos; (d) detectar el primer producto de extensión del cebador y el segundo producto de extensión del cebador en cada uno de los sitios; y (e) repetir las etapas (b) a (d), determinando así las diferentes secuencias del primer molde y del segundo molde en cada uno de los sitios.

20 También se proporciona en la presente memoria un método para secuenciar moldes de ácido nucleico que incluye las etapas de (a) proporcionar un primer molde de ácido nucleico y un segundo molde de ácido nucleico, en donde el primer molde de ácido nucleico tiene una secuencia que es diferente de la secuencia del segundo molde de ácido nucleico; (b) extender un primer cebador unido al primer molde usando una primera especie de polimerasa y un primer conjunto de análogos de nucleótidos, produciendo de este modo un primer producto de extensión de cebador que tiene un primer análogo de nucleótido; (c) extender un segundo cebador unido al segundo molde usando una segunda especie de polimerasa y un segundo conjunto de análogos de nucleótidos, produciendo de este modo un segundo producto de extensión del cebador que tiene un segundo análogo de nucleótido, en el que la primera especie de polimerasa es diferente de la segunda especie de polimerasa y en el que el primer conjunto de análogo de nucleótido es diferente del segundo conjunto de análogo de nucleótido, (d) detectar el primer producto de extensión de cebador y el segundo producto de extensión de cebador usando un detector que tiene una resolución que es menor que la separación espacial entre el primer producto de extensión de cebador y el segundo producto de extensión del cebador; y (e) repetir las etapas (b) a (d), determinando con ello las diferentes secuencias del primer molde y del segundo molde.

35 Como se ha expuesto anteriormente, un método de la presente descripción puede incluir una etapa de proporcionar moldes de ácido nucleico primero y segundo, en donde las secuencias para los dos moldes son diferentes. Las dos secuencias molde pueden ser porciones de una sola molécula de ácido nucleico o, alternativamente, las dos secuencias molde pueden estar localizadas en moléculas separadas. Como se expone con más detalle en otra parte del presente documento, las dos secuencias molde pueden estar en una proximidad que esté demasiado cerca para resolverse espacialmente con el sistema de detección utilizado. Sin embargo, los métodos de detección ortogonales de la presente descripción permiten que estas secuencias molde sean distinguidas. El esquema de detección ortogonal se ejemplifica para dos secuencias molde, pero se puede usar con dos o más secuencias molde. Por consiguiente, un sistema o método expuesto en la presente memoria puede incluir al menos 2, 3, 4, 5, 10 o más secuencias molde que estén en estrecha proximidad, por ejemplo en una única molécula de ácido nucleico, en un único sitio de una matriz o, por el contrario, en una proximidad que esté demasiado cerca para resolver espacialmente con el sistema de detección utilizado.

45 Los ácidos nucleicos diana usados en la presente invención pueden estar compuestos de ADN, ARN o análogos de los mismos. La fuente de los ácidos nucleicos diana puede ser ADN genómico, ARN mensajero u otros ácidos nucleicos de fuentes nativas. En algunos casos, los ácidos nucleicos diana que se derivan de tales fuentes pueden amplificarse antes de su uso en un método o composición de la presente invención.

50 Ejemplos de muestras biológicas de las que se pueden derivar ácidos nucleicos diana incluyen, por ejemplo, las de mamíferos tales como roedor, ratón, rata, conejo, cobaya, ungulado, caballo, oveja, cerdo, cabra, vaca, gato, primate, humano o primate no humano; plantas tal como *Arabidopsis thaliana*, maíz, sorgo, avena, trigo, arroz, canola o soja; algas tales como *Chlamydomonas reinhardtii*; nematodos tal como *Caenorhabditis elegans*; insectos tal como *Drosophila melanogaster*, mosquito, mosca de la fruta, abeja de la miel o araña; peces tal como el pez cebra; reptiles; anfibios tal como una rana o *Xenopus laevis*; un dictyostelium discoideum; hongos tal como *Pneumocystis carinii*, *Takifugu rubripes*, levadura, *Saccharomyces cerevisiae* o *Schizosaccharomyces pombe*; o un *Plasmodium falciparum*. Los ácidos nucleicos diana también pueden derivarse de un procarionta tal como de una bacteria, *Escherichia coli*, *staphylococci* o *mycoplasma pneumoniae*; un archae; un virus tal como el virus de la hepatitis C o el virus de la inmunodeficiencia humana; o un viroide. Los ácidos nucleicos diana pueden derivarse de un cultivo homogéneo o población de los organismos anteriores o, alternativamente, de una colección de varios organismos diferentes, por ejemplo, en una comunidad o ecosistema. Los ácidos nucleicos se pueden aislar usando métodos conocidos en la técnica incluyendo, por ejemplo, los descritos en Sambrook y col., *Molecular Cloning: A*

Laboratory Manual, 3ª edición, Cold Spring Harbor Laboratory, Nueva York (2001) o Ausubel et al, Current Protocols in Molecular Biology, John Wiley y Sons, Baltimore, Md. (1998).

En realizaciones particulares, una muestra de ácido nucleico puede ser modificada o preparada para su uso en uno o más de los métodos expuestos en este documento. En algunos casos se desea añadir uno o más sitios de unión del cebador a un ácido nucleico. Pueden usarse técnicas de biología molecular conocidas para introducir sitios de unión de cebadores aguas arriba de las respectivas secuencias molde, por ejemplo, mediante la inserción de un adaptador que tiene el sitio de unión del cebador, mutación para crear el sitio de unión del cebador, ligación de un adaptador que tiene el sitio de unión del cebador, etc. Métodos útiles se describen en Sambrook et al, supra y Ausubel y col., Supra. El Ejemplo I proporciona una ilustración de una técnica basada en la marcación. La marcación es particularmente útil para introducir uno o más sitios de unión del cebador y puede llevarse a cabo, por ejemplo, utilizando las técnicas expuestas en las patentes de EE.UU. Números 6.294.385 y 8.383.345, y la publicación PCT N° WO 2012/106546.

Se entenderá que en algunos casos las regiones de secuencia que aparecen naturalmente que residen aguas arriba de las respectivas secuencias molde pueden ser explotadas como sitios de unión de cebadores en un método expuesto en este documento. Pueden usarse métodos similares a los ejemplificados anteriormente para sitios de unión de cebadores para introducir otros elementos de secuencia deseados tales como elementos promotores para secuencias de extensión o de marcador basadas en ARN polimerasa. Un método ejemplar para crear fragmentos de ácido nucleico que tiene cada uno una primera secuencia molde en proximidad con un sitio de cebado de ADN y una segunda secuencia molde en proximidad con un sitio de cebado de ARN y promotor de ARN polimerasa se muestra en la Fig. 4 y se describe en el Ejemplo I.

Los sitios de cebado universales son particularmente útiles para aplicaciones múltiple de los métodos expuestos en este documento. El término "universal", cuando se usa en referencia a ácidos nucleicos, significa una región de secuencia que es común a dos o más moléculas de ácido nucleico, donde las moléculas también tienen regiones de secuencias diferentes. Una secuencia universal presente en diferentes miembros de una colección de moléculas puede permitir la replicación, amplificación o detección de múltiples secuencias diferentes usando una única especie de cebador universal que es complementario a la secuencia universal. Así, un cebador universal incluye una secuencia que se puede hibridar específicamente con una secuencia universal. Ejemplos de métodos para unir secuencias universales a una colección de ácidos nucleicos diana se pueden encontrar en la publicación de solicitud de patente de EE.UU. N°. 2007/0128624 A1.

Pueden usarse cualquiera de una variedad de promotores según sea apropiado para la ARN polimerasa particular que se va a usar. Por ejemplo, se puede usar un promotor bacteriano con una ARN polimerasa bacteriana o un promotor eucariota puede usarse con una ARN polimerasa eucariótica. Un promotor estará localizado generalmente cerca del molde que se va a detectar, aguas arriba del sitio de unión del cebador de ARN y en la misma cadena que el molde. Se pueden usar técnicas estándar de síntesis de ácidos nucleicos y/o biológicas moleculares para crear una construcción promotora funcional en un ácido nucleico diana. Un promotor particularmente útil es un promotor bidireccional tal como los presentes en pares de genes bidireccionales de mamíferos. Un promotor bidireccional puede ser útil para aplicaciones de secuenciación final pareadas tales como las expuestas con más detalle a continuación.

En algunas realizaciones, los ácidos nucleicos diana se pueden obtener como fragmentos de uno o más ácidos nucleicos mayores. La fragmentación puede llevarse a cabo usando cualquiera de una variedad de técnicas conocidas en la técnica incluyendo, por ejemplo, nebulización, sonicación, escisión química, escisión enzimática o cizallamiento físico. La fragmentación también puede resultar del uso de una técnica de amplificación particular que produce amplicones copiando solamente una porción de un ácido nucleico más grande. Por ejemplo, la amplificación por PCR produce fragmentos que tienen un tamaño definido por la longitud del fragmento entre los cebadores flanqueantes usados para la amplificación.

Una población de ácidos nucleicos diana, o amplicones de los mismos, puede tener una longitud de cadena promedio que se desea o es apropiada para una aplicación particular de los métodos o composiciones expuestos en este documento. Por ejemplo, la longitud media de la cadena puede ser inferior a aproximadamente 100.000 nucleótidos, 50.000 nucleótidos, 10.000 nucleótidos, 5.000 nucleótidos, 1.000 nucleótidos, 500 nucleótidos, 100 nucleótidos o 50 nucleótidos. Alternativamente o adicionalmente, la longitud media de la cadena puede ser mayor que aproximadamente 10 nucleótidos, 50 nucleótidos, 100 nucleótidos, 500 nucleótidos, 1.000 nucleótidos, 5.000 nucleótidos, 10.000 nucleótidos, 50.000 nucleótidos o 100.000 nucleótidos. La longitud media de la cadena para la población de ácidos nucleicos diana, o sus amplicones, puede estar en un intervalo entre un valor máximo y mínimo establecido anteriormente. Se entenderá que los amplicones generados en un sitio de amplificación (o fabricados de otro modo o utilizados en la presente memoria) pueden tener una longitud media de la cadena que esté en un intervalo entre un límite superior e inferior seleccionado entre los ejemplificados anteriormente.

En algunos casos, se puede producir una población de ácidos nucleicos diana o configurarse de otra forma para que tenga una longitud máxima para sus miembros. Por ejemplo, la longitud máxima para los miembros que se fabrican o utilizan como se establece en la presente memoria puede ser inferior a aproximadamente 100.000 nucleótidos, 50.000 nucleótidos, 10.000 nucleótidos, 5.000 nucleótidos, 1.000 nucleótidos, 500 nucleótidos, 100 nucleótidos o 50



nucleótidos. Alternativamente o adicionalmente, se puede producir una población de ácidos nucleicos diana, o sus amplicones, bajo condiciones o ser configuradas de otro modo para que tengan una longitud mínima para sus miembros. Por ejemplo, la longitud mínima para los miembros que se usan en una o más etapas de un método establecido en la presente memoria o que están presentes en una composición particular puede ser más de  
5 aproximadamente 10 nucleótidos, 50 nucleótidos, 100 nucleótidos, 500 nucleótidos, 1.000 nucleótidos, 5.000 nucleótidos, 10.000 nucleótidos, 50.000 nucleótidos o 100.000 nucleótidos. La longitud de cadena máxima y mínima para los ácidos nucleicos diana en una población puede estar en un intervalo entre un valor máximo y mínimo establecido anteriormente. Se entenderá que los amplicones generados en un sitio de amplificación (o de otro modo fabricados o utilizados en este documento) pueden tener longitudes de cadena máximas y/o mínimas en un intervalo  
10 entre los límites superior e inferior ejemplificados anteriormente.

Puede usarse cualquiera de una variedad de técnicas de amplificación conocidas para aumentar la cantidad de secuencias de molde presentes para su uso en un método expuesto en este documento. Las técnicas ejemplares incluyen, pero no se limitan a, reacción en cadena de la polimerasa (PCR), amplificación de círculo rodante (RCA), amplificación de desplazamiento múltiple (MDA) o amplificación de cebador aleatoria (RPA) de moléculas de ácido  
15 nucleico que tienen secuencias de molde. Se entenderá que la amplificación de los ácidos nucleicos diana antes del uso en un método o composición expuestos en este documento es opcional. Como tal, los ácidos nucleicos diana no se amplificarán antes de su uso en algunas realizaciones de los métodos y composiciones en este documento expuestos. Los ácidos nucleicos diana pueden derivarse opcionalmente de bibliotecas sintéticas. Los ácidos nucleicos sintéticos pueden tener composiciones nativas de ADN o ARN o pueden ser análogos de los mismos.  
20 Pueden utilizarse también métodos de amplificación en fase sólida, que incluyen, por ejemplo, amplificación en clúster, amplificación de puente u otros métodos expuestos a continuación en el contexto de los métodos basados en matrices.

Un ácido nucleico usado en un método expuesto en este documento puede ser una fase en solución o una fase sólida. El ácido nucleico, en fase de disolución, es generalmente soluble, pero también puede estar en una forma suspendida que sea capaz de precipitarse, como es el caso de algunas especies de ácidos nucleicos grandes tales como los cromosomas o los Nanoballs de ácidos nucleicos (véase, por ejemplo, publicación de patente de EE.UU. N°.  
25 2007/0099208 A1). Un ácido nucleico que está en fase sólida puede estar en un soporte en fase sólida o sobre éste. Ejemplos de soportes en fase sólida incluyen los fabricados a partir de vidrio, nitrocelulosa, sílice, metal, plástico y otros materiales expuestos en otra parte en este documento, por ejemplo, con respecto a los formatos de matrices y células de flujo. Similarmente, un ácido nucleico puede estar en un soporte semisólido o sobre éste, tal como en un gel. Geles ejemplares que son útiles incluyen, pero no se limitan a, aquellos que tienen una estructura coloidal, tal como agarosa; estructura de malla polimérica, tal como gelatina; o estructura de polímero reticulado, tal como poli(acrilamida). Los hidrogeles son particularmente útiles tales como los expuestos en la publicación de patente de EE.UU. N°.  
30 2011/0059865 A1 y solicitud de patente de EE.UU. N°. Ser. 13/784.368.

La fijación de un ácido nucleico a un soporte, ya sea rígido o semirrígido, puede ocurrir por medio de un enlace(s) covalente o no covalente. Los enlaces ejemplares se exponen en las patentes de EE.UU. Nos. 6.737.236; 7.259.258; 7.375.234 y 7.427.678; y la publicación de patente de EE.UU. N°. 2011/0059865 A1. En algunas realizaciones, un ácido nucleico u otro componente de reacción puede unirse a un gel u otro soporte semisólido que a su vez esté unido o adherido a un soporte de fase sólida. En tales realizaciones, se entenderá que el ácido  
40 nucleico u otro componente de reacción está en fase sólida.

Los sistemas de detección ortogonales pueden basarse en el uso de dos o más sistemas de reactivos para la extensión del cebador, en el que los componentes de los dos sistemas de reactivos no reaccionan sustancialmente inespecíficamente. Por ejemplo, un primer sistema reactivo puede incluir una ADN polimerasa, desoxirribonucleótidos y un cebador de ADN; y un segundo sistema reactivo puede incluir una ARN polimerasa, ribonucleótidos y un cebador de ARN. Ambos sistemas son capaces de actuar sobre un molde de ADN, por ejemplo, uno que tiene un primer sitio de cebado que es complementario al cebador de ADN y un segundo sitio de cebado que es complementario al cebador de ARN. Sin embargo, la ADN polimerasa es específica para el cebador de ADN y desoxirribonucleótidos de tal manera que extiende selectivamente el cebador de ADN con los desoxirribonucleótidos en lugar de los ribonucleótidos. Por el contrario, la ARN polimerasa es específica para el  
45 cebador de ARN y los ribonucleótidos de tal manera que extiende selectivamente el cebador de ARN con los ribonucleótidos en lugar de los desoxirribonucleótidos. De forma similar, la ortogonalidad se puede conseguir usando otros sistemas de reactivos específicos tales como una polimerasa modificada que incorpora selectivamente HNA (ácidos nucleicos de 1,5-anhidrohexitol) en un cebador hecho de monómeros de HNA. La extensión del cebador basado en HNA es ortogonal a la ADN polimerasa y los sistemas de extensión de ARN polimerasa. Las condiciones ejemplares y los reactivos que pueden usarse para la extensión de cebadores a base de HNA se describen en Pinheiro et al, Science, 336 (6079): 341-344 (2012) y Cozens et al., Pro Nat. Acad. Sci. USA, 109 (21): 8067 - 8072 (2012).  
50  
55

De acuerdo con las realizaciones ejemplares expuestas anteriormente, los desoxirribonucleótidos pueden considerarse una clase ortogonal de nucleótidos con respecto a los ribonucleótidos y HNA. De manera similar, en el contexto de realizaciones particulares, las clases de ADN polimerasas y ARN polimerasas son ortogonales entre sí y las clases de cebadores de ADN y cebadores de ARN son ortogonales entre sí. Generalmente, la ortogonalidad puede explotarse en un método expuesto en este documento cuando una primera polimerasa sea selectiva para una  
60

primera clase de análogos de nucleótidos comparado con una segunda clase de análogos de nucleótidos y en donde una segunda polimerasa es selectiva para la segunda clase de análogos de nucleótidos comparado con la primera clase de análogos de nucleótidos. De forma similar, puede existir ortogonalidad cuando la primera polimerasa sea selectiva para una primera clase de cebador en comparación con una segunda clase de cebador y cuando la segunda polimerasa sea selectiva para la segunda clase de cebador en comparación con la primera clase de cebador.

Puede utilizarse cualquiera de una variedad de polimerasas en el método o composición expuesto en el presente documento, incluyendo, por ejemplo, enzimas basadas en proteínas aisladas de sistemas biológicos y variantes funcionales de las mismas. La referencia a una polimerasa particular, tal como las ejemplificadas a continuación, se entenderá que incluye variantes funcionales de la misma a menos que se indique lo contrario. Una función particularmente útil de una polimerasa es catalizar la polimerización de una cadena de ácido nucleico utilizando un ácido nucleico existente como molde. Otras funciones que son útiles se describen en otra parte del presente documento. Ejemplos de polimerasas útiles incluyen ADN polimerasas, transcriptasas inversas y ARN polimerasas.

Una polimerasa que tiene una actividad de exonucleasa de corrección intrínseca de 3' a 5' puede ser útil para algunas realizaciones. Las polimerasas que carecen sustancialmente de actividad de exonucleasa de corrección de 3' a 5' también son útiles en algunas realizaciones, por ejemplo, en la mayoría de las realizaciones de secuenciación. La ausencia de actividad de exonucleasa puede ser una característica de tipo salvaje o una característica impartida por una estructura de polimerasa modificada o variante. Por ejemplo, el fragmento de Klenow exo menos es una versión mutada del fragmento de Klenow que carece de actividad de exonucleasa de corrección 3' a 5'.

Dependiendo de la realización que se vaya a usar, una polimerasa puede ser tanto termófila como inactivable por calor. Las polimerasas termofílicas son típicamente útiles para condiciones a altas temperaturas o en condiciones de termociclado tales como las empleadas para las técnicas de reacción en cadena de la polimerasa (PCR). Ejemplos de polimerasas termofílicas incluyen, pero sin limitación, ADN polimerasa de 9°N, ADN polimerasa Taq, ADN polimerasa Phusion®, ADN polimerasa Pfu, ADN polimerasa RB69, ADN polimerasa KOD y ADN polimerasa VentR®. La mayoría de las polimerasas aisladas de organismos no termofílicos son inactivables por calor. Ejemplos son las ADN polimerasas de fagos. Se entenderá que las polimerasas de cualquiera de una variedad de fuentes pueden ser modificadas para aumentar o disminuir su tolerancia a las condiciones a altas temperaturas. Las polimerasas particularmente útiles para incorporar nucleótidos que tienen marcadores y/o restos terminales reversibles se describen en el documento de EE.UU. 2006/0281109 A1, que se incorpora a la presente memoria como referencia.

Otro sistema de reactivo ortogonal de extensión de cebador es un sistema basado en ligasa que es selectivo para la incorporación de oligonucleótidos en lugar de nucleótidos monoméricos que se incorporan mediante los sistemas de extensión basados en polimerasa descritos anteriormente. Un sistema reactivo de ADN ligasa es totalmente ortogonal con un sistema de reactivos basado en ARN polimerasa cuando se usa en condiciones en las que el cebador de ADN se extiende por la ADN ligasa pero no por la ARN polimerasa y en el que un cebador de ARN se extiende por la ARN polimerasa pero no por la ADN ligasa. La extensión por ligación puede llevarse a cabo en una aplicación de secuenciación usando una población de oligonucleótidos de sonda parcialmente aleatorios que tienen un esquema de codificación de una o dos bases. Las técnicas de extensión basadas en ligación que pueden usarse para la detección en una reacción de extensión tal como en un contexto de secuenciación se exponen en McKernan et al, Genome Research 19 (9): 1527-41 (2009); Shendure et al., Science 309: 1728-1732 (2005); y las patentes de Estados Unidos N° 5.599.675 y 5.750.341.

La manipulación ortogonal y la detección de acuerdo con la presente descripción no requieren que dos secuencias de molde difieran en cada posición a lo largo de su longitud. Por el contrario, el mismo resto de base puede estar presente en posiciones que se detectan en un primer molde y un segundo molde, respectivamente. Las dos posiciones se pueden distinguir basándose en las características distinguibles de los marcadores presentes en los sistemas de reactivos ortogonales y en la especificidad de los sistemas de reactivos para extender el cebador apropiado. Esta información se puede utilizar a su vez para detectar de manera distinguible las dos secuencias de molde diferentes, incluso si las dos posiciones se detectan simultáneamente usando un detector que tiene una resolución que es demasiado baja para resolver puntos a una distancia equivalente a la separación de las dos secuencias de molde.

Puede utilizarse cualquiera de una variedad de marcadores. Un resto de marcador que es particularmente útil cuando se usa para la detección de un análogo de nucleótidos, puede ser cualquier parte del análogo de nucleótidos que proporcione una característica distinguible cuando se compara con otras moléculas presentes en su entorno. La característica distinguible puede ser, por ejemplo, una señal óptica tal como la absorbancia de radiación, la emisión de fluorescencia, la emisión de luminiscencia, la vida de fluorescencia, la polarización de fluorescencia, o similares; afinidad de unión para un ligando o receptor; propiedades magnéticas; propiedades eléctricas; carga; masa; radiactividad o similares. Ejemplos de restos de marcadores incluyen, sin limitación, un fluoróforo, luminóforo, cromóforo, isótopo radiactivo, marcador de masa, marcador de carga, marcador de rotación, receptor, ligando o similares. El resto marcador puede formar parte de un nucleótido que sea una unidad monomérica presente en un

polímero de ácido nucleico o el resto marcador puede ser una parte de un análogo de nucleótido libre (por ejemplo, un nucleótido trifosfato).

Los fluoróforos son particularmente útiles e incluyen, por ejemplo, nanocristales fluorescentes; puntos cuánticos, fluorescencia, rodamina, tetrametilrhodamina, eosina, eritrosina, cumarina, metil-cumarinas, pireno, verde de Malacita, Cy3, Cy5, estilbena, Lucifer Yellow, Cascade Blue, Texas Red, colorantes Alexa, colorantes SETA, colorantes Atto, ficoeritrina, Bodipy, y sus análogos. Las sondas ópticas útiles se describen en Lakowicz, Principles of Fluorescence Spectroscopy, 3ª Ed. Springer (2006); Haugland, Handbook of Fluorescent Probes and Research Products 9th Ed., Molecular Probes, Inc, (2002); Shapiro, Practical Flow Cytometry, 4ª Ed., John Wiley & Sons (2003); documentos WO 98/59066; WO 91/06678 o la publicación de solicitud de patente de EE.UU. N.º. 2010/0092957 A1.

Otros marcadores, algunas de las cuales son marcadores no ópticos, pueden utilizarse en las realizaciones de los métodos y composiciones expuestas en la presente memoria. Los ejemplos incluyen, sin limitación, un marcador isotópico tal como un isótopo radioactivo o pesado naturalmente no abundante; sustancia magnética; material rico en electrones tal como un metal; marcador electroquimioluminiscente tal como Ru(bpy)<sup>32+</sup>; o fracción que puede ser detectada en base a una característica magnética nuclear, paramagnética, eléctrica, de carga a masa o térmica. Los marcadores también pueden incluir partículas magnéticas o nanopartículas codificadas ópticamente. Dichos marcadores pueden detectarse usando métodos apropiados conocidos por los expertos en la técnica. Por ejemplo, se puede detectar un marcador cargado usando un detector eléctrico tal como los usados en sistemas de secuenciación comercialmente disponibles de Ion Torrent (Guilford, CT, una subsidiaria de Life Technologies) o sistemas de detección descritos en la publicación de solicitud de patente de EE.UU. Nos. 2009/0026082 A1; 2009/0127589 A1; 2010/0137143 A1; y 2010/0282617 A1. Se entenderá que para algunas realizaciones un análogo de nucleótido no necesitará tener uno o más de los marcadores expuestos en este documento.

Un resto de marcador se puede unir a un nucleótido en una variedad de formas. Las uniones ejemplares y las composiciones de marcadores que son útiles para los nucleótidos se exponen en Bentley et al., Nature 456: 53-59 (2008), documentos WO 04/018497; US 7.057.026; WO 91/06678; WO 07/123744; US 7.329.492; US 7.211.414; US 7.315.019; US 7.405.281 y US 2008/0108082.

En realizaciones particulares, por ejemplo, aquellas que utilizan la extensión del cebador cíclico en un enfoque de síntesis secuencial ortogonal, los nucleótidos pueden incluir restos terminadores reversibles. Los restos terminadores reversibles proporcionan una manera conveniente de controlar una reacción de extensión para añadir sólo un solo nucleótido a un cebador hasta que se lleva a cabo una etapa de desbloqueo posterior. Esto se puede entender en el contexto de un enfoque de secuenciación como sigue. Para iniciar un primer ciclo de secuenciación, uno o más nucleótidos marcados, ADN polimerasa, etc., pueden suministrarse a una serie de moldes de ácido nucleico unidos a un cebador. Opcionalmente, los nucleótidos pueden incluir un resto terminador reversible tal que la extensión posterior no pueda ocurrir hasta que se libere un agente de desbloqueo para eliminar el resto. Se pueden detectar dos o más marcadores añadidos a los sitios mediante las reacciones de extensión del cebador, por ejemplo, utilizando métodos o aparatos descritos en el presente documento. Se puede poner en contacto un reactivo de desbloqueo con la matriz (antes o después de la detección) para eliminar el terminador reversible. Los lavados se pueden llevar a cabo entre las diversas etapas de suministro. El ciclo puede entonces repetirse n veces para extender los cebadores por n nucleótidos, detectando así secuencias de longitud n. Ejemplos de técnicas de secuenciación y reactivos útiles se describen, por ejemplo, en Bentley et al., Nature 456: 53-59 (2008), los documentos WO 04/018497; US 7.057.026; WO 91/06678; WO 07/123744; US 7.329.492; US 7.211.414; US 7.315.019; US 7.405.281 y US 2008/0108082.

Un método de secuenciación ortogonal descrito en este documento se puede utilizar en un enfoque de secuenciación de extremos a pares. Generalmente, la secuenciación de extremos a pares implica la determinación de las secuencias en dos extremos de una región de secuencia de molde, en la que la longitud de la región de secuencia de molde es conocida. Se conocen métodos para fragmentar una muestra de ácido nucleico diana (por ejemplo, muestra de ADN genómico), fijar cebadores para acomodar lecturas de extremos emparejados y secuencia de lectura desde los extremos de los fragmentos y pueden llevarse a cabo como se describe, por ejemplo, en las patentes de EE.UU. Nos. 7.754.429; 8.017.335; y 8.192.930.

En el caso de una realización de síntesis secuencial ortogonal, pueden construirse fragmentos de ácido nucleico para tener dos secuencias de molde y se pueden obtener lecturas emparejadas de cada una de los dos moldes para obtener 4 lecturas de un solo fragmento. Las lecturas de extremos emparejados pueden facilitarse mediante el uso de un promotor bidireccional flanqueado por sitios de unión de la ARN polimerasa. Un ejemplo de construcción se muestra en la Fig. 5. En este ejemplo, la construcción incluye los sitios de cebado Lectura 1 y Lectura 1' para una primera lectura ortogonal. En la primera lectura ortogonal, un cebador de ADN puede hibridarse con el sitio de cebado Lectura 1 para permitir la lectura catalizada por ADN polimerasa de la secuencia en un primer extremo del Molde 1 (indicado por la flecha cerrada en el diagrama superior). También en la primera lectura ortogonal, el sitio de cebado Lectura 1' puede hibridarse con un cebador de ARN y debido a la proximidad del promotor bidireccional la ARN polimerasa puede leer un primer extremo del molde 2 (indicado por la flecha abierta en el diagrama superior). Se puede obtener una segunda lectura ortogonal por hibridación de un cebador de ADN al sitio de cebado Lectura 2 y la lectura del segundo extremo del Molde 2 e hibridación de un cebador de ARN al sitio de cebado Lectura 2' y la

lectura del segundo extremo del molde 1 (indicado por la flecha cerrada en el diagrama inferior). La proximidad del promotor bidireccional al sitio de cebado Lectura 2' permite que se produzca la extensión de la ARN polimerasa (indicada por la flecha abierta en el diagrama inferior). La construcción ejemplificada en la Fig. 5 puede hacerse, por ejemplo, usando los métodos descritos en la Fig. 4.

- 5 Un promotor bidireccional no es necesario para las lecturas de extremos emparejados usando una ARN polimerasa en una realización de síntesis de secuenciación ortogonal. Más bien, los sitios de cebado de ARN y sus promotores pueden situarse en los extremos de una construcción de 2 moldes y el adaptador que enlaza los dos moldes puede contener sitios de cebado de ADN. Tomando el constructo de la Fig. 5 a modo de ejemplo, las posiciones de los sitios de cebado Lectura 1 y Lectura 1' pueden intercambiarse, las posiciones de los sitios de cebado Lectura 2 y Lectura 2' se pueden intercambiar, el promotor bidireccional puede ser retirado y los promotores de ARN separados pueden estar situados aguas arriba de los sitios de cebado Lectura 1' y Lectura 2', respectivamente.

15 Una reacción de extensión de ácido nucleico, u otra reacción cíclica, que se lleve a cabo utilizando los procedimientos expuestos en la presente memoria, puede proceder durante uno o más ciclos. En realizaciones particulares, una reacción multiciclo puede incluir al menos 2 ciclos, 5 ciclos, 10 ciclos, 50 ciclos, 100 ciclos, 500 ciclos, 1.000 ciclos, 5.000 ciclos, 10.000 ciclos o más. Alternativamente o adicionalmente, una reacción puede tener un límite superior por el cual no más de 1 ciclo, 2 ciclos, 5 ciclos, 10 ciclos, 50 ciclos, 100 ciclos, 500 ciclos, 1.000 ciclos, 5.000 ciclos, o 10.000 ciclos aparecen. En algunas realizaciones, cada ciclo resultará en la incorporación de un único análogo de nucleótidos en un cebador extendido. En este caso, el número mínimo o máximo de ciclos ejemplificados anteriormente se puede entender para ejemplificar el número mínimo o máximo de nucleótidos incorporados en un producto de extensión en una reacción catalizada por polimerasa.

20 Algunas realizaciones pueden utilizar reacciones de extensión no cíclicas tales como la extensión de base única (SBE) o las reacciones de extensión de cebador específicas de alelo (ASPE). Los restos terminadores reversibles se pueden usar para la extensión no cíclica. Dado que una etapa de desbloqueo no es necesaria para estas reacciones no cíclicas, los nucleótidos pueden estar en su lugar terminados de forma no reversible. Por ejemplo, se pueden usar didesoxinucleótidos. Los ejemplos de reactivos y técnicas relacionadas para SBE, ASPE y otras técnicas de extensión no cíclicas útiles se describen, por ejemplo, en la patente de EE.UU. N° 7.670.810 y la publicación de solicitud de patente de EE.UU. Nos. 2003/0108867; 2003/0108900; 2003/0170684; 2003/0207295; o 2005/0181394.

25 Un ejemplo de un producto comercialmente disponible que utiliza una técnica de extensión no cíclica y que puede modificarse para aumentar el contenido de información a través de los métodos de detección ortogonales expuestos en la presente memoria es el producto de genotipificación Infinium® disponible en Illumina, Inc. (San Diego, CA).

30 Las reacciones cíclicas y no cíclicas por igual pueden incluir etapas en las que los componentes de reacción están separados entre sí o eliminados del entorno de reacción. Uno o más componentes de reacción pueden separarse, por ejemplo, por separación de componentes de fase sólida de componentes en fase líquida. Opcionalmente, se pueden incluir etapas de lavado con el fin de eliminar de forma más completa el componente(s) de fase líquida no deseado del componente(s) en fase sólida. Un recipiente de reacción particularmente útil para tales separaciones es una célula de flujo, tal como las comúnmente usadas en procedimientos de secuenciación cíclica. Ejemplos de células de flujo, métodos para su fabricación y métodos para su uso se describen en la publicación de solicitud de patente de EE.UU. Nos. 2010/0111768 A1 y 2012/0270305 A1; y el documento WO 05/065814.

35 Independientemente de que se utilicen métodos de separación en fase sólida, los componentes de reacción pueden eliminarse mediante cualquiera de una variedad de otras técnicas conocidas en la técnica, incluyendo extracción líquido-líquido, extracción en fase sólida, cromatografía, filtración, centrifugación o similares.

40 La detección puede llevarse a cabo en un método de la presente descripción usando un aparato adecuado para el marcador particular en uso. Por ejemplo, se puede usar un detector óptico tal como un detector de fluorescencia, detector de absorbancia, detector de luminiscencia o similar para detectar marcadores ópticos apropiados. Los sistemas diseñados para la detección basada en matrices son particularmente útiles. Por ejemplo, los sistemas ópticos para uso con los métodos en este documento expuestos pueden construirse para incluir diversos componentes y conjuntos como se describe en la patente de EE.UU. Nos. 8. 241.573; 7.329.860 y 8.039.817; y la publicación de solicitud de patente de EE.UU. Nos. 2009/0272914 A1 y 2012/0270305 A1.

45 Como se ha expuesto anteriormente, un método de la presente descripción puede incluir dos etapas de extensión de cebador ortogonales. Por ejemplo, se expone un método que incluye inter alia las etapas de (b) extender un primer cebador unido a un primer ácido nucleico usando una primera especie de polimerasa y un primer conjunto de análogos de nucleótidos, produciendo de este modo un primer producto de extensión de cebador que tiene un primer análogo de nucleótidos en cada uno de los sitios; y (c) extender un segundo cebador unido a un segundo ácido nucleico usando una segunda especie de polimerasa y un segundo conjunto de análogos de nucleótidos, produciendo de este modo un segundo producto de extensión de cebador que tiene un segundo análogo de nucleótido en cada uno de los sitios, en donde la primera especie de polimerasa es diferente de la segunda especie de polimerasa y en el que el primer conjunto de análogos de nucleótidos es diferente del segundo conjunto de análogos de nucleótidos. En algunas realizaciones, las etapas (b) y (c) se llevan a cabo simultáneamente. Alternativamente, las etapas (b) y (c) pueden llevarse a cabo secuencialmente, en cualquier orden. En cualquier

caso, la ortogonalidad de las reacciones de extensión del cebador permite distinguir los dos productos de extensión. Por tanto, ambos productos de extensión pueden estar simultáneamente presentes durante una etapa de detección y no necesitan ser resueltos espacialmente por el detector utilizado.

5 Una reacción multiplex puede utilizar un soporte en fase sólida. Un soporte en fase sólida puede ser útil para separar reacciones individuales de modo que cada una puede ser examinada por separado o individualmente. Por ejemplo, se pueden unir varios ácidos nucleicos diferentes en una mezcla al soporte en fase sólida. Los ácidos nucleicos se pueden unir al soporte en fase sólida en un formato de matriz.

10 En algunas realizaciones, se proporciona una matriz de sitios, en la que cada sitio incluye un primer molde de ácido nucleico y un segundo molde de ácido nucleico y en donde el primer molde de ácido nucleico tiene una secuencia que es diferente de la secuencia del segundo molde de ácido nucleico. Ejemplos de matrices que pueden ser útiles incluyen, sin limitación, una matriz BeadChip disponible de Illumina®, Inc. (San Diego, CA) o matrices tales como las descritas en las patentes de Estados Unidos N° 6.266.459; 6.355.431; 6.770.441; 6.859.570; o 7.622.294; o la publicación PCT N° WO 00/63437. Otros ejemplos de matrices comercialmente disponibles que pueden usarse incluyen, por ejemplo, una matriz Affymetrix® GeneChip® u otra matriz sintetizada de acuerdo con técnicas a veces denominadas tecnologías VLSIPS™ (Very Large Scale Immobilized Polymer Synthesis). Una matriz de detección también se puede usar de acuerdo con algunas realizaciones. Una matriz de detección ejemplar es una matriz CodeLink™ disponible en Amersham Biosciences. Otra matriz que es útil es una que se fabrica utilizando métodos de impresión de inyección de tinta tal como la Tecnología SurePrint™ disponible de Agilent Technologies.

20 Otras matrices útiles incluyen aquellas que se usan en aplicaciones de secuenciación de ácidos nucleicos. Por ejemplo, las matrices que tienen amplicones de fragmentos genómicos (a menudo denominados grupos) son particularmente útiles; tales como las descritas en Bentley et al., Nature 456: 53-59 (2008), documentos WO 04/018497; US 7.057.026; WO 91/06678; WO 07/123744; US 7.329.492; US 7.211.414; US 7.315.019; US 7.405.281, o US 2008/0108082.

25 Los grupos de ácidos nucleicos pueden crearse mediante métodos de amplificación en fase sólida. Por ejemplo, un ácido nucleico que tiene una o más secuencias de molde a detectar puede unirse a una superficie y amplificarse utilizando la amplificación de puente. Los métodos de amplificación de puentes útiles se describen, por ejemplo, en la patente de Estados Unidos N° 5.641.658; publicación de patente de Estados Unidos No. 2002/0055100; patente de Estados Unidos No. 7.115.400; publicación de patente de Estados Unidos No. 2004/0096853; publicación de patente de Estados Unidos No. 2004/0002090; publicación de patente de Estados Unidos No. 2007/0128624; y publicación de patente de Estados Unidos No. 2008/0009420. Otro método útil para amplificar ácidos nucleicos sobre una superficie es la amplificación en círculo rodante (RCA), por ejemplo, como se describe en Lizardi et al, Nat. Genet. 19: 225-232 (1998) y la publicación de solicitud de patente de EE.UU. No. 2007/ 0099208 A1. Otro tipo de matriz que es útil es una matriz de partículas producidas a partir de una técnica de amplificación por PCR en emulsión. Ejemplos se describen en Dressman et al, Proc. Natl. Acad. Sci. USA 100: 8817-8822 (2003), documentos WO 05/010145, US 2005/0130173 o US 2005/0064460. Aunque las matrices anteriores han sido descritas en el contexto de aplicaciones de secuenciación, se entenderá que las matrices pueden usarse en otras realizaciones incluyendo, por ejemplo, aquellas que usan una técnica de extensión de cebador no cíclica.

40 La detección puede llevarse a cabo a niveles de conjunto o de molécula única en una matriz. La detección a nivel de conjunto es la detección que se produce de manera que varias copias de una única secuencia de molde se detectan en cada sitio individual y las copias individuales en el sitio no se distinguen entre sí. Por lo tanto, la detección en conjunto proporciona una señal media de una secuencia de molde particular en el sitio. Por ejemplo, el sitio puede contener al menos 10, 100, 1000 o más copias de una secuencia de molde particular. Por supuesto, un sitio puede contener múltiples secuencias de molde diferentes cada una de las cuales está presente como un conjunto. Alternativamente, la detección a nivel de molécula única incluye la detección que tiene lugar de manera que las secuencias individuales de molde se resuelven individualmente en la matriz, cada una en un sitio diferente. De este modo, la detección de una sola molécula proporciona una señal de una molécula individual que se distingue de una o más señales que pueden surgir de una población de moléculas dentro de la cual está presente la molécula individual. Por supuesto, incluso en una matriz de molécula única, un sitio puede contener varias secuencias de molde diferentes (por ejemplo, dos o más regiones de secuencia de molde situadas a lo largo de una única molécula de ácido nucleico).

50 Una matriz de sitios puede aparecer como una cuadrícula de puntos o parches. Los sitios pueden estar situados en un patrón repetitivo o en un patrón no repetitivo irregular. Los patrones particularmente útiles son los patrones hexagonales, patrones rectilíneos, patrones de rejilla, patrones que tienen simetría reflexiva, patrones que tienen simetría rotacional, o similares. Los patrones asimétricos también pueden ser útiles. El tamaño de los sitios y/o la separación entre los sitios en una matriz puede variar para conseguir una densidad alta, densidad media o densidad más baja. Las matrices de alta densidad se caracterizan por tener sitios separados por menos de aproximadamente 15 µm. Las matrices de densidad media tienen sitios separados por aproximadamente 15 a 30 µm, mientras que las matrices de baja densidad tienen sitios separados por más de 30 µm. Una matriz útil en algunas realizaciones puede tener sitios que estén separados por menos de 100 µm, 50 µm, 10 µm, 5 µm, 1 µm o 0,5 µm. Una realización de los métodos expuestos en la presente memoria puede usarse para fotografiar una matriz a una resolución suficiente para distinguir los sitios en las densidades o intervalos de densidad anteriores. Sin embargo, la etapa de detección

usará típicamente un detector con una resolución espacial que será demasiado baja para resolver puntos a una distancia equivalente a la separación entre el primer producto de extensión de cebador y el segundo producto de extensión de cebador en cada uno de los sitios. En realizaciones particulares, los sitios de una matriz pueden tener cada uno un área que es mayor que aproximadamente  $100 \text{ nm}^2$ ,  $250 \text{ nm}^2$ ,  $500 \text{ nm}^2$ ,  $1 \text{ }\mu\text{m}^2$ ,  $2,5 \text{ }\mu\text{m}^2$ ,  $5 \text{ }\mu\text{m}^2$ ,  $10 \text{ }\mu\text{m}^2$ ,  $100 \text{ }\mu\text{m}^2$ , o  $500 \text{ }\mu\text{m}^2$ . Alternativamente o adicionalmente, los sitios de una matriz pueden tener cada uno un área que sea más pequeña que aproximadamente  $1 \text{ mm}^2$ ,  $500 \text{ }\mu\text{m}^2$ ,  $100 \text{ }\mu\text{m}^2$ ,  $25 \text{ }\mu\text{m}^2$ ,  $10 \text{ }\mu\text{m}^2$ ,  $5 \text{ }\mu\text{m}^2$ ,  $1 \text{ }\mu\text{m}^2$ ,  $500 \text{ nm}^2$ , o  $100 \text{ nm}^2$ . De hecho, el sitio puede tener un tamaño que esté en un intervalo entre un límite superior e inferior seleccionado de los ejemplificados anteriormente.

Los métodos expuestos en la presente memoria pueden usar matrices con sitios en cualquiera de una variedad de densidades incluyendo, por ejemplo, al menos aproximadamente 10 sitios/ $\text{cm}^2$ , 100 sitios/ $\text{cm}^2$ , 500 sitios/ $\text{cm}^2$ , 1.000 sitios/ $\text{cm}^2$ , 5.000 sitios/ $\text{cm}^2$ , 10.000 sitios/ $\text{cm}^2$ , 50.000 sitios/ $\text{cm}^2$ , 100.000 sitios/ $\text{cm}^2$ , 1.000.000 sitios/ $\text{cm}^2$ , 5.000.000 sitios/ $\text{cm}^2$ , o superior.

Un sistema de detección ortogonal, tal como un sistema utilizado para la síntesis secuencial ortogonal, puede utilizar diferentes marcadores para distinguir los diferentes nucleótidos que se añaden a cada cebador. En una realización, cada especie de nucleótido tendrá un marcador óptico único que producirá una señal única para distinguir esa especie de nucleótido. Un ejemplo es el enfoque de SBOS de 4 tintes descrito en el Ejemplo 1, a continuación, y mostrado en la Fig. 1A y Fig. 1B. En este ejemplo, se utiliza un primer conjunto de 4 tintes fluorescentes diferentes para distinguir entre sí a los 4 análogos de dNTP diferentes y se utiliza un segundo conjunto de 4 tintes fluorescentes diferentes para distinguir entre sí los 4 análogos de rNTP diferentes. Los dos conjuntos de colorantes son únicos de tal manera que los 8 colorantes producen 8 señales distinguibles, respectivamente.

En realizaciones en las que todos los nucleótidos están marcados de forma distinguible, tal como el enfoque de SBOS de 4 tintes, se puede poner en contacto un par de secuencias de molde con todos los nucleótidos y luego se puede realizar la detección después. En este caso la capacidad de distinguir todos los nucleótidos debido a marcadores ópticos únicos proporciona el beneficio de manipulaciones fluidicas relativamente sencillas, con lo que todos los nucleótidos pueden ser suministrados a las secuencias molde de modo que estén simultáneamente presentes. En una realización SBOS relativamente sencilla y preferida, todos los 8 nucleótidos se administran simultáneamente; sin embargo, uno o más subconjuntos se pueden entregar secuencialmente si se desea. La detección puede producirse durante o después del suministro de los nucleótidos. Este proceso fluido relativamente simple se acomoda mediante un dispositivo de detección relativamente complejo que tiene la capacidad de distinguir todas las señales. Por ejemplo, puede utilizarse un sistema de detección de fluorescencia capaz de distinguir 8 señales fluorescentes diferentes para un enfoque SBOS que utiliza 8 nucleótidos marcados fluorescentemente diferentes. Los expertos en la técnica sabrán o serán capaces de determinar un aparato de detección de fluorescencia apropiado para conseguir este tipo de diferenciación de señal. Por ejemplo, las propiedades de excitación y emisión de los marcadores fluorescentes pueden adaptarse adecuadamente con una combinación de longitudes de onda de excitación producidas y longitudes de onda de emisión detectadas por un fluorómetro. En Lakowicz, Principles of Fluorescence Spectroscopy, 3<sup>a</sup> Ed. Springer (2006); Haugland, Handbook of Fluorescent Probes and Research Products 9th Ed., Molecular Probes, Inc, (2002); y Shapiro, Practical Flow Cytometry, 4<sup>a</sup> Ed., John Wiley & Sons (2003) se proporcionan guías ejemplares para la óptica y marcadores útiles para la detección de fluorescencia de longitud de onda múltiple.

Los principios ejemplificados anteriormente para un sistema en el que todos los nucleótidos están marcados de forma distinguible, se pueden extender fácilmente a un formato de matriz. Se espera que una matriz que tenga un número y una variedad suficientes de diferentes secuencias de molde incorporará todos los nucleótidos marcados cuando se trate con sistemas de reacción de extensión de cebador. Más específicamente, en un enfoque SBOS basado en matrices, que tiene una amplia variedad de ácidos nucleicos a través de los sitios de la matriz y que tiene dos moldes diferentes por sitio, todas las combinaciones posibles del tinte de 2-tintes se espera que se produzcan en la matriz después de un ciclo de extensión de cebador en el que todos los 8 nucleótidos se suministraron a la matriz. Los sitios se pueden distinguir espacialmente utilizando dispositivos ópticos conocidos en la técnica, por ejemplo, los descritos en las patentes de EE.UU. Nos. 8.241.573; 7.329.860 y 8.039.817; y la publicación de solicitud de patente de EE.UU. Nos. 2009/ 0272914 A1 y 2012/0270305 A1. Dichos sistemas de detección pueden modificarse fácilmente para acomodar la detección fluorescente de 8 colores tal como se ha expuesto anteriormente. Un sistema de detección que se modifica de esta manera será capaz de multiplexar la detección ortogonal de manera que se distingan dos moldes diferentes (por ejemplo, mediante secuenciación) en múltiples sitios, teniendo cada uno una composición de secuencia diferente.

En algunas realizaciones, el número de señales diferentes que se distinguen en un método particular es menor que el número de especies de nucleótidos diferentes utilizadas en ese método. Por ejemplo, múltiples especies de nucleótidos diferentes pueden tener el mismo marcador y/o un subconjunto de las especies de nucleótidos puede no estar marcado. Un ejemplo de una configuración que usa el mismo marcador para múltiples especies de nucleótidos diferentes es el caso de un método de extensión de cebador ortogonal (o SBOS) en el que 4 desoxirribonucleótidos diferentes tienen un primer marcador en común y 4 ribonucleótidos diferentes tienen un segundo marcador en común. En esta configuración, los 4 desoxirribonucleótidos diferentes pueden distinguirse entre sí mediante ciclos secuenciales de liberación de uno de los desoxirribonucleótidos y detectando los desoxirribonucleótidos antes de administrar el desoxirribonucleótido subsiguiente. Siempre y cuando se puedan distinguir el primer marcador y el

segundo marcador en este ejemplo, los desoxirribonucleótidos y ribonucleótidos pueden administrarse en pares (1 cada uno de una única especie de desoxirribonucleótido y una sola especie de ribonucleótido), en 4 ciclos de administración y detección. Por lo tanto, los miembros de un primer conjunto de análogos de nucleótidos utilizados en una reacción de extensión de cebadores (por ejemplo, dNTPs) pueden incluir sólo un tipo de marcador óptico que se detecte y un segundo conjunto de análogos de nucleótidos, que sea ortogonal al primer conjunto (por ejemplo, rNTPs) puede incluir también sólo un tipo de marcador óptico que se detecta, en el que el marcador utilizado en el primer conjunto es ópticamente distinguible del marcador utilizado en el segundo conjunto.

La escala de grises permite el uso de múltiples especies de nucleótidos diferentes que tienen el mismo marcador. En este caso, se pueden distinguir diferentes especies de nucleótidos basándose en la intensidad de la señal del marcador detectada. Por ejemplo, cada especie de nucleótido se puede suministrar como una mezcla de proporción única de esa especie de nucleótido en forma marcada y no marcada. La variación en la relación de nucleótidos marcados: no marcados para cada especie dará como resultado una salida de señal en escala de grises única para cada mezcla. A modo de ejemplo más específico, un primer nucleótido puede marcarse completamente (sin mezclar el primer nucleótido marcado y no marcado), un segundo nucleótido puede marcarse al 75% (una mezcla de un segundo nucleótido marcado al 75% y un segundo nucleótido no marcado del 25%), un tercer nucleótido puede marcarse al 50% (una mezcla de un tercer nucleótido marcado al 50% y un tercer nucleótido no marcado al 50%) y un cuarto nucleótido puede marcarse al 25% (una mezcla de un cuarto nucleótido marcado con un 25% y un cuarto nucleótido no marcado con un 75%). Estas 4 especies de nucleótidos pueden distinguirse basándose en las diferencias resultantes en la intensidad de la señal, por lo que una población de cebadores (por ejemplo, en un sitio de matriz) producirá la señal completa debido a la incorporación del primer nucleótido; un 75% debido a la incorporación del segundo nucleótido, un 50% de señal debido a la incorporación del tercer nucleótido y un 25% de señal debido a la incorporación del cuarto nucleótido.

En realizaciones particulares, al menos una de las especies de nucleótidos puede estar completamente no marcada. Por lo tanto, en el caso en el que los marcadores ópticos estén presentes en los otros nucleótidos en un conjunto de nucleótidos, también puede haber un nucleótido "oscuro". La extensión de un cebador para incorporar un nucleótido oscuro, o, de otro modo, no marcado, se puede determinar por la inferencia basada en la ausencia de un marcador que se esperaría si los otros nucleótidos en el conjunto fueran incorporados por la reacción de extensión. Por lo tanto, en algunas realizaciones sólo un subconjunto de los nucleótidos usados en una reacción de extensión de cebador expuesta en este documento necesita tener un marcador.

El uso de especies de nucleótidos completamente no marcados se puede combinar con la escala de grises. Por ejemplo, tres de cuatro especies de nucleótidos diferentes en un conjunto pueden tener cantidades no distintas distinguibles de un marcador particular (por ejemplo, relaciones de nucleótidos marcados y no marcados en una mezcla) y la cuarta especie de nucleótido puede carecer de ese marcador. Alternativamente o adicionalmente, la escala de grises puede combinarse con el uso de varios marcadores ópticamente distinguibles. Por ejemplo, algunas especies de nucleótidos pueden representarse en una reacción de extensión como una mezcla de nucleótidos del mismo tipo pero que tienen marcadores diferentes. Tal configuración se ejemplifica en el Ejemplo I siguiente en el que una especie de nucleótido se proporciona como una mezcla de 50% de rCTP-F<sub>rojo</sub>/50% de rCTP-F<sub>azul</sub>. Otros ejemplos de escala de grises y marcadores mixtos que pueden modificarse para su uso en un método ortogonal de la presente descripción se exponen en el documento de EE.UU. 2013/0079232 A1.

Alternativamente o adicionalmente al uso de múltiples marcadores diferentes, escala de grises y/o especies no marcadas, una realización presentada en este documento puede usar un nucleótido que tenga un ligando, un enlazador escindible u otro resto que proporcione ganancia o pérdida de un marcador debido a un tratamiento definido. Los sistemas de reactivos de este tipo se ilustran en el Ejemplo I siguiente en el que algunas especies de nucleótidos tienen un ligando tal que pueden distinguirse de otros nucleótidos basándose en la ausencia inicial de una señal detectable seguida por la aparición de una señal después del tratamiento con un receptor marcado apropiadamente. El Ejemplo I también ilustra el uso de un nucleótido que se puede distinguir basándose en una señal detectable inicial que se pierde posteriormente o al menos se reduce debido al tratamiento con un reactivo que modifica el marcador (por ejemplo, por escisión química de un enlazador entre el marcador y el nucleótido). En este caso, las otras especies de nucleótidos del conjunto no son susceptibles a la modificación (por ejemplo, carecen del enlazador escindible) y se distinguen basándose en la persistencia de la generación de señales después del tratamiento.

Como se ha ejemplificado anteriormente y en el Ejemplo I, en algunas realizaciones, un marcador se puede unir a un análogo de nucleótidos a través de un enlazador escindible. En realizaciones particulares, se pueden usar enlazadores fotoescindibles en lugar del enlazador químicamente escindible ejemplificado anteriormente. En algunas realizaciones, el enlazador se selecciona entre enlazadores lábiles ácidos (incluyendo enlazadores de dialcoxibencilo, enlazadores de Sieber, enlazadores de indol, enlazadores de Sieber de t-butilo), enlazadores electrofílicamente escindibles, enlazadores nucleofílicamente escindibles, enlazadores fotoescindibles, enlazadores que se escinden en condiciones reductoras o condiciones oxidativas, enlazadores de captura de seguridad y enlazadores que se escinden mediante mecanismos de eliminación. En algunas de tales realizaciones, el enlazador se selecciona de un enlazador disulfuro (-S-S-), éster, nitrobenzeno, imina, péptido y polinucleótido enzimáticamente o químicamente escindible, tal como ADN.

En algunas realizaciones, los miembros de un primer conjunto de análogos de nucleótidos usados en una reacción de extensión de cebador (por ejemplo, dNTPs) incluirán sólo un tipo de marcador óptico que se detecta y un segundo conjunto de análogos de nucleótidos, que es ortogonal al primer conjunto (por ejemplo, rNTPs) también incluirá sólo un tipo de marcador óptico que se detecta, en donde el marcador usado en el primer conjunto es ópticamente distinguible del marcador usado en el segundo conjunto. En esta realización, el tipo de marcador óptico se puede unir a sustancialmente todos los análogos de nucleótidos de una primera especie en el primer conjunto, el tipo de marcador óptico se puede unir a un subconjunto de los análogos de nucleótidos de una segunda especie en el primer conjunto, sustancialmente todos los análogos de nucleótidos de una tercera especie en el primer conjunto pueden unirse a un ligando, y sustancialmente todos los análogos de nucleótidos de una cuarta especie en el primer conjunto no están unidos al tipo de marcador óptica o al ligando.

En otra realización, los miembros de un primer conjunto de análogos de nucleótidos usados en una reacción de extensión de cebadores (por ejemplo dNTPs) incluirán sólo dos tipos de marcadores ópticos que se detectan y un segundo conjunto de análogos de nucleótidos, que es ortogonal al primer conjunto (por ejemplo, rNTPs) también incluirá sólo dos tipos de marcadores ópticos que se detectan. En esta realización, un primer de los dos tipos de marcadores ópticos puede estar unido a sustancialmente todos los análogos de nucleótidos de una primera especie en el primer conjunto, un segundo de los dos tipos de marcadores ópticos puede estar unido a sustancialmente todos los análogos de nucleótidos de una segunda especie en el primer conjunto, el primer de los dos tipos de marcadores ópticos y el segundo de los dos tipos de marcadores ópticos se pueden unir a análogos de nucleótidos de una tercera especie en el primer conjunto y sustancialmente todos los análogos de nucleótidos de una cuarta especie en el primer conjunto no están unidos a uno de los dos tipos de marcadores ópticos o al segundo de los dos tipos de marcadores ópticos.

Se comprenderá a partir de los ejemplos anteriores que la reducción del número de marcadores diferentes en un sistema de detección ortogonal puede proporcionar la ventaja de reducir la complejidad del dispositivo de detección necesario para distinguir la adición de diferentes nucleótidos a un cebador unido a molde. Sin embargo, en muchas realizaciones esto se logra aumentando la complejidad de las etapas fluidicas de tal manera que el número de manipulaciones fluidicas utilizadas durante las etapas de detección se incrementa en comparación con las etapas fluidicas utilizadas cuando cada una de las especies de nucleótidos tiene un marcador único. Una ventaja general de los presentes métodos es que un experto en la técnica puede seleccionar una combinación apropiada de marcadores, pasos fluidos y dispositivos de detección para adaptarse a una aplicación o circunstancia particular.

La presente descripción proporciona mezclas de reacción (también denominadas en este documento sistemas reactivos) que incluyen diversas combinaciones de componentes. En varios casos, se describen componentes de reacción y varias combinaciones de los componentes en el contexto de métodos ejemplares. Se entenderá que las mezclas de reacción y sus componentes no necesitan limitarse a su uso en los procedimientos ejemplificados en la presente memoria. También se contemplan otros usos. Por consiguiente, los componentes se pueden ensamblar en una variedad de combinaciones útiles, por ejemplo, para crear kits. Los kits pueden ser útiles para el almacenamiento, transporte o transacción comercial de los componentes expuestos en este documento. Opcionalmente, los kits pueden incluir instrucciones para llevar a cabo uno o más de los métodos expuestos en este documento.

Los siguientes ejemplos pretenden ilustrar, pero no limitar, la presente invención.

#### **Ejemplo 1**

Secuenciación por síntesis ortogonal de dos cebadores

Este ejemplo describe una nueva plataforma de secuenciación que permite duplicar la salida de la secuenciación cuando se compara con la secuenciación por síntesis tradicional (SBS). Esta plataforma capitaliza en SBS porque la información de la secuenciación se deriva de un alargamiento por etapas del cebador de secuenciación (Figura 1A). Sin embargo, en la nueva plataforma, el evento de elongación en el primer sitio se produce en paralelo a un segundo evento de elongación de secuenciación, que ocurre en el segundo sitio aguas abajo del primer sitio (Figura 1B).

La ortogonalidad entre los sitios 1 y 2 de SBS se proporciona mediante el uso de dos diferentes polimerasas y combinaciones de sustrato. Como se ejemplifica a continuación, el sistema puede usar una ADN polimerasa de SBS y nucleótidos totalmente funcionales (FFN), tales como los disponibles de Illumina, Inc. (San Diego, CA), en combinación con una ARN polimerasa (p. ej. pol. T7 ARN) y rNTPs marcados correspondientes (Figura 1C). Aunque la discriminación en la especificidad de sustrato entre las ARN y las ADN polimerasas no es absoluta, se ha demostrado que la diferencia en la preferencia de sustrato entre rNTP y dNTP es tan alta como 10000 veces en ciertas condiciones (Joyce, Proc. Nat'l. Acad. Sci. USA, 94: 1619 - 1622 (1997) y Gao et al, Proc. Nat'l. Acad. Sci. USA, 94: 407 - 411 (1997)).

Alternativamente, se puede conseguir un sistema ortogonal explotando análogos de ADN/ARN sintéticos (XNAs) y una polimerasa modificada correspondiente. Un sistema de este tipo es posible utilizando la síntesis molde de ADN de HNAs largos (ácidos nucleicos de 1,5 anhidrohexitol) con una ADN polimerasa evolucionada. Las condiciones ejemplares y los reactivos que pueden usarse para la síntesis de molde de ADN de HNA largos se describen en



Pinheiro et al, Science, 336 (6079): 341 - 344 (2012) y Cozens et al., Proc. Nat'l. Acad. Sci. USA, 109 (21): 8067 - 8072 (2012).

#### Discriminación de la señal

5 Las siguientes configuraciones ejemplares pueden usarse para la discriminación de señales en una plataforma de síntesis por secuenciación ortogonal (SBOS).

(1) Enfoque basado en la química SBOS de 4-tintes. En este enfoque, se usa un conjunto de cuatro desoxirribonucleótidos terminadores reversibles (rtdNTPs) en los que cada una de las especies de rtdNTP tiene un fluoróforo que es ópticamente distinguible de los fluoróforos usados para las otras tres especies de rtdNTP. Esto es similar a la combinación de nucleótidos usada en plataformas SBS de 4 tintes comercialmente disponibles de Illumina, Inc. (San Diego, CA). También se usa un conjunto de cuatro ribonucleótidos terminadores reversibles (rtrNTPs) y cada una de las especies rtrNTP tiene un fluoróforo que es ópticamente distinguible de los fluoróforos usados para las otras tres especies de rtrNTP. Además, los fluoróforos usados para los rtdNTPs son ópticamente distinguibles de los fluoróforos para el rtrNTPs. Como tal, hay 8 fluoróforos diferentes en uso a través de los dos conjuntos de nucleótidos. La Tabla 1 muestra un ejemplo de conjunto donde el subíndice se refiere a la longitud de onda de emisión para cada fluoróforo (F).

Tabla 1: Nucleótidos para la configuración SBS de 4 tintes

Conjunto rtdNTP	Conjunto rtrNTP
rtdTTP-F <sub>lejos de rojo</sub>	rtrUTP-F <sub>verde</sub>
rtdCTP-F <sub>cerca de rojo</sub>	rtrCTP-F <sub>azul</sub>
rtdGTP-F <sub>naranja</sub>	rtrGTP-F <sub>indigo</sub>
rtdGTP-F <sub>amarillo</sub>	rtrGTP-F <sub>violeta</sub>

En el enfoque de 4 tintes, los componentes ópticos utilizados para distinguir los 8 diferentes fluoróforos pueden ser relativamente complejos, por ejemplo, utilizando hasta ocho canales de emisión diferentes y/o hasta ocho líneas de excitación diferentes.

(2) Enfoque basado en la química de SBOS de 1-tinte. Con el fin de reducir la complejidad del dispositivo óptico descrito para el enfoque de 4 tintes, puede utilizarse otro enfoque que utiliza menos fluoróforos que las especies de nucleótidos. Un enfoque con un colorante puede usar un conjunto de rtdNTP que tienen un único fluoróforo de un primer tipo (por ejemplo, un fluoróforo que emite azul). Las diferentes especies de rtdNTPs se pueden distinguir entre sí debido a la presencia del fluoróforo en una primera especie en el conjunto de rtdNTPs (por ejemplo, fluoróforo que emite de azul de rtdTTP), la presencia de un ligando de unión en una segunda especie en el conjunto de RtdNTPs (por ejemplo rtdCTP-Biotina), la ausencia del fluoróforo y el ligando de unión en una tercera especie en el conjunto de rtdNTPs (por ejemplo, rtdGTP no marcado) y la unión del fluoróforo a través de un enlazador escindible a una cuarta especie en el conjunto de rtdNTPs (por ejemplo, rtdATP-enlazador de disulfuro-fluoróforo que emite en azul). Ejemplos de combinaciones de especies de nucleótidos marcadas con fluoróforo, marcadas con ligandos y no marcadas que pueden usarse para crear un conjunto de nucleótidos para la detección de 1-tinte se exponen con más detalle en el documento de EE.UU. 2013/0079232 A1.

El enfoque con 1 tinte puede utilizar además un conjunto de rtrNTPs que tienen un segundo tipo de fluoróforo. El segundo tipo de fluoróforo (por ejemplo, un fluoróforo que emite en rojo) es ópticamente distinguible del fluoróforo usado para miembros del conjunto de rtdNTP. Además, las especies individuales en el conjunto de rtrNTPs pueden distinguirse entre sí por una combinación de especies marcadas y no marcadas similar a la combinación ejemplificada anteriormente para el conjunto rtdNTP, con la siguiente modificación. Con el fin de evitar la introducción de un nuevo sistema Biotina-Estreptavidina, uno de los rtrNTP puede ser una mezcla de los dos fluoróforos en uso. Un conjunto ejemplar de nucleótidos se muestra en la Tabla 2 (un conjunto similar se muestra en la Figura 2A).

Tabla 2: Nucleótidos para la configuración SBS de 1 tinte

Conjunto rtdNTP	Conjunto rtrNTP
rtdTTP-F <sub>azul</sub>	rtrUTP-F <sub>rojo</sub>
rtdCTP-Biotina	50% rtrCTP-F <sub>rojo</sub> / 50% rtrCTP-F <sub>azul</sub>
rtdGTP (no marcado)	rtrGTP (no marcado)
rtdATP-S-S-F <sub>azul</sub>	rtrATP-S-S-F <sub>rojo</sub>

Para un enfoque SBOS usando la configuración de la Tabla 2, los componentes ópticos sólo necesitan incluir 2 canales de emisión diferentes y sólo 1 ó 2 líneas de excitación diferentes. En este enfoque, como se muestra en la Fig. 2A, se obtienen un total de 4 imágenes por ciclo (2 por color). Dos imágenes se registran antes y dos después del tratamiento con estreptavidina-F<sub>azul</sub>. En la Fig. 2A y Fig. 2B, "oscuro" indica la ausencia del marcador del fluoróforo, "NR550C4" es el fluoróforo que emite en azul, "THP" es Tris(3-hidroxipropil)fosfina que escinde el enlace disulfuro ("SS"), "Strep" es estreptavidina, "d-FFN" es rtdNTP, "r-FFN" es rtrNTP y "rojo" es el fluoróforo emisor rojo.

Aunque la mezcla de especies marcadas con rtrCTP se ejemplifica anteriormente como una relación molar 1:1 (es decir, 50% F<sub>rojo</sub> y 50% de F<sub>azul</sub>), se entenderá que se pueden usar otras relaciones en su lugar. Por ejemplo, puede ser deseable ajustar la relación para acomodar diferentes propiedades ópticas para los dos fluoróforos (por ejemplo, uno de los fluoróforos puede estar presente a un ligero exceso molar para acomodar una intensidad de emisión relativamente más baja que el otro fluoróforo a la longitud de onda de detección usada). La proporción también puede desviarse de 1:1 para acomodar diferentes propiedades bioquímicas de los nucleótidos marcados (por ejemplo una afinidad diferente por la polimerasa).

(3) Química basada en SBOS de 2-tintes. Este enfoque proporciona otra combinación de nucleótidos con menos fluoróforos que las especies de nucleótidos. Se puede usar un conjunto de rtrNTP que tienen fluoróforos con dos emisiones diferentes (por ejemplo, emisión de rojo lejano y azul). De forma similar, se puede usar un conjunto de rtdNTP con fluoróforos con dos emisiones diferentes. Las cuatro emisiones diferentes son ópticamente distinguibles entre los dos conjuntos de rNTP. Un conjunto ejemplar de nucleótidos se muestra en la Tabla 3.

Tabla 3: Nucleótidos para la configuración SBS de 2 tintes

Conjunto rtdNTP	Conjunto rtrNTP
50% rtdTTP-F <sub>verde</sub> / 50% rtdTTP-F <sub>rojo</sub>	rtrUTP-F <sub>lejos de rojo</sub>
rtdGTP-F <sub>rojo</sub>	50% rtrCTP-F <sub>lejos de rojo</sub> / 50% rtrCTP-F <sub>azul</sub>
rtdGTP (no marcado)	rtrGTP (no marcado)
rtdATP-F <sub>verde</sub>	rtrATP-F <sub>azul</sub>

El enfoque de 2 tintes, ejemplificado anteriormente, proporciona una ventaja sobre el enfoque de 1 tinte previamente ejemplificado en que las etapas de procesamiento químico, tales como la unión a estreptavidina, y la escisión de THP, no se usan durante la fase de detección del enfoque de 2 tintes. Esto se demuestra por el protocolo del ciclo de extensión mostrado para el enfoque de 2 tintes en la Fig. 3. Sin embargo, en el caso de este enfoque particular de 2 tintes, se utiliza un dispositivo óptico capaz de distinguir cuatro colores diferentes. Como se ilustra en esta comparación, pueden utilizarse diferentes aproximaciones para ajustar la complejidad óptica o la complejidad fluidica para adaptarse a una plataforma de secuenciación particular. Por ejemplo, añadiendo etapas fluidicas para modificar nucleótidos durante un ciclo de detección de SBS se pueden usar menos componentes ópticos y/o más simples. Por el contrario, en los casos en que se dispone de componentes ópticos más complejos, se pueden utilizar menos manipulaciones fluidas y/o más sencillas.

#### Preparación de la muestra

Se puede preparar una muestra de ADN genómico u otra muestra de ácido nucleico para SBS de 2 – cebadores como se ejemplifica en la Fig. 4. El método permite la creación de un molde de ADN que tiene dos sitios de cebado diferentes. El método puede llevarse a cabo bajo condiciones que proporcionan fragmentos en un intervalo de tamaños deseado. Como se muestra en la Fig. 4, se puede usar una transposasa para marcar una muestra de ADN genómico con dos secuencias de marcador diferentes. Las condiciones pueden seleccionarse para producir fragmentos que tienen un promedio de aproximadamente 300 nucleótidos de longitud y que tienen un primera marcador en un extremo y un segundo marcador en el otro extremo. Los marcadores formarán salientes de una sola

5 hebra a la que los adaptadores pueden hibridar. La unión de los adaptadores hibridados como se muestra en la Fig. 4 producirá concatámeros de 2 fragmentos que tienen el siguiente orden de regiones de secuencia: una secuencia p5, un sitio de cebado Read 1, un fragmento 1 de ADN, un promotor de ARN polimerasa, un sitio de cebado Read 1', un fragmento 2 de ADN, índice y secuencia p7. Las secuencias p5 y p7 al final de los concatámeros permiten la  
10 amplificación de captura y puente, por ejemplo, utilizando células de flujo de secuenciación y kits disponibles de Illumina, Inc. (San Diego, CA). Los sitios de cebado Read 1 son complementarios a un cebador de ADN que se extiende por ADN polimerasa y el sitio de cebado Read 1' es complementario a un cebador de ARN que se extiende por ARN polimerasa, respectivamente, en la reacción ortogonal de SBS. El promotor de ARN polimerasa está situado aguas abajo del sitio de cebado de ARN polimerasa para activar la actividad de ARN polimerasa. El  
15 fragmento 1 de ADN está aguas abajo del sitio de cebado de ADN polimerasa y como tal está posicionado para la detección por la extensión de ADN polimerasa en la reacción de SBS y el fragmento 2 de ADN está aguas abajo del sitio de cebado de la ARN polimerasa y como tal está posicionado para la detección por la extensión de la ARN polimerasa en la reacción SBS. El índice está disponible opcionalmente para el seguimiento de muestras.

15 En general, se espera que la plataforma proporcionada por este ejemplo aumente la salida de secuenciación duplicando la información de secuenciación por ciclo; es decir, un ciclo de SBS 150 de 2 iniciadores sería equivalente a un ciclo de 2x150 de pares de SBS comercialmente disponible (Illumina, Inc., San Diego CA), pero con ahorros adicionales en tiempo de ejecución y uso de reactivos. Si se implementa en un formato de extremo pareado, un ciclo de 2x75 2 extremo de pares de SBS de 2-cebadores equivaldría a un ciclo de 2x150 pareados tradicional.

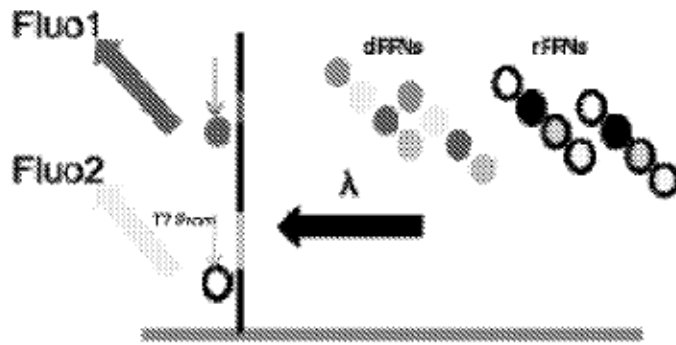
20 La expresión "que comprende" está destinada en este documento a ser abierta, incluyendo no sólo los elementos citados, sino que además abarca cualquier elemento adicional.

Aunque la invención se ha descrito con referencia a los ejemplos proporcionados anteriormente, debe entenderse que pueden realizarse diversas modificaciones sin apartarse de la invención. Por consiguiente, la invención está limitada únicamente por las reivindicaciones.

## REIVINDICACIONES

1. Un método para secuenciar moldes de ácido nucleico, que comprende:
  - (a) proporcionar un conjunto de sitios, en el que cada sitio comprende un primer molde de ácido nucleico y un segundo molde de ácido nucleico,
  - 5 en el que el primer molde de ácido nucleico tiene una secuencia que es diferente de la secuencia del segundo molde de ácido nucleico;
  - (b) extender un primer cebador unido al primer molde utilizando una primera especie de polimerasa y un primer conjunto de análogos de nucleótidos, produciendo de este modo un primer producto de extensión del cebador que comprende un primer análogo de nucleótido en cada uno de los sitios;
  - 10 (c) extender un segundo cebador unido al segundo molde usando una segunda especie de polimerasa y un segundo conjunto de análogos de nucleótidos, produciendo de este modo un segundo producto de extensión de cebador que comprende un segundo análogo de nucleótido en cada uno de los sitios,
  - en el que la primera especie de polimerasa es diferente de la segunda especie de polimerasa y en el que el primer conjunto de análogos de nucleótidos es diferente del segundo conjunto de análogos de nucleótidos;
  - 15 (d) detectar el primer producto de extensión del cebador y el segundo producto de extensión del cebador en cada uno de los sitios; y
  - (e) repetir las etapas (b) a (d), determinando así las diferentes secuencias del primer molde y el segundo molde en cada uno de los sitios.
2. El método de la reivindicación 1, en el que la primera polimerasa es selectiva para el primer conjunto de análogos de nucleótidos comparado con el segundo conjunto de análogos de nucleótidos y en el que la segunda polimerasa es selectiva para el segundo conjunto de análogos de nucleótidos comparado con el primer conjunto de análogos de nucleótidos.
3. El método de la reivindicación 1, en el que la primera polimerasa es selectiva para el primer cebador en comparación con el segundo cebador y en el que la segunda polimerasa es selectiva para el segundo cebador en comparación con el primer cebador.
4. El método de la reivindicación 1, en el que la detección usa un detector que tiene una resolución espacial que es demasiado baja para resolver puntos a una distancia equivalente a la separación entre el primer producto de extensión de cebador y el segundo producto de extensión de cebador en cada uno de los sitios.
5. El método de la reivindicación 4, en el que el detector es un detector óptico, y opcionalmente en el que los análogos de nucleótidos comprenden marcadores ópticos.
6. El método de la reivindicación 5, en el que los marcadores ópticos del primer conjunto de análogos de nucleótidos son diferentes de los marcadores ópticos del segundo conjunto de análogos de nucleótidos.
7. El método de la reivindicación 5, en el que un subconjunto de los análogos de nucleótidos en el primer conjunto de análogos de nucleótidos comprende marcadores ópticos y opcionalmente en el que un subconjunto de los análogos de nucleótidos en el segundo conjunto de análogos de nucleótidos comprende marcadores ópticos.
8. El método de la reivindicación 7, en el que el primer conjunto de análogos de nucleótidos comprende marcadores ópticos que son diferentes de los marcadores ópticos del segundo conjunto de análogos de nucleótidos.
9. El método de la reivindicación 4, en el que un píxel del detector adquiere señales tanto del primer producto de extensión del cebador como del segundo producto de extensión del cebador.
10. El método de la reivindicación 1, en el que la primera polimerasa es una ARN polimerasa y la segunda polimerasa es una ADN polimerasa, y opcionalmente en el que el primer cebador es un cebador de ARN y el segundo cebador es un cebador de ADN.
11. El método de la reivindicación 1, en el que las etapas (b) y (c) se llevan a cabo simultáneamente.
12. El método de la reivindicación 1, en el que las etapas (b) y (c) se llevan a cabo secuencialmente.
- 45 13. El método de la reivindicación 1, en el que los sitios tienen un área que no es mayor de  $100 \mu\text{m}^2$ .
14. El método de la reivindicación 1, en el que los sitios comprenden copias múltiples del primer molde de ácido nucleico y del segundo molde de ácido nucleico.

15. El método de la reivindicación 1, en el que la primera polimerasa es selectiva para el primer molde comparado con el segundo molde y en el que la segunda polimerasa es selectiva para el segundo molde en comparación con el primer molde.



**SBS de 2 cebadores**

ADNpol + ADNcebador + dNTPs marcados + ARNpol  
(es decir, T7) + ARNcebador + rNTPs marcados

Figura 1A

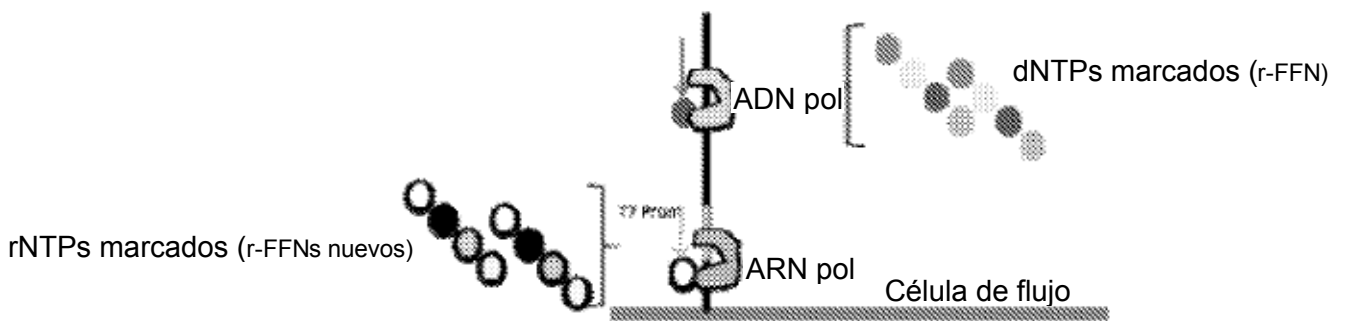
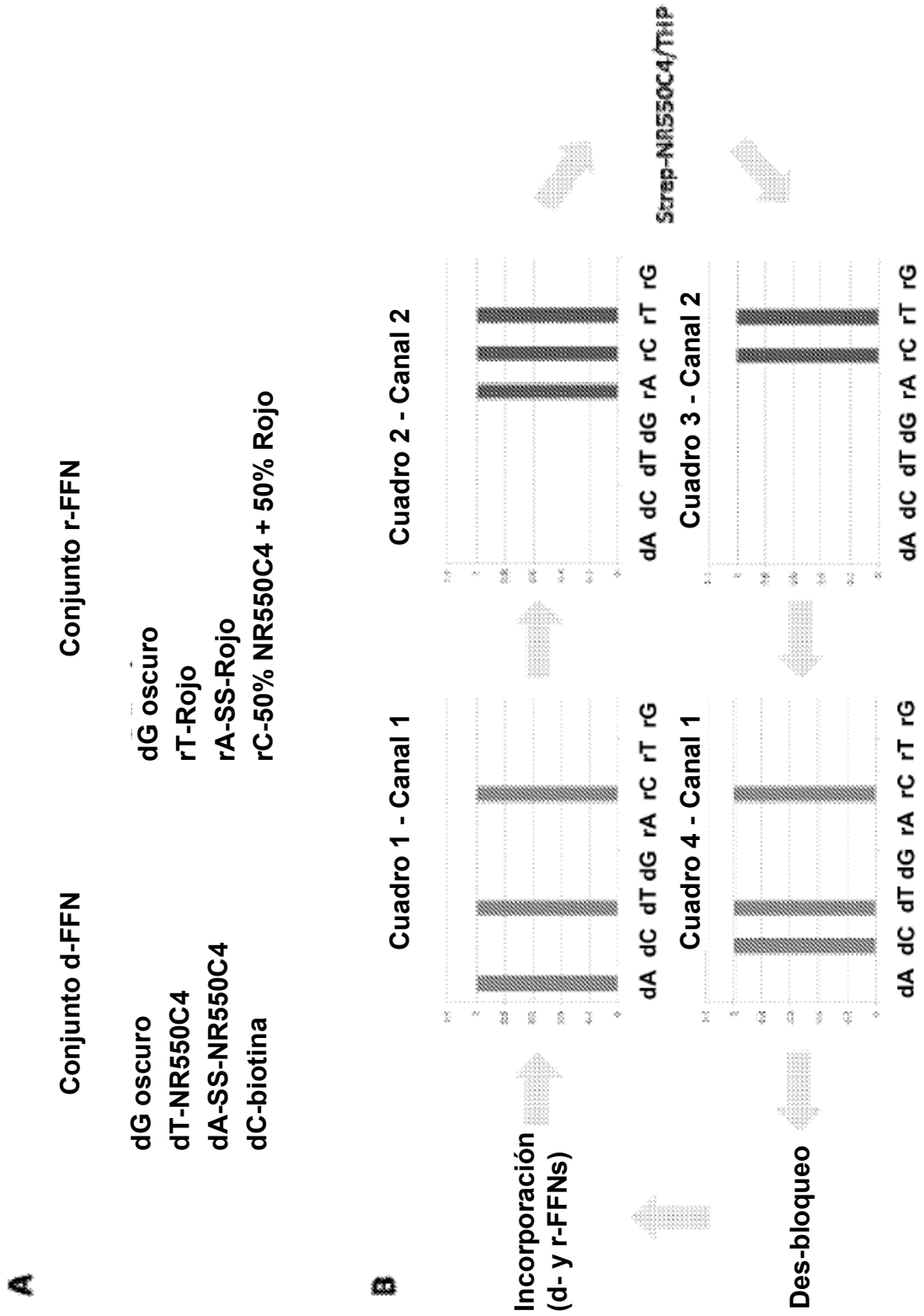
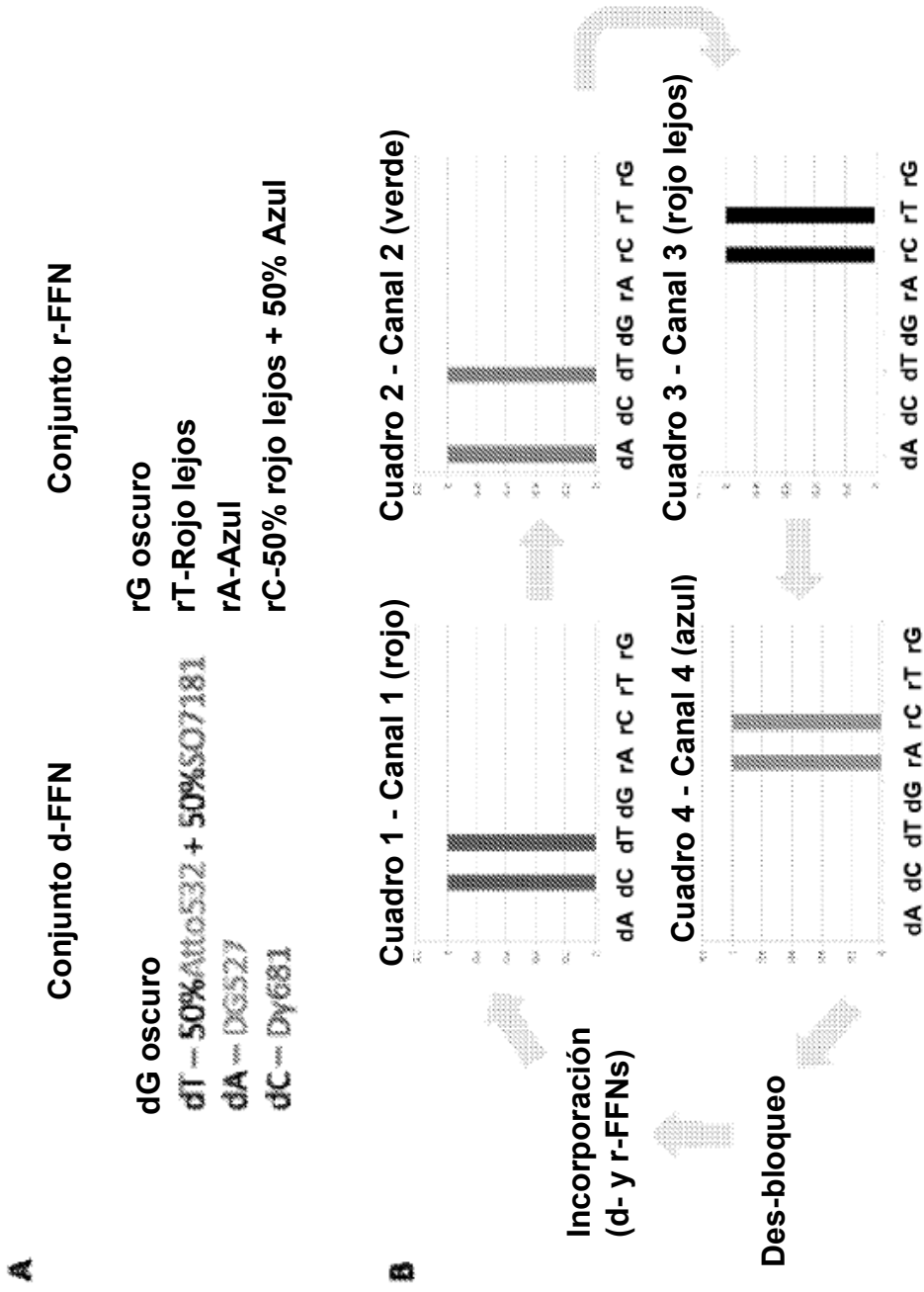


Figura 1B



Figuras 2A y 2B



Figuras 3A y 3B



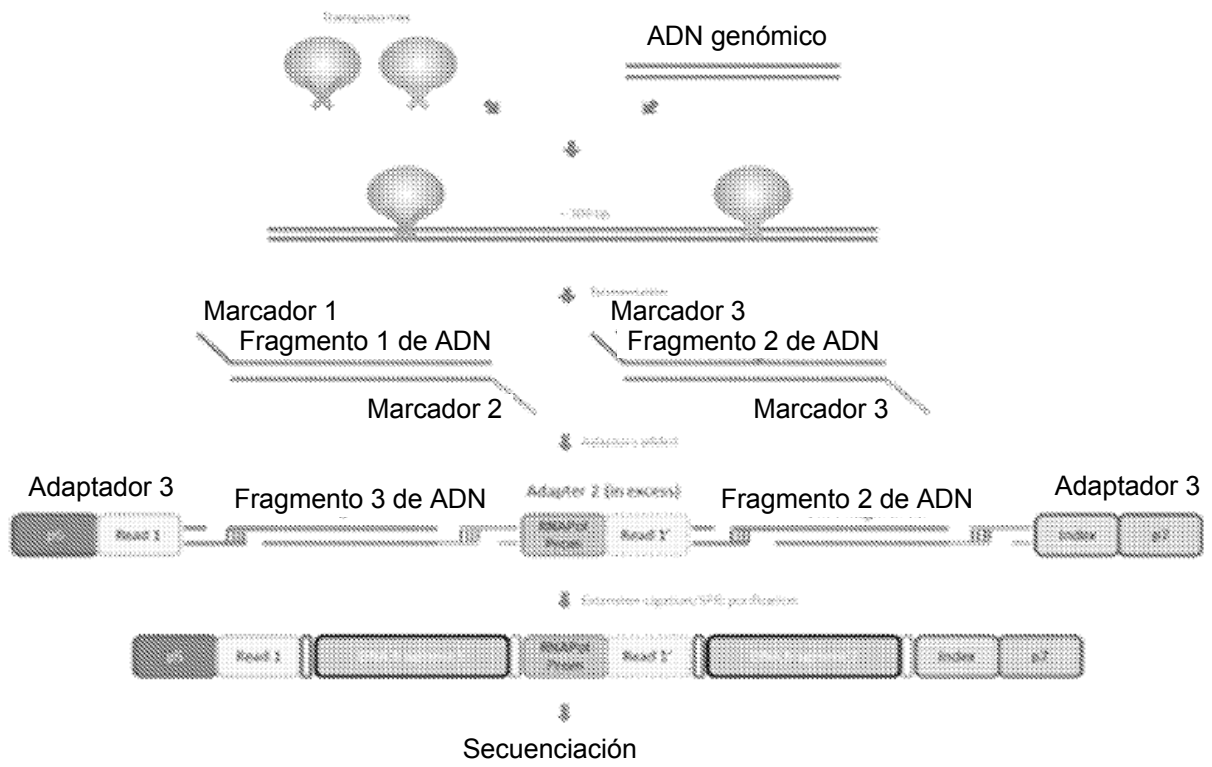
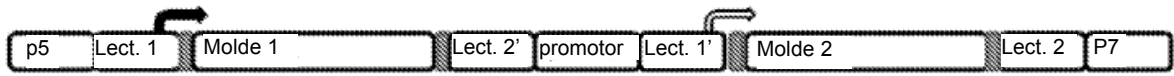


Figura 4

Lectura 1 PE ortogonal



Lectura 2 PE ortogonal

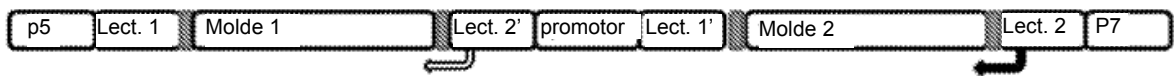


Figura 5