

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 639 938**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **21.09.2012 PCT/US2012/056556**

87 Fecha y número de publicación internacional: **28.03.2013 WO13044018**

96 Fecha de presentación y número de la solicitud europea: **21.09.2012 E 12766838 (2)**

97 Fecha y número de publicación de la concesión europea: **12.07.2017 EP 2768972**

54 Título: **Métodos y composiciones para la secuenciación de ácidos nucleicos**

30 Prioridad:

23.09.2011 US 201161538294 P
03.04.2012 US 201261619878 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
30.10.2017

73 Titular/es:

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

72 Inventor/es:

KAIN, ROBERT C.;
LIU, XIAOHAI;
FENG, WENYI;
HIRSCHBEIN, BERNARD;
ELTOUKHY, HELMY A.;
WU, XIAOLIN;
SMITH, GEOFFREY PAUL;
BOUTELL, JONATHAN MARK;
JOSEPH, THOMAS;
SMITH, RANDALL;
SHEN, MIN-JUI RICHARD;
TREGIDGO, CAROLYN y
KLAUSING, KAY

74 Agente/Representante:

ELZABURU, S.L.P

Observaciones :

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

ES 2 639 938 T3

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

DESCRIPCIÓN

Métodos y composiciones para la secuenciación de ácidos nucleicos

Antecedentes

5 La detección de analitos, tales como las secuencias de ácidos nucleicos que están presentes en una muestra biológica, se ha utilizado como un método para identificar y clasificar microorganismos, diagnosticar enfermedades infecciosas, detectar y caracterizar anomalías genéticas, identificar cambios genéticos asociados al cáncer, estudiar la susceptibilidad genética frente a enfermedades, y medir la respuesta frente a diversos tipos de tratamiento. Una técnica común para detectar analitos, tales como las secuencias de ácido nucleico en una muestra biológica, es la secuenciación de ácidos nucleicos.

10 La metodología de la secuenciación de ácidos nucleicos ha evolucionado significativamente desde los métodos de degradación química utilizados por Maxam y Gilbert y los métodos de alargamiento de cadenas utilizados por Sanger. Hoy en día se utilizan varias metodologías de secuenciación que permiten el procesamiento paralelo de miles de ácidos nucleicos, todos en una única secuencia. La instrumentación que realiza tales métodos es típicamente grande y costosa ya que los métodos actuales suelen basarse en grandes cantidades de reactivos
15 costosos y conjuntos múltiples de filtros ópticos para registrar la incorporación de los ácidos nucleicos en las reacciones de secuenciación.

Se ha puesto de manifiesto que la necesidad de tecnologías de secuenciación de ADN de alto rendimiento, más pequeñas y menos costosas será beneficiosa para obtener los frutos de la secuenciación del genoma. La asistencia sanitaria personalizada se está moviendo hacia la vanguardia y se beneficiará de estas tecnologías; la
20 secuenciación del genoma de un individuo para identificar posibles mutaciones y anomalías será crucial para identificar si una persona tiene una enfermedad en particular, seguido del desarrollo de terapias posteriores adaptadas a esa persona. Para acomodar un esfuerzo tan potente, la secuenciación debe avanzar y ser accesible a tecnologías de alto rendimiento no sólo por sus capacidades de alto rendimiento, sino también en términos de facilidad de uso, eficiencia de tiempo y costo, y accesibilidad de los clínicos a los instrumentos y reactivos.

25 El documento WO 2009/054922 describe un método para determinar la identidad de cada una de una serie de residuos de nucleótidos consecutivos en un ácido nucleico que comprende:

a) poner en contacto una pluralidad de ácidos nucleicos con (i) un análogo de didesoxinucleótido (ddNTP), que comprende un fluoróforo separable unido a una base, en el que cada base tiene un fluoróforo diferente unido y en el que el fluoróforo unido a cada tipo de base difiere en su espectros de excitación o emisión del
30 fluoróforo unido a los otros tipos de bases, (ii) un análogo de desoxinucleótido trifosfato (dNTP), una polimerasa de ácido nucleico y al menos dos cebadores,

b) identificar el fluoróforo del análogo de ddNTP que ha formado el enlace fosfodiéster, identificando así la identidad del nucleótido consecutivo,

c) separar el fluoróforo del análogo de ddNTP que ha formado el enlace fosfodiéster;

35 d) repetir iterativamente las etapas a) hasta c);

e) repetir los pasos a) y b).

El documento CN 101 570 784 describe un método de secuenciación de ligación de ADN basado en la codificación de la combinación de señales que utiliza la codificación de combinación de señales. El documento CN 101 565
40 746 describe un método de secuenciación de ADN para códigos combinados de señales con una comprobación de paridad.

El documento WO2009/097368A describe métodos y composiciones para adquirir la información de la secuencia nucleotídica de secuencias diana usando menos marcadores para distinguir entre nucleótidos, detectando nucleótidos en múltiples posiciones de detección en una secuencia diana.

Sumario

45 Las reacciones de secuenciación basadas en fluorescencia existentes distinguen entre la incorporación de diferentes nucleótidos en una hebra de ácido nucleico creciente uniendo un resto fluorescente a cada uno de los cuatro nucleótidos A, T, C y G. Típicamente, cada uno de los restos fluorescentes excita y emite a diferentes longitudes de onda y, por lo tanto, la secuencia diana es determinada. Por el contrario, la presente descripción proporciona la determinación de una secuencia, por ejemplo, una secuencia de ácido nucleico, utilizando un conjunto colorante
50 mínimo, fuentes de luz de excitación mínimas y filtros de emisión óptica mínima, al tiempo que permite la diferenciación de la incorporación de los cuatro nucleótidos en una reacción de secuenciación. La presente descripción proporciona métodos y composiciones susceptibles a cualquier sistema fluorescente en el que se desea detectar más de un analito. Sin embargo, se encuentran ventajas particulares cuando se aplican los métodos de la presente invención a metodologías de secuenciación tales como las secuencias por metodologías de síntesis.

Los instrumentos y sistemas para detectar la secuenciación de fluorescencia de cuatro colores son grandes y costosos de ejecutar, no sólo debido al coste del instrumento, sino también a los reactivos, y por lo tanto no son muy atractivos para ubicaciones más pequeñas y más limitadas de capital. Los métodos y composiciones que reducirían los costes y/o el tamaño asociados con la detección de fluorescencia de cuatro colores, por ejemplo para la secuenciación de genomas, proporcionarían a los investigadores herramientas más eficientes en términos de eficiencia de tiempo, menor uso de reactivos, instrumentación menos costosa y similares para uso en sus empresas de investigación.

Las realizaciones de la presente descripción proporcionan esas opciones proporcionando a los investigadores métodos y composiciones para la determinación de una secuencia polimérica, por ejemplo una secuencia de ácido nucleico, que comprende usar un conjunto de colorantes mínimo, fuentes de luz mínimas y filtros de excitación/emisión mínimos, mientras se permite la diferenciación de los tipos de monómeros (por ejemplo, nucleótidos diferentes) incorporados en una reacción de secuenciación.

Las realizaciones descritas en este documento proporcionan la determinación de la secuencia de un ácido nucleico en base a la temporización de eventos y la memorialización de dichos eventos en un "espacio de tiempo". La presente descripción proporciona realizaciones para el uso de un colorante, o una pluralidad de colorantes del mismo o similar espectro de emisión/emisión, o dos o más colorantes de diferentes espectros de fluorescencia, para determinar la presencia de analitos, por ejemplo nucleótidos, en una muestra, utilizando eventos de imágenes basadas en el espacio de tiempo. Como se describe en este documento, las reacciones de secuenciación de espacio de tiempo utilizan una o más químicas y eventos o etapas de formación de imágenes para diferenciar entre una pluralidad de analitos, por ejemplo cuatro nucleótidos, que se incorporan en una hebra de ácido nucleico en crecimiento durante una reacción de secuenciación.

En algunas realizaciones, se pueden detectar menos de cuatro colores diferentes en una mezcla que tiene cuatro nucleótidos diferentes, mientras que todavía resulta en la determinación de los cuatro nucleótidos diferentes, por ejemplo en una reacción de secuenciación. Como primer ejemplo, se puede detectar un par de tipos de nucleótidos a la misma longitud de onda, pero se distinguen basándose en una diferencia de intensidad para un miembro del par comparado con el otro, o basándose en un cambio en un miembro del par (por ejemplo, mediante modificación química, modificación fotoquímica o modificación física) que hace que aparezca o desaparezca la señal aparente en comparación con la señal detectada para el otro miembro del par. Como segundo ejemplo, se pueden detectar tres de cuatro tipos de nucleótidos diferentes bajo condiciones particulares mientras que un cuarto tipo de nucleótido carece de un marcador que sea detectable en esas condiciones, o se detecte mínimamente bajo esas condiciones (por ejemplo, detección mínima debido a fluorescencia de fondo, etc.). La incorporación de los tres primeros tipos de nucleótidos en un ácido nucleico se puede determinar basándose en la presencia de sus respectivas señales y la incorporación del cuarto tipo de nucleótido en el ácido nucleico se puede determinar basándose en la ausencia o detección mínima de cualquier señal. Como un tercer ejemplo, un tipo de nucleótido puede incluir marcador o marcadores que se detectan en dos canales diferentes, mientras que otros tipos de nucleótidos se detectan en no más de uno de los canales.

En un aspecto de la presente invención, se proporciona un método de secuenciación basada en fluorescencia por síntesis (SBS), según se define en la reivindicación 1, que utiliza un primer tipo de nucleótido que se detecta en un primer canal (por ejemplo, dATP que tiene un marcador que se detecta en el primer canal cuando es excitado por una primera longitud de onda de excitación), un segundo tipo de nucleótido que se detecta en un segundo canal (por ejemplo dCTP que tiene un marcador que se detecta en el segundo canal cuando se excita por una segunda longitud de onda de excitación), un tercer tipo de nucleótido que se detecta tanto en el primer como en el segundo canal (por ejemplo dTTP que tiene al menos un marcador que se detecta en ambos canales cuando se excita por la primera y/o segunda longitud de onda de excitación) y un cuarto tipo de nucleótido que carece de un marcador que no es, o es mínimamente detectado en cualquier canal (por ejemplo, dGTP que no tiene marcador).

De esta manera, se puede tratar una serie de características del ácido nucleico con los cuatro tipos de nucleótidos de tal manera que se produzca un evento de extensión sustancialmente en todas las características antes de un evento de detección y las características se detecten en tan sólo un evento de imagen, en tan poco como dos eventos de imagen durante el evento de detección. Una primera imagen obtenida usando la primera longitud de onda de excitación y emisión en el primer canal puede detectar y mostrar características que incorporan el primer y/o tercer tipo de nucleótidos (p. ej., A y/o T). Una segunda imagen obtenida usando la segunda longitud de onda de excitación y emisión en el segundo canal puede detectar y mostrar características que incorporan el segundo y/o tercer tipo de nucleótido (p. ej., C y/o T). La identificación inequívoca del tipo de nucleótido incorporado en cada característica se puede determinar, por ejemplo, comparando las dos imágenes para llegar a lo siguiente: características que aparecen (es decir, se detectan) al máximo en el primer canal que incorpora el primer tipo de nucleótido (p. ej., A), características que se muestran al máximo en el segundo canal que incorpora el segundo tipo de nucleótido (p. ej., C), características que aparecen en ambos canales que incorporan el tercer tipo de nucleótido (p. ej., T) y características que no se muestran o que son mínimamente detectables en cualquiera de los canales que incorporan el cuarto tipo de nucleótido (p. ej., G).

Alternativamente, la incorporación de los cuatro nucleótidos se puede determinar usando solamente un evento de formación de imágenes combinado. Por ejemplo, la incorporación de los tipos de nucleótidos marcados puede

determinarse exponiendo los nucleótidos incorporados a dos longitudes de onda de excitación de una vez (por ejemplo, simultáneamente) y capturando los espectros de emisión en una imagen combinada. La identificación inequívoca de los tipos de nucleótidos incorporados podría determinarse como se ha indicado anteriormente; las características que aparecen en un canal de la imagen combinada indican la incorporación de ese tipo nucleotídico marcado (por ejemplo, A), las características que aparecen en el segundo canal de la imagen combinada indican la incorporación de ese tipo nucleotídico marcado (por ejemplo, C) y las características que aparecen en ambos canales indicaría la incorporación de un tercer tipo de nucleótidos (por ejemplo, T). Como uno de los tipos de nucleótidos no está marcado (por ejemplo, G) la incorporación se determina por la ausencia de, o por las características mínimamente medibles, en ambos canales para ese nucleótido no marcado. Obsérvese que la localización de las características que incorporan G en este ejemplo se puede determinar a partir de otros ciclos (donde al menos uno de los otros tres tipos de nucleótidos se incorpora).

En la presente descripción, se proporcionan métodos para determinar la secuencia de un polinucleótido que comprende detectar en una reacción de secuenciación la incorporación de tres tipos diferentes de conjugados de nucleótidos detectables en un polinucleótido y determinar la incorporación de un cuarto tipo de nucleótido basado en el patrón de detección de los tres tipos diferentes de nucleótidos detectables en el polinucleótido, determinando con ello la secuencia de un polinucleótido, en el que la incorporación de tres tipos diferentes de conjugados de nucleótidos detectables se detecta a partir de un estado de señal y en el que la incorporación del cuarto tipo de nucleótido se determina a partir de un estado de oscuridad.

En otra realización, la presente descripción proporciona métodos para determinar la secuencia de un polinucleótido que comprende aplicar a una muestra de polinucleótido para secuenciar una solución que comprende cuatro tipos de nucleótidos modificados en los que tres tipos de nucleótidos modificados están conjugados con uno o más restos de detección y uno o más enlazantes colocados entre el nucleótido y uno o más restos de detección, y en el que un cuarto tipo de nucleótido carece de un resto de detección, detectando un patrón de incorporación de dichos nucleótidos modificados en una reacción de secuenciación capturando de este modo un primer patrón detectable, aplicando una o más composiciones a la reacción de secuenciación, por lo tanto, cambiando el primer patrón detectable, detectando un segundo patrón detectable y determinando la secuencia de la muestra polinucleotídica basada en los patrones detectables.

En algunas realizaciones, el polinucleótido para la secuenciación comprende uno o más de ácidos desoxirribonucleicos, ácidos desoxirribonucleicos modificados, ácidos ribonucleicos y ácidos ribonucleicos modificados. En algunas realizaciones, el polinucleótido para la secuenciación es una preparación de biblioteca de ADN genómico. En algunas realizaciones, el conjugado de nucleótidos comprende tipos de nucleótidos seleccionados del grupo que consiste en dATP, dTTP, dUTP, dCTP, dGTP o sus análogos de nucleótidos no naturales. En algunas realizaciones, el análogo de nucleótido no natural comprende un resto terminador reversible y se selecciona del grupo que consiste en rbATP, rbTTP, rbCTP, rbUTP y rbGTP. En algunas realizaciones, la incorporación de nucleótidos es una secuencia por síntesis, secuencia por ligación y secuencia por hibridación o una combinación de los mismos. En algunas realizaciones, los tres conjugados de tipo nucleótido se detectan detectando un resto fluorescente. En algunas realizaciones, el resto fluorescente es el mismo para los tres conjugados de nucleótidos mientras que en otras realizaciones el resto fluorescente es uno o más restos fluorescentes diferentes. En algunas realizaciones, el uno o más restos fluorescentes diferentes son detectados por el mismo filtro de emisión. En algunas realizaciones, el resto fluorescente comprende un resto de sistema de transferencia de energía de resonancia fluorescente. En algunas realizaciones, la incorporación del cuarto nucleótido se determina por falta de detección. En algunas realizaciones, los conjugados de ácido nucleico detectables se detectan por fluorescencia. En algunas realizaciones, la fluorescencia se detecta mediante un primer y un segundo evento de formación de imágenes, en otras realizaciones, los primeros y segundos eventos de formación de imágenes se separan en el tiempo. En algunas realizaciones, el primer evento de formación de imágenes detecta un patrón de fluorescencia que es diferente del patrón de fluorescencia detectado por el segundo evento de formación de imágenes. En algunas realizaciones, la incorporación de uno o más nucleótidos está determinada por la diferencia en el patrón de fluorescencia entre el primer y segundo eventos de formación de imágenes. En algunas realizaciones, uno o más conjugados de tipo nucleotídico comprenden además una o más secuencias enlazantes, en otras realizaciones la una o más secuencias enlazantes comprenden uno o más de un enlazante escindible y un enlazante espaciador. En algunas realizaciones, el enlazante escindible comprende uno o más grupos de enlace escindibles seleccionados del grupo que consiste en un disulfuro, un diol, un diazo, un éster, una sulfona, una azida, un alilo y un éter de sililo, mientras que en realizaciones preferidas el grupo de unión escindible es un disulfuro. En algunas realizaciones, el enlazador espaciador es uno o más de polietilenglicol o sus concatámeros y el ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoil-1-azido-etoxi]-etoxi}-acético. En algunas realizaciones, uno o más enlazantes espaciadores comprenden además uno o más grupos de enlace escindibles en los que el grupo de enlace escindible se selecciona del grupo que consiste en un disulfuro, un diol, un diazo, un éster, una sulfona, una azida, un alilo y un éter de sililo. En algunas realizaciones, el enlazante espaciador es polietilenglicol o sus concatámeros, mientras que en otras realizaciones el enlazante espaciador es el ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoil-1-azido-etoxi]-etoxi}-acético. En algunas realizaciones, uno o más conjugados de nucleótidos comprenden un enlazador de polietilenglicol y un enlazante de ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoil-1-azido-etoxi]-etoxi}-acético que puede o no puede comprender además un hapteno y un resto fluorescente. En algunas realizaciones, el hapteno se selecciona del grupo que consiste en biotina, digoxigenina y dinitrofenol. En algunas realizaciones, uno o más conjugados de

nucleótidos comprende un conjugado de estreptavidina-resto fluorescente, mientras que en otras realizaciones, uno o más conjugados de nucleótidos comprenden un conjugado anticuerpo anti-hapteno-resto fluorescente seleccionado del grupo que consiste en anti-digoxigenina y anti-dinitrofenol. En algunas realizaciones, el conjugado de nucleótidos que comprende un enlazante de polietilenglicol y un enlazador de ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético comprende además dos restos fluorescentes. En algunas realizaciones, los dos restos fluorescentes constituyen un sistema de transferencia de energía de resonancia de fluorescencia.

Un aspecto adicional de la presente descripción proporciona una composición para uso en el método SBS descrito anteriormente para secuenciar un ácido nucleico y como se define en la reivindicación 12, que comprende tres tipos de nucleótidos modificados detectables por un resto fluorescente y un cuarto tipo de nucleótido modificado, en el que dicho cuarto tipo de nucleótido modificado no es detectable por un resto fluorescente y en el que la incorporación de los cuatro tipos de nucleótidos modificados en la composición en una reacción de secuenciación se determina mediante la detección fluorescente de los tres tipos de nucleótidos modificados detectables en la composición. En algunas realizaciones, el ácido nucleico de composición comprende ADN de una preparación de biblioteca de ADN. En algunas realizaciones, el tipo de nucleótido modificado comprende un resto terminador reversible y se selecciona del grupo que comprende rbATP, rbTTP, rbUTP, rbCTP y rbGTP. En algunas realizaciones, la reacción de secuenciación es una secuencia por síntesis, secuencia por ligación o secuencia por hibridación. En algunas realizaciones, el resto fluorescente es el mismo para los tres nucleótidos modificados. En algunas realizaciones, el resto fluorescente es uno o más restos fluorescentes diferentes que se detectan preferiblemente por el mismo filtro de emisión. En algunas realizaciones, la incorporación de tres tipos de nucleótidos modificados se determina mediante un primer patrón de formación de imágenes fluorescentes y un segundo patrón de formación de imágenes fluorescentes. En algunas realizaciones, la incorporación del cuarto tipo de nucleótido está determinada por los patrones de formación de imágenes de fluorescencia de los otros tres tipos de nucleótidos. En algunas realizaciones, las composiciones descritas en la presente memoria que comprenden uno o más de los tipos de nucleótidos modificados comprenden además una o más secuencias enlazantes. En algunas realizaciones, la una o más secuencias enlazantes comprenden uno o más de un enlazante escindible y un enlazante espaciador, en el que el enlazante escindible comprende uno o más grupos de enlace escindibles seleccionados del grupo que consiste en un disulfuro, un diol, un diazo, un éster, una sulfona, una azida, un alilo y un éter de sililo, preferiblemente el grupo de enlace escindible es disulfuro. En algunas realizaciones, el enlazador espaciador es uno o más de polietilenglicol o sus concatámeros y el ácido 2-{2-[3-(2-amino-etilcarbomil)fenoxil-1-azido-etoxi]-etoxi-acético, en el que los concatámeros de polietilenglicol incluyen entre cuatro y doce moléculas de polietilenglicol. En algunas realizaciones, el uno o más enlazantes espaciadores comprenden además uno o más grupos de enlace escindibles como se ha descrito anteriormente. En algunas realizaciones, uno o más de tres tipos de nucleótidos modificados comprenden un enlazante de polietilenglicol y un enlazador de ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético, mientras que algunas realizaciones preferidas comprenden además un hapteno y un resto fluorescente. En algunas realizaciones, el hapteno se selecciona del grupo que consiste en biotina, digoxigenina y dinitrofenol. En algunas realizaciones, el hapteno se detecta mediante un conjugado de pareja de unión de hapteno-resto fluorescente o un conjugado de anticuerpo anti-hapteno-resto fluorescente. En algunas realizaciones, el anticuerpo anti-hapteno se selecciona entre anti-digoxigenina y anti-dinitrofenol. En algunas realizaciones, dicha pareja de unión de hapteno es estreptavidina. En algunas realizaciones, dichos tipos de nucleótidos modificados detectables por un resto fluorescente se conjugan con uno o más de un enlazante escindible y un enlazante espaciador o una combinación de los mismos, en el que un enlazante se conjuga con un resto fluorescente o un hapteno, y en el que un nucleótido modificado que no es detectable por un resto fluorescente no está conjugado de esa manera.

Una realización adicional, según se describe en el presente documento, proporciona un método para determinar una pluralidad de secuencias de ácido nucleico que comprende proporcionar una muestra que comprende una pluralidad de diferentes ácidos nucleicos, comprendiendo cada ácido nucleico una plantilla y un cebador; realizar un ciclo de una reacción de secuenciación, en el que el ciclo comprende extender los cebadores para los ácidos nucleicos en la muestra para formar una pluralidad de cebadores extendidos que tienen al menos cuatro tipos de nucleótidos diferentes, formando así una muestra extendida, adquiriendo una primera colección de señales de la muestra extendida, en la que no más de tres de los diferentes tipos de nucleótidos en los cebadores extendidos están en un estado de señal y en el que al menos uno de los diferentes tipos de nucleótidos en los cebadores extendidos está en estado oscuro; tratar la muestra extendida con un reactivo modificador, en el que se modifica al menos uno de los diferentes tipos de nucleótidos en los cebadores extendidos, produciendo de este modo una muestra modificada, y adquiriendo una segunda colección de señales de la muestra modificada, en la que al menos uno de los diferentes tipos de nucleótidos está en estado diferente en la primera colección de señales comparada con la segunda colección de señales; y determinando las secuencias para la pluralidad de diferentes ácidos nucleicos mediante la evaluación de la primera colección de señales y la segunda colección de señales de los ciclos. En algunas realizaciones, la pluralidad de ácidos nucleicos diferentes está unida a un sustrato. En algunas realizaciones, la extensión de los cebadores comprende la adición catalizada por polimerasa de los diferentes tipos de nucleótidos. En algunas realizaciones, los diferentes tipos de nucleótidos comprenden restos de bloqueo reversibles, con lo que se añade un tipo de nucleótido único a cada uno de los cebadores extendidos en cada uno de los ciclos. En algunas realizaciones, la extensión de los cebadores comprende la adición catalizada por ligasa de oligonucleótidos que comprenden los diferentes tipos de nucleótidos. Como se discute en este documento, durante la adquisición de la

primera colección de señales de la muestra extendida, no más de dos de los diferentes tipos de nucleótidos en los cebadores extendidos están en un estado de señal, mientras que alternativamente durante la adquisición de la primera colección de señales de la muestra extendida, al menos dos de los diferentes tipos de nucleótidos en los cebadores extendidos están en un estado oscuro. En algunas realizaciones, uno de los diferentes tipos de nucleótidos en los cebadores extendidos se encuentra en estado oscuro durante la adquisición de la primera colección de señales de la muestra extendida. En algunas realizaciones, el tratamiento de la muestra extendida con un reactivo modificador comprende eliminar un marcador de un tipo de nucleótido o añadir un marcador a un tipo de nucleótido. En algunas realizaciones, al menos dos de los diferentes tipos de nucleótidos en los cebadores extendidos se modifican por el tratamiento de la muestra extendida con un reactivo modificador, mientras que en otras realizaciones no más de 3 de los diferentes tipos de nucleótidos en los cebadores extendidos se modifican por el tratamiento de la muestra extendida con un reactivo modificador. En algunas realizaciones, la extensión de los cebadores para los ácidos nucleicos en la muestra forma una pluralidad de cebadores extendidos que no tienen más de cuatro tipos de nucleótidos diferentes, mientras que en otras realizaciones la extensión de los cebadores para los ácidos nucleicos en la muestra forma una pluralidad de cebadores extendidos que tienen al menos cinco tipos de nucleótidos diferentes. En algunas realizaciones, dos de los diferentes tipos de nucleótidos complementan el mismo nucleótido en el ácido nucleico y en el que un primer de los dos tipos de nucleótidos diferentes está en un estado de señal durante la adquisición de la primera colección de señales y en el que un segundo de los dos tipos de nucleótidos diferentes está en un estado oscuro durante la adquisición de la primera colección de señales. En algunas realizaciones, el primer de los dos tipos de nucleótidos diferentes está en estado oscuro durante la adquisición de la segunda colección de señales. En algunas realizaciones, el segundo de los dos tipos de nucleótidos diferentes está en un estado de señal durante la adquisición de la segunda colección de señales. En realizaciones preferidas, se repite una o más veces un ciclo de reacción de secuenciación como el descrito anteriormente.

En otra realización, la presente descripción proporciona un método para determinar la secuencia de un polinucleótido que comprende detectar por eventos de formación de imágenes la incorporación de tres tipos diferentes de conjugados de nucleótidos detectables en un polinucleótido y determinar la incorporación de un cuarto tipo de nucleótido basado en el patrón de detección de los tres tipos diferentes de nucleótidos detectables en el polinucleótido, donde la detección comprende menos eventos de formación de imágenes que los diferentes tipos de conjugados de nucleótidos detectables. En algunas realizaciones, el polinucleótido comprende uno o más de ácidos desoxirribonucleicos, ácidos desoxirribonucleicos modificados, ácidos ribonucleicos o ácidos ribonucleicos modificados. En algunas realizaciones, el conjugado de nucleótidos comprende los tipos de nucleótidos seleccionados del grupo que consiste en dATP, dTTP, dUTP, dCTP, dGTP o análogos de nucleótidos no naturales de los mismos, en los que el análogo de nucleótido no natural comprende un resto terminador reversible y se selecciona del grupo que consiste en rbATP, rbTTP, rbCTP, rbUTP y rbGTP. En algunas realizaciones, la incorporación de nucleótidos es una secuencia por síntesis, secuencia por ligación o secuencia por hibridación. En algunas realizaciones, los tres conjugados de tipo nucleótido se detectan detectando un resto fluorescente, en el que el resto fluorescente es el mismo para los tres conjugados de nucleótidos o en el que el resto fluorescente es uno o más restos fluorescentes diferentes. En algunas realizaciones, uno o más restos fluorescentes diferentes son detectados por el mismo filtro de emisión. En algunas realizaciones, el resto fluorescente comprende un resto de sistema de transferencia de energía de resonancia fluorescente. La incorporación del cuarto nucleótido se determina por falta de detección. En algunas realizaciones, los conjugados de ácido nucleico detectables se detectan por fluorescencia en la que la fluorescencia es detectada por los eventos de formación de imágenes. En algunas realizaciones, los eventos de formación de imágenes comprenden un primer y un segundo evento de formación de imágenes, por ejemplo, que están separados en el tiempo. En algunas realizaciones, el primer evento de formación de imágenes detecta un patrón de fluorescencia que es diferente del patrón de fluorescencia detectado por el segundo evento de formación de imágenes. En algunas realizaciones, la incorporación de uno o más nucleótidos está determinada por la diferencia en el patrón de fluorescencia entre el primer y segundo eventos de formación de imágenes. En algunas realizaciones, uno o más conjugados de tipo nucleotídico comprenden además una o más secuencias enlazantes que comprenden uno o más de un enlazante escindible y un enlazante espaciador. En algunas realizaciones, el enlazante escindible comprende uno o más grupos de enlace escindibles seleccionados del grupo que consiste en un disulfuro, un diol, un diazo, un éster, una sulfona, una azida, un alilo y un éter de sililo, preferiblemente el grupo de enlace escindible es un disulfuro. En algunas realizaciones, el enlazador espaciador es uno o más de polietilenglicol o sus concatámeros y el ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético. En algunas realizaciones, uno o más enlazantes espaciadores comprenden además uno o más grupos de enlace escindibles en los que el grupo de enlace escindible se selecciona del grupo que consiste en un disulfuro, un diol, un diazo, un éster, una sulfona, una azida, un alilo y un éter de sililo. En algunas realizaciones, el enlazador espaciador es polietilenglicol o sus concatámeros o el ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético o ambos. En algunas realizaciones, el conjugado de nucleótidos que comprende un enlazante de polietilenglicol y un enlazante de ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético comprende además un hapteno y un resto fluorescente, en el que el hapteno se selecciona del grupo que consiste en biotina, digoxigenina y dinitrofenol. En algunas realizaciones, uno o más conjugados de nucleótidos comprenden un conjugado de estreptavidina-resto fluorescente. En algunas realizaciones, uno o más conjugados de nucleótidos comprenden un conjugado de anticuerpo anti-hapteno-resto fluorescente seleccionado del grupo que consiste en anti-digoxigenina y anti-dinitrofenol. En algunas realizaciones, el conjugado de nucleótidos que comprende un enlazante de polietilenglicol y un enlazante de ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético

comprende además dos restos fluorescentes. En algunas realizaciones, los dos restos fluorescentes constituyen un sistema de transferencia de energía de resonancia de fluorescencia. En algunas realizaciones, uno o más conjugados de nucleótidos comprenden además un hapteno o un resto fluorescente, en el que el hapteno se selecciona del grupo que consiste en biotina, digoxigenina y dinitrofenol. En algunas realizaciones, uno o más conjugados de nucleótidos comprenden un conjugado de estreptavidina-resto fluorescente. En algunas realizaciones, la detección de uno o más conjugados de nucleótidos comprende un conjugado de anticuerpo anti-hapteno-resto fluorescente seleccionado del grupo que consiste en anti-digoxigenina y anti-dinitrofenol.

Figuras

La Figura 1 muestra mapas de calor tipo nube ejemplares, o diagramas de nubes, para ciclos en una reacción de secuenciación. Las gráficas representan el compuesto de imagen 1 (eje x) e imagen 2 (eje y), de tal manera que las representaciones representan la imagen de fluorescencia después de un ciclo completo. La ubicación de A, C, G y T en el diagrama de nubes se demuestra en el mapa del diagrama de nube inferior.

La Figura 2 muestra gráficas ejemplares que informan del porcentaje de tasas de error (eje Y) (parte superior) y de los ciclos basales en blanco (parte inferior) en una reacción de secuenciación en un ciclo por base de ciclo (eje X).

La Figura 3 muestra A) los espectros de emisión para dos colorantes ejemplares y B) un diagrama de nubes ejemplar para un ciclo de secuenciación cuando se practica la realización de usar dos colorantes de diferentes espectros de fluorescencia para la secuenciación.

La Figura 4 muestra A) los espectros de emisión para dos conjuntos de colorantes ejemplares y B) un diagrama de nubes ejemplar para un ciclo de secuenciación cuando se practica la realización de usar dos conjuntos de colorantes de diferentes espectros de emisión para la secuenciación.

La Figura 5 muestra A) la tasa de error de los ciclos basales para un experimento usando un colorante en una reacción de secuenciación, B) patrones fluorescentes ejemplares para cada uno de los nucleótidos modificados en un primer evento de imagen (Imagen 1) usando sólo un colorante, C) patrones ejemplares fluorescentes para cada uno de los nucleótidos modificados en un segundo evento de imagen (imagen 2) utilizando sólo un colorante, y D) un diagrama de nubes combinando el primer y segundo eventos de formación de imágenes de una reacción de secuenciación en la que sólo se utilizan un colorante y dos eventos de formación de imágenes para diferenciar entre los cuatro nucleótidos diferentes presentes para su incorporación durante una reacción de secuenciación.

Descripción detallada

Las tecnologías actuales basadas en la fluorescencia utilizadas para diferenciar entre diferentes analitos en una muestra, tal como se encuentran en las tecnologías de secuenciación (es decir, tecnologías de secuenciación por fluorescencia) se basan, por ejemplo, en la calidad de una señal generada por un resto de detección asociado con un tipo particular de nucleótido. Por ejemplo, las tecnologías de secuenciación fluorescente tradicionales utilizan fragmentos fluorescentes claramente identificables, cada uno unido a uno de los cuatro nucleótidos A, T, C y G que se utilizan en una reacción de secuenciación. Los nucleótidos marcados fluorescentemente utilizados durante una reacción de secuenciación, independientemente de su método de utilización, son típicamente excitados y medidos por uno de cuatro filtros ópticos (es decir, uno para cada colorante distinto) en un instrumento de secuenciación fluorescente. La tecnología de secuenciación por síntesis (SBS), así como la tecnología de secuenciación de terminadores de colorante que utiliza dideoxinucleótidos, son ejemplares de las tecnologías de secuenciación basadas en fluorescencia de cuatro canales. La instrumentación de secuenciación basada en fluorescencia es típicamente grande, costosa y poco atractiva para los medios más pequeños y más limitados de capital. Las nuevas tecnologías de secuenciación usualmente utilizan métodos, sistemas y composiciones innovadores para avanzar en la precisión (es decir, menos errores), tener mayor capacidad de procesamiento (es decir, más secuencias de genomas por período de tiempo dado) y/o reducir costos (es decir, <\$ 10.000/genoma), y deseablemente tener una ocupación de espacio que no exceda el pequeño espacio en la mesa de trabajo de un investigador.

La presente descripción proporciona soluciones para avanzar en el campo de la secuenciación de ácidos nucleicos. En las realizaciones se describen métodos y composiciones que utilizan restos de detección mínimos, por ejemplo preferiblemente un colorante, o una pluralidad de colorantes con características de detección similares, al detectar y diferenciar múltiples analitos diferentes, tales como diferentes tipos de nucleótidos, en una muestra, por ejemplo para secuenciación de muestras. Además, la presente descripción proporciona métodos para determinar la incorporación de cuatro nucleótidos en una reacción de secuenciación utilizando menos de cuatro filtros de detección y menos etapas de formación de imágenes. El uso de menos de cuatro filtros y, por lo tanto, menos etapas de formación de imágenes permite realizar la secuenciación en formatos más pequeños, ya que hay que tener menos filtros de excitación y emisión. Se contempla que los métodos y sistemas descritos en la presente memoria disminuyen las necesidades de hardware del instrumento, disminuyen el tamaño del instrumento, el uso de reactivos y los costes al tiempo que aumentan la producción de datos.

En realizaciones particulares, se proporcionan métodos para determinar una secuencia de subunidades monoméricas en un polímero. Los métodos se ejemplifican en este documento con respecto a polímeros de ácido nucleico y sus subunidades nucleotídicas, pero pueden llevarse a cabo para otros polímeros y sus subunidades.

Aunque los métodos pueden usarse para muestras que tienen una única secuencia polimérica, los métodos proporcionan ventajas particulares cuando se usan para distinguir varios tipos de subunidades diferentes en una muestra que tiene polímeros con muchas secuencias diferentes (es decir, una muestra de polímero multiplex). Por ejemplo, en algunas realizaciones los métodos proporcionan la capacidad de distinguir un número de diferentes tipos de subunidades en una muestra que es mayor que el número de diferentes tipos de señales que se adquieren de la muestra. En el caso de una muestra de ácido nucleico, se puede realizar una etapa de adquisición de datos en la muestra para adquirir una colección de menos de cuatro tipos de señal diferentes y, sin embargo, la ubicación de la secuencia para los cuatro tipos de nucleótidos diferentes puede determinarse para la muestra.

Varios aspectos de los métodos, individualmente o en combinación, proporcionan la capacidad de distinguir una serie de diferentes tipos de subunidades (por ejemplo, diferentes tipos de nucleótidos, diferentes tipos de didesoxinucleótidos, tipos de didesoxinucleótidos modificados, tipos de nucleótidos modificados unidos de forma reversible, etc.) en una muestra de polímero que es mayor que el número de tipos de señal diferentes adquiridos a partir de la muestra de polímero. Los aspectos pueden incluir, pero no se limitan a, correlacionar uno o más tipos de subunidades monoméricas con un estado oscuro, correlacionar uno o más tipos de subunidades monoméricas con un estado gris o correlacionar uno o más tipos de subunidades monoméricas con un cambio de estado entre un estado oscuro, estado gris o estado de señal. Un "estado de señal", cuando se usa en referencia a un evento de detección, significa una condición en la que se produce una señal específica en el evento de detección. Por ejemplo, una subunidad de nucleótidos puede estar en un estado de señal y ser detectable cuando está unida a un marcador fluorescente que se detecta en una etapa de detección de fluorescencia por excitación y emisión de ese marcador fluorescente en un método de secuenciación. La expresión "estado oscuro", cuando se usa en referencia a un evento de detección, significa una condición en la que no se produce una señal específica en el evento de detección. Por ejemplo, una subunidad de nucleótidos puede estar en estado oscuro cuando el nucleótido carece de un marcador fluorescente y/o no emite fluorescencia que se detecta específicamente en un paso de detección fluorescente de un método de secuenciación. La detección del estado oscuro también puede incluir cualquier fluorescencia de fondo que pueda estar presente sin un marcador fluorescente. Por ejemplo, algunos componentes de reacción pueden demostrar una fluorescencia mínima cuando se excitan a ciertas longitudes de onda. Como tal, aunque no haya un resto fluorescente presente puede haber fluorescencia de fondo de tales componentes. Además, la fluorescencia de fondo puede ser debida a la dispersión de luz, por ejemplo a partir de reacciones de secuenciación adyacentes, que pueden ser detectadas por un detector. Como tal, un "estado oscuro" puede incluir una fluorescencia de fondo tal como cuando un resto fluorescente no está específicamente incluido, tal como cuando un nucleótido que carece de un marcador fluorescente se utiliza en los métodos descritos en la presente memoria. Sin embargo, se contempla que dicha fluorescencia de fondo es diferenciable desde un estado de señal y, como tal, la incorporación de nucleótidos de un nucleótido no marcado (o nucleótido "oscuro") es todavía discernible. La expresión "estado gris", cuando se usa en referencia a un evento de detección, significa una condición en la que se produce una señal atenuada en el evento de detección. Por ejemplo, una población de nucleótidos de un tipo particular puede estar en un estado gris cuando una primera subpoblación de los nucleótidos unidos a un marcador fluorescente se detecta en un paso de detección de fluorescencia de un método de secuenciación mientras que una segunda subpoblación de los nucleótidos carece del marcador fluorescente y no emite fluorescencia que se detecta específicamente en la etapa de detección de fluorescencia.

En realizaciones particulares, el método para secuenciar un polímero se lleva a cabo en ciclos, en el que un ciclo individual incluye una o más etapas usadas para distinguir un monómero en una posición particular en el polímero. Un ciclo puede comprender un evento de detección en algunas realizaciones. Sin embargo, un ciclo de secuenciación no necesita incluir un evento de detección, por ejemplo, si la detección se lleva a cabo después de que se llevan a cabo etapas para distinguir uno o más monómeros en un polímero. Por ejemplo, un evento de detección puede ocurrir a la mitad de un ciclo, al final de un ciclo, al final de 1 ½ ciclos, al final de 2 ½ ciclos, al final de tres ciclos, etc. Un aspecto adicional de los métodos que pueden proporcionar la capacidad de distinguir un número de diferentes tipos de subunidades en una muestra polimérica, que es mayor que el número de diferentes tipos de señal adquiridos a partir de la muestra de polímero, es el uso de dos o más etapas de adquisición de señal y al menos una etapa de modificación de nucleótidos durante un ciclo de secuenciación individual. Como tal, un método de secuenciación puede incluir varios ciclos de adición de nucleótidos y los ciclos pueden incluir etapas ortogonales de adquisición de señales de la muestra de secuenciación, modificando después uno o más nucleótidos en la muestra de secuenciación para cambiar su estado (por ejemplo, entre un estado de señal, estado de oscuridad o estado gris) y, a continuación, adquirir un segundo conjunto de señales de la muestra de secuenciación. Varios ejemplos se exponen con más detalle a continuación en los que tipos de nucleótidos particulares están en un estado de señal debido a un marcador fluorescente unido, tipos de nucleótidos particulares están en un estado oscuro debido a la ausencia del marcador, se convierten nucleótidos en particular de un estado de señal a un estado oscuro mediante la escisión de un enlazante que une un marcador fluorescente y/o nucleótidos particulares se convierten de un estado oscuro a un estado de señal por unión de un receptor (por ejemplo, anticuerpo o estreptavidina) que recluta un marcador fluorescente al nucleótido que de otro modo no tenía el marcador.

En lugar de detectar diferencias en la calidad de una señal fluorescente, por ejemplo, tal como se practica para algunas tecnologías de secuenciación fluorescentes, la presente descripción proporciona la detección de múltiples

analitos diferentes (es decir, nucleótidos, proteínas o fragmentos de los mismos) en una reacción distinguiendo entre diferencias en la detección de un resto fluorescente o dos restos fluorescentes de espectros de excitación/emisión iguales o similares (es decir, excitados por el mismo láser y emisión capturados por el mismo filtro óptico), en momentos diferentes durante una reacción, por ejemplo antes y después de un cambio en las condiciones de reacción. En algunas realizaciones, los métodos para detectar y determinar un analito comprenden detectar la salida de fluorescencia en dos momentos diferentes durante un ciclo de reacción.

Típicamente, se llevará a cabo un ciclo de reacción suministrando al menos cuatro tipos de nucleótidos a una muestra de ácido nucleico en presencia de una polimerasa, por ejemplo una ADN o ARN polimerasa, durante una reacción de extensión del cebador. La presencia de al menos cuatro tipos de nucleótidos proporciona una ventaja de aumentar la fidelidad de la polimerasa en comparación con el uso de menos de cuatro tipos de nucleótidos. El uso de etapas ortogonales para convertir uno o más tipos de nucleótidos incorporados de un estado a otro, permite que múltiples tipos de nucleótidos estén presentes simultáneamente durante una reacción de extensión de polimerasa, aumentando así la fidelidad, permitiendo también que se detecte un único tipo de marcador en cada ciclo, lo que sirve para proporcionar una óptica más simplificada. El uso de una óptica simplificada es preferencial en comparación con sistemas que se basan en una óptica más compleja para registrar la salida de múltiples marcadores diferentes para distinguir diferentes tipos de nucleótidos que están presentes simultáneamente en una reacción de extensión. Se contempla además que en algunas realizaciones menos que cuatro tipos diferentes de nucleótidos puedan estar presentes durante una reacción de extensión de polimerasa.

A continuación se describen algunas realizaciones ilustrativas. Las composiciones y sus métodos de uso no se limitan a estas realizaciones.

En algunas realizaciones, los métodos para secuenciar un ácido nucleico comprenden el uso de un resto fluorescente para la detección directa o indirecta de tres tipos diferentes de nucleótidos y un tipo de nucleótido que no se detecta por la presencia de una señal fluorescente sino que se detecta por la falta o ausencia de una señal fluorescente. En algunas realizaciones, los métodos para secuenciar un ácido nucleico comprenden el uso de dos o más restos fluorescentes diferentes que comprenden los mismos o similares espectros de excitación/emisión para la detección directa o indirecta de tres tipos diferentes de nucleótidos y un tipo de nucleótido que no es detectado por la presencia de una señal fluorescente, sino que se detecta por una falta o ausencia de señal fluorescente. Los mismos o similares espectros de excitación y emisión son tales que un láser excita los dos o más restos fluorescentes diferentes y un filtro óptico capta sus señales de fluorescencia emitidas. La detección de la fluorescencia para determinar la secuencia de una muestra de ácido nucleico se realiza en un espacio de tiempo, por ejemplo en momentos diferentes durante una reacción de secuenciación (es decir, antes y después de un cambio en las condiciones de reacción tales como escisión enzimática, cambio en el pH ambiental, adición de reactivos adicionales), proporcionando patrones de fluorescencia tales como patrones de transiciones de fluorescencia, sus patrones acumulativos determinando la secuencia de la diana del ácido nucleico. Como tal, los métodos descritos en la presente memoria son eficaces en términos de tiempo y de coste y permiten simplificar la instrumentación de secuenciación asociada.

Una aplicación ejemplar de la utilización de las diferencias del patrón de fluorescencia en el espacio de tiempo para determinar una secuencia de ácido nucleico diana es una secuencia por metodologías y tecnologías de síntesis (SBS). Como tal, las realizaciones descritas en este documento encuentran utilidad particular en la secuenciación mediante aplicaciones de fluorescencia de síntesis. Aunque las realizaciones descritas en este documento son ejemplares de métodos innovadores de secuenciación fluorescente, las realizaciones descritas también encuentran utilidad para una variedad de otras aplicaciones en las que se desea la detección de más de un analito (es decir, un nucleótido, una proteína o fragmentos de los mismos) en una muestra.

En las realizaciones de desarrollo para la secuenciación usando un conjunto de colorantes mínimo, la experimentación reveló estrategias alternativas para distinguir entre incorporaciones de nucleótidos usando solo uno o dos restos fluorescentes. Estas estrategias proporcionan que los cuatro tipos de nucleótidos estén simultáneamente presentes en un ciclo de secuencia, y para el uso de colorantes mínimos y conjuntos de filtros ópticos. En algunas realizaciones, no se utilizan más de tres restos fluorescentes para determinar la incorporación de los cuatro tipos de nucleótidos que están presentes durante una reacción, usando uno o dos filtros de excitación y emisión. En realizaciones preferidas, no se utiliza más de un resto fluorescente (o dos o tres de los mismos o similares espectros de excitación/emisión) para determinar la incorporación de los cuatro tipos de nucleótidos que están presentes durante una reacción, usando un rango de excitación de luz y un filtro de emisión de detección. Se entenderá que, en algunas realizaciones, pueden usarse más de un resto fluorescente (o fragmentos de más de un rango de excitación o rango de emisión).

En algunas realizaciones, la secuenciación utilizando un conjunto de colorantes mínimo se realiza sobre un sustrato, tal como un vidrio, un plástico, un chip semiconductor o un sustrato derivado compuesto. En algunas realizaciones, se proporciona una especie de ácido nucleico sobre un sustrato, por ejemplo, para la secuenciación de una diana única. En otras realizaciones, la secuenciación también puede estar en un formato multiplex, en el que se detectan y secuencian múltiples objetivos de ácidos nucleicos en paralelo, por ejemplo en un formato de tipo de celda de flujo o de matriz. Las realizaciones descritas en la presente memoria son particularmente ventajosas cuando se practica una secuenciación paralela o una secuenciación paralela masiva. Las plataformas que practican la secuenciación

paralela fluorescente incluyen, pero no se limitan a, las ofrecidas por Illumina, Inc. (p. ej., HiSeq, Genome Analyzer, MiSeq, plataformas iScan), Life Technologies (por ejemplo SOLiD), Helicos Biosciences (p. ej., Heliscope), 454/Roche Life Sciences (Branford, CT) y Pacific Biosciences (por ejemplo, SMART). Las células de flujo, los chips y otros tipos de superficies que pueden acomodar múltiples especies de ácidos nucleicos son ejemplares de los sustratos utilizados para la secuenciación paralela. En los formatos multiplex, en los que se secuencian múltiples especies de ácidos nucleicos en paralelo, las secuencias diana amplificadas clonalmente (por ejemplo, mediante PCR en emulsión (emPCR) o amplificación en puente) típicamente se inmovilizan covalentemente sobre un sustrato. Por ejemplo, cuando se practica la PCR en emulsión, el objetivo de interés se inmoviliza en unas perlas, mientras que las dianas amplificadas clonalmente se inmovilizan en canales de una célula de flujo o localizaciones específicas en una matriz o chip.

Las células de flujo para su uso con las composiciones y los métodos que se describen en este documento pueden usarse en la secuenciación de varias maneras. Por ejemplo, una muestra de ADN, tal como una biblioteca de ADN, se puede aplicar a una célula de flujo o dispositivo fluido que comprende uno o más canales de flujo grabados, en el que la célula de flujo puede comprender además una población de moléculas de sonda unidas covalentemente a su superficie. Las sondas fijadas en los canales de la célula de flujo están situadas ventajosamente en diferentes posiciones direccionables en el canal y las moléculas de la biblioteca de ADN pueden añadirse a los canales de la célula de flujo en las que las secuencias complementarias pueden unirse (como se describe en la presente memoria descriptiva además como se describe en el documento WO 2012/096703). Otro ejemplo de una célula de flujo para uso en la presente solicitud comprende una célula de flujo CMOS como se describe en la Solicitud de Patente de EE.UU. US2013/0058051. Pueden realizarse amplificaciones de puente como se describe en este documento seguido por la secuenciación por métodos de síntesis y composiciones como se describe en la presente memoria. Los métodos para crear y utilizar las células de flujo para la secuenciación son conocidos en la técnica. Se contempla que los métodos y composiciones como se describen en la presente memoria descriptiva no estén limitados a ninguna fabricación o método particular de metodologías de secuenciación dirigidas a celdas de flujo.

La secuenciación utilizando los métodos y composiciones descritos en la presente memoria también puede realizarse en una placa de microvaloración, por ejemplo en placas o láminas de reacción de alta densidad (Margulies et al, 2005, Nature 437 (7057): 376-380). Por ejemplo, los objetivos genómicos pueden ser preparados por tecnologías emPCR. Las placas de reacción o láminas pueden crearse a partir de material de fibra óptica capaz de capturar y registrar la luz generada a partir de una reacción, por ejemplo a partir de una reacción fluorescente o luminiscente. El material de núcleo puede grabarse para proporcionar pocillos de reacción discretos capaces de contener al menos una perla de reacción de emPCR. Dichas láminas/placas pueden contener más de 1,6 millones de pocillos. Las láminas/placas creadas pueden cargarse con las perlas de reacción de secuenciación secuencial emPCR y montarse en un instrumento en el que se proporcionan los reactivos de secuenciación y se produce la secuenciación.

Se proporciona un ejemplo de sustratos dispuestos para la secuenciación de dianas que utilizan composiciones y métodos como se describen en la presente memoria cuando se practican sustratos modelados que comprenden nanobolas de ADN en un chip o lámina, tal como se realiza por Complete Genomics (Mountain View, CA). Como se describe en Drmanac et al., 2010, Science 327 (5961): 78-81, una oblea de silicio se puede recubrir con capas con dióxido de silicio y titanio y posteriormente ser modelado usando fotolitografía y técnicas de grabado en seco. La oblea se puede tratar con HMDS y recubrirse con una capa fotorresistente para definir regiones discretas para la silanización y posterior unión covalente de nanobolas de ADN para la secuenciación. Cualquier experto en la técnica apreciará que existen muchos métodos para crear láminas/chips con posiciones discretas para la inmovilización de ácidos nucleicos para uso en metodologías de secuenciación y los presentes métodos no están limitados por el método en el que se prepara un sustrato para la secuenciación.

Para fines de ilustración y no destinados a limitar realizaciones como se describe en la presente memoria, un ciclo de secuenciación de estrategia general puede describirse mediante una secuencia de etapas. El ejemplo siguiente se basa en una secuencia por reacción de secuenciación de síntesis, sin embargo los métodos como se describen en este documento no se limitan a ninguna metodología de reacción de secuenciación particular.

Los cuatro tipos de nucleótidos A, C, T y G, nucleótidos típicamente modificados diseñados para reacciones de secuenciación tales como nucleótidos bloqueados reversiblemente (rb) (p. ej., rbA, rbT, rbC, rbG) en los que tres de los cuatro tipos están marcados fluorescentemente, son simultáneamente añadidos junto con otros componentes de reacción, a una localización en la que se localiza la secuencia plantilla de interés y tiene lugar la reacción de secuenciación (por ejemplo, la célula de flujo, el chip, la lámina, etc.). Después de la incorporación de un nucleótido en una cadena de ácido nucleico de secuencia creciente basada en la secuencia diana, la reacción se expone a luz y se observa y registra la fluorescencia; esto constituye un primer evento de formación de imágenes y un primer patrón de detección de fluorescencia. Después del primer evento de formación de imágenes, se pueden añadir uno o más reactivos químicos adicionales a la reacción de secuenciación, por lo que el reactivo o los reactivos añadidos pueden cambiar la intensidad de la fluorescencia u otro aspecto químico de la primera reacción, lo que provoca un cambio identificable y mensurable en la fluorescencia (es decir, un cambio de transición de la fluorescencia). La ubicación de la reacción se ilumina una vez más y se captura y registra cualquier cambio en la fluorescencia; constituyendo un segundo evento de formación de imágenes (es decir, un segundo patrón de detección de fluorescencia). Los bloqueadores presentes en los nucleótidos incorporados se eliminan y se lavan junto con otros

reactivos presentes después del segundo evento de formación de imágenes en preparación para el siguiente ciclo de secuenciación. Reactivos químicos ejemplares incluyen, pero no se limitan a, reactivos de escisión, conjugados de pareja de unión-resto fluorescente u otros reactivos que pueden causar directa o indirectamente un cambio identificable y medible en la fluorescencia desde el primer evento de formación de imágenes al segundo evento de formación de imágenes. Se comparan los patrones de fluorescencia de los dos eventos de formación de imágenes y se determina la incorporación de nucleótidos y, por lo tanto, se determina la secuencia del ácido nucleico diana para ese ciclo en particular. El ciclo de estrategia general ejemplar utiliza preferiblemente un resto fluorescente (o más de uno de igual o similar excitación/emisión) y un filtro de detección de emisiones para determinar la incorporación de los cuatro tipos de nucleótidos diferentes en una reacción de secuenciación.

Una vía para diferenciar entre las diferentes estrategias para detectar la incorporación de nucleótidos en una reacción de secuenciación utilizando un colorante fluorescente (o dos o más colorantes de los mismos o similares espectros de excitación/emisión) es caracterizando las incorporaciones en términos de la presencia o ausencia relativa, o niveles intermedios, de transición de fluorescencia que se produce durante un ciclo de secuenciación. Como tal, las estrategias de secuenciación pueden ser ejemplificadas por su perfil fluorescente para un ciclo de secuenciación. Para las estrategias descritas en la presente memoria, "1" y "0" denota un estado fluorescente en el que un nucleótido está en un estado de señal (por ejemplo, detectable por fluorescencia) (1) o si un nucleótido se encuentra en estado oscuro (por ejemplo, no detectado o detectado mínimamente en una etapa de formación de imágenes) (0). Un estado "0" no se refiere necesariamente a una falta total, o ausencia de señal. Aunque en algunas realizaciones puede haber una falta total o ausencia de señal (por ejemplo, fluorescencia). También se contempla que la señal de fluorescencia mínima o disminuida (por ejemplo, la señal de fondo) se incluya en el alcance de un estado "0" mientras se pueda distinguir de manera fiable un cambio en la fluorescencia de la primera a la segunda imagen (o viceversa).

En una realización, se ejemplifica una estrategia ejemplar para detectar y determinar la incorporación de nucleótidos en una reacción de secuenciación usando un colorante fluorescente (o dos colorantes de los mismos o similares espectros de excitación/emisión) y dos eventos de formación de imágenes mediante las siguientes cuadrículas y tablas. Las cuadrículas representan la representación espacial teórica de los datos de secuenciación tal como se visualiza en el mapa de calor, o diagramas de nubes, por ejemplo, como se ve en la Figura 1.

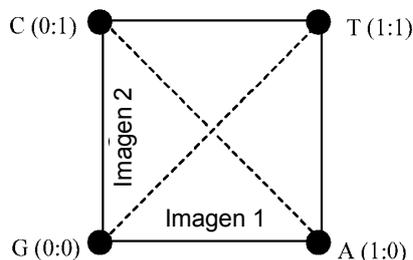


	Imagen 1	Imagen 2
A	1	0
C	0	1
G	0	0
T	1	1

En algunas realizaciones de secuenciación por síntesis (SBS), se añaden simultáneamente cuatro tipos de trifosfato de nucleótido modificado, en este caso trifosfatos de nucleótido bloqueados reversiblemente (rbNTP) a una reacción de SBS. Los rbNTPs compiten por la incorporación en la hebra de ácido nucleico en crecimiento durante la extensión dirigida por plantilla de un cebador. Se contempla que la extensión competitiva en presencia de una variedad suficiente de tipos de nucleótidos para complementar todos los tipos de nucleótidos en el ácido nucleico plantilla mejora la fidelidad de incorporación en comparación con la adición de nucleótidos uno a la vez a una reacción de secuenciación. Los cuatro tipos rbNTP poseen un terminador 3' que comprende, en la posición ribosa de la muestra 3', tanto funcionalidades alcóxido como azido que se pueden separar por escisión con un reactivo fosfina, creando de este modo un nucleótido bloqueado de forma reversible y una vez más funcional para la posterior elongación (es decir, totalmente funcional o ff). Los nucleótidos totalmente funcionales, ffNTP, están comercialmente disponibles en Illumina, Inc. y son ejemplares de nucleótidos bloqueados reversiblemente, o rbNTPs. En realizaciones preferidas, tres de los cuatro rbNTP comprenden marcadores fluorescentes unidos mediante enlazantes. Los enlazantes pueden comprender uno o más grupos de escisión, o ningún grupo de escisión. Por

ejemplo, un enlazante que asocia uno o más rbNTP a un fluoróforo puede comprender una azida y/o un grupo alcoxi, por ejemplo sobre el mismo carbono, de tal manera que los enlazantes puedan escindirse después de cada ciclo de incorporación por medio de un reactivo de fosfina como se menciona anteriormente, liberando de este modo el resto fluorescente para un alargamiento posterior de la secuencia.

- 5 Por ejemplo, la timina rbNTP inicial (rbTTP) puede marcarse fluorescentemente a través de un enlazante en el que el enlazante comprende un sitio de escisión de azida/alcoxi. Otro rbNTP inicialmente marcado fluorescentemente, por ejemplo adenina o rbATP, comprende un enlazante que además del grupo alcoxi/azida comprende además un segundo sitio de escisión como un grupo disulfuro situado entre, por ejemplo, el grupo alcoxi/azida y el marcador fluorescente. El marcador fluorescente asociado con rbATP puede ser el mismo que el marcador fluorescente asociado con rbTTP, o puede ser un marcador fluorescente similar en que comparten características de espectros de excitación y emisión similares. Un tercer rbNTP, por ejemplo citosina o rbCTP, comprende un resto hapteno, tal como una biotina, en el extremo de un enlazante que contiene alcoxi/azida. En este ejemplo, el rbCTP de partida no está marcado fluorescentemente y por lo tanto no fluoresce en un primer evento de formación de imágenes. Sin embargo, el tratamiento subsiguiente con una estreptavidina marcada fluorescentemente provoca la unión del conjugado de estreptavidina-resto fluorescente al resto de biotina en el conjugado de rbCTP y después de dicho tratamiento las posiciones en las que se incorporó rbCTP se fluorescen cuando se exponen a la longitud de onda apropiada de luz y se registra la fluorescencia durante el segundo evento de formación de imágenes. El cuarto rbNTP, en este caso guanina o rbGTP carece de un resto fluorescente y puede o no estar conjugado con un enlazante, se considera un rbNTP "oscuro" y no fluoresce, o tiene fluorescencia disminuida o mínima, en ambos eventos de formación de imágenes.

La estrategia ejemplar mencionada anteriormente se puede describir adicionalmente de acuerdo con la construcción rbNTP, por ejemplo:

rbTTP-enlazante CS1-FM

rbATP-enlazante CS1-CS2-FM

25 rbCTP-enlazante-CS1-B

rbGTP

- en donde CS1 es un primer sitio de escisión (por ejemplo, azida/alcoxi), CS2 es un segundo sitio de escisión (por ejemplo, enlace SS), FM es un resto fluorescente y B es biotina. Se contempla que uno de los sitios de escisión es opcional. Un sitio de escisión opcional (por ejemplo, dos sitios de escisión presentes en un enlazante) puede proporcionar funcionalidad adicional a un ciclo de secuenciación incluyendo, pero sin limitarse a, escisión de todos los restos fluorescentes en un ciclo subsiguiente, reacciones de escisión alternativas en ciclos subsiguientes de secuenciación y/o combinando reacciones de escisión en uno o más ciclos de secuenciación, o combinaciones de los mismos.

- Un esquema de detección ejemplar para un ciclo de secuenciación para el análisis en tiempo real de la secuencia por incorporación de nucleótidos de síntesis utilizando la estrategia antes mencionada comprende dos eventos de formación de imágenes y en realizaciones particulares no más de dos eventos de formación de imágenes. Los rbNTPs conjugados, rbTTP, rbATP y rbCTP y rbGTP no conjugados (o tal vez conjugados sólo con el enlazante) se añaden simultáneamente al comienzo de un ciclo de secuenciación. La luz de la longitud de onda de excitación para el resto fluorescente se aplica a la reacción de secuenciación y se registra una primera imagen (imagen 1). La primera imagen registra fluorescencia (1) para las incorporaciones de rbATP y rbTTP, pero ninguna fluorescencia o una fluorescencia mínima para la incorporación de rbCTP o rbGTP. Después del primer evento de formación de imágenes, se añade DTT por ejemplo a la reacción que escinde CS2 (enlace disulfuro) en el enlazante de rbATP liberando de este modo el FM y el rbATP de transición de detectable (1) a indetectable (0) para el segundo evento de formación de imágenes. La escisión de rbATP y el paso de transición fluorescente resultante proporcionan la diferenciación de rbATP de los otros eventos de incorporación de rbNTP durante un ciclo de secuenciación. Adicionalmente, después de la primera etapa de formación de imágenes, se añade a la reacción una estreptavidina (SA)-FM. La SA se une a la B de la composición rbCTP haciendo transición de rbCTP de indetectable (0) a detectable (1) y permitiendo la detección de las localizaciones en las que se incorporó rbCTP en la reacción y proporcionando la diferenciación de los eventos de incorporación de rbCTP durante un ciclo de secuenciación. En este ejemplo, no hay cambios de transición para rbTTP o rbGTP. Como tal, después de la aplicación de la DTT ejemplar y del SA-FM se toma una segunda imagen del ciclo de secuenciación que da como resultado señales fluorescentes para incorporaciones de rbTTP y rbCTP y ninguna fluorescencia para las incorporaciones de rbATP y rbGTP. Después de la segunda imagen, se utilizan las transiciones de fluorescencia, o la falta de ellas, para determinar qué nucleótido se incorporó en qué localización en la secuencia por reacción de síntesis. Cada ciclo posterior sigue el mismo patrón de extensión de polimerasa-imagen 1-tratamiento químico-imagen 2-ciclo siguiente hasta que la operación de secuenciación se haya completado. El ciclo puede incluir opcionalmente una etapa de determinación de nucleótidos. Adicionalmente o alternativamente, la determinación de nucleótidos o la secuencia de nucleótidos puede ocurrir después de que uno o más ciclos estén completos. También se pueden incluir otras

etapas por ciclo incluyendo, pero no limitándose a, desbloqueo, lavado y/o pasos adicionales usados en métodos de secuencia por síntesis conocidos en la técnica.

Se contempla que cualquier número de sitios de escisión potenciales y sus compuestos de escisión se pueden utilizar en la estrategia mencionada anteriormente, y los mencionados son sólo a modo de ejemplo. Por ejemplo, se pueden usar agentes reductores además de DTT (por ejemplo, TCEP, BME, etc.) o reactivos que participan en las reacciones de intercambio de tiol-disulfuro para liberar un resto fluorescente como se ha descrito anteriormente. Además, también pueden utilizarse parejas de unión a haptenos además de biotina-estreptavidina (por ejemplo, digoxigenina, dinitrofenol y anticuerpos de los mismos). Además, pueden utilizarse uno o más restos fluorescentes. Sin embargo, si se utilizan dos o más, es preferible que tengan los mismos o similares espectros de absorción y emisión. Las realizaciones preferidas utilizan un resto fluorescente para la detección de todos los nucleótidos incorporados, o un filtro óptico que detecta la emisión de una pluralidad de restos fluorescentes.

Se contempla que los reactivos de reacción (es decir, reactivos de escisión, reactivos de marcado, etc.) añadidos entre los eventos de formación de imágenes se pueden proporcionar por separado, por ejemplo secuencialmente o combinados y añadidos como un reactivo completo (por ejemplo, una mezcla maestra que comprenda todos los productos químicos necesarios para la escisión completa, el marcado, etc.). Las realizaciones preferidas comprenden la adición de una solución reactiva completa o mezcla maestra entre las etapas de formación de imágenes.

En otra realización a modo de ejemplo, se ejemplifica una segunda estrategia para detectar y determinar la incorporación de nucleótidos en una reacción de secuenciación utilizando un colorante fluorescente (o dos colorantes de los mismos o similares espectros de excitación/emisión) y dos eventos de formación de imágenes por la siguiente tabla y cuadrícula de detección.

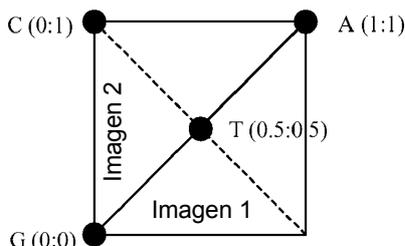


	Imagen 1	Imagen 2
A	1	1
C	0	1
G	0	0
T	0,5	0,5

Para la segunda estrategia, como se ejemplifica en la primera, se añaden simultáneamente las cuatro composiciones de nucleótido trifosfato (rbNTPs) completamente funcionales a una reacción SBS. Los rbNTPs compiten por su incorporación en la hebra de ácido nucleico en crecimiento. Los rbNTPs poseen un terminador 3' que comprende tanto funcionalidades alcoxi como azido que son removibles por escisión con un reactivo de fosfina creando de este modo un nucleótido que es una vez más funcional para un alargamiento adicional. En realizaciones preferidas, tres de los cuatro rbNTP comprenden marcadores fluorescentes unidos mediante enlazantes. Los enlazantes pueden comprender uno o más sitios de escisión. Por ejemplo, un enlazante que une uno o más rbNTP a un fluoróforo puede comprender una azida y/o un grupo alcoxi, por ejemplo sobre el mismo carbono, de modo que los enlazantes puedan ser escindidos después de cada ciclo de incorporación por medio de un reactivo de fosfina liberando así el resto fluorescente para la prolongación de la secuencia adicional.

En la segunda estrategia, el grupo inicial de rbNTP típicamente comprende una mezcla de moléculas de rbTTP. Por ejemplo, un grupo de rbTTP comprende una relación 2:1 de un rbTTP marcado con fluorescencia (es decir, a través de un enlazante) y rbTTP marcado no fluorescentemente (es decir, rbTTP oscuro). Se contempla que puede usarse cualquier relación de rbNTP fluorescente:no fluorescente. Por ejemplo, también funcionarán relaciones de 2:1, 1:0,5, 0,5:1 y 1:2, cuya diferencia cambiaría la salida de intensidad de la imagen sin cambiar la capacidad de detectar y diferenciar la incorporación de nucleótidos. Un rbATP marcado fluorescentemente, un rbGTP no marcado u oscuro y un rbCTP marcado con biotina completan la mezcla de nucleótidos. Un tratamiento subsiguiente con una

estreptavidina marcada fluorescentemente provoca la unión de estreptavidina-resto fluorescente al resto biotina en el conjugado rbCTP y después de dicho tratamiento las posiciones en las que se incorporó rbCTP se fluorescen cuando se exponen a la longitud de onda apropiada de luz y la fluorescencia se registra durante el segundo evento de formación de imagen.

5 La estrategia ejemplar mencionada anteriormente puede comprender las construcciones rbNTP:

rbTTP-enlazante FM/rbTTP-oscuro

rbATP-enlazante-FM

rbCTP-enlazante-B

rbGTP-oscuro

10 Un esquema de detección ejemplar para un ciclo de secuenciación para el análisis en tiempo real de la secuencia por incorporación de nucleótidos de síntesis utilizando la estrategia antes mencionada comprende dos eventos de formación de imágenes y en realizaciones particulares no más de dos eventos de formación de imágenes. Los cuatro tipos de rbNTP se añaden simultáneamente al comienzo de un ciclo de secuenciación. La luz de la longitud de onda de excitación para el resto fluorescente se aplica a la reacción de secuenciación y se registra una primera imagen (imagen 1). La primera imagen incluye fluorescencia (1) para incorporaciones de rbATP y rbTTP (a 50% de intensidad de fluorescencia), pero ninguna fluorescencia para rbCTP, rbGTP y 1/2 de las incorporaciones de rbTTP. Después de la primera etapa de formación de imágenes, se añade a la reacción un fluoróforo SA-FM marcado con estreptavidina. La SA se une a la B de la composición rbCTP haciendo transición de rbCTP de indetectable (0) a detectable (1) durante el segundo evento de formación de imágenes y permitiendo la detección de localizaciones donde rbCTP se incorporó en la reacción y proporcionando la diferenciación de eventos de incorporación de rbCTP durante un ciclo de secuenciación. En este ejemplo, no hay cambios de transición para rbTTP, rbATP o rbGTP. Después de la segunda imagen, se utilizan las transiciones de fluorescencia, o la falta de ellas, para determinar qué nucleótido se incorporó en qué localización en la secuencia por reacción de síntesis y se identifica la secuencia de interés. Cada ciclo subsiguiente sigue el mismo patrón de prolongación de polimerasa-imagen 1-tratamiento-imagen 2-ciclo siguiente hasta que la secuenciación total del objetivo deseado esté completa. El ciclo puede incluir opcionalmente una etapa de determinación de nucleótidos. Adicionalmente o alternativamente, la determinación de nucleótidos o la secuencia de nucleótidos puede ocurrir después de que uno o más ciclos estén completos. También se pueden incluir otras etapas por ciclo incluyendo, pero no limitándose a, desbloqueo, lavado y/o pasos adicionales usados en los métodos de secuencia por síntesis conocidos en la técnica.

30 En otra realización, se ejemplifica una tercera estrategia para detectar y determinar la incorporación de nucleótidos en una reacción de secuenciación utilizando un colorante fluorescente (o dos colorantes de los mismos o similares espectros de excitación/emisión) y dos etapas de formación de imágenes mediante la siguiente tabla y cuadrícula de detección.

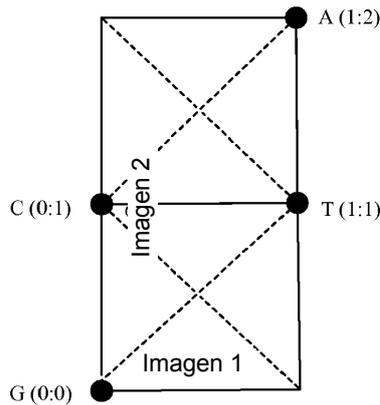


	Imagen 1	Imagen 2
A	1	2
C	0	1
G	0	0

	Imagen 1	Imagen 2
T	1	1

Como se ejemplifica en la primera y segunda, las cuatro composiciones de nucleótido trifosfato (rbNTPs) totalmente funcionales se añaden simultáneamente a una reacción SBS. Los rbNTPs compiten por su incorporación en la hebra de ácido nucleico en crecimiento. Los rbNTPs poseen un terminador 3' que es removible, creando de este modo un nucleótido que es una vez más funcional para el alargamiento posterior. La tercera estrategia difiere de las estrategias ejemplares anteriores al incorporar, por ejemplo, por conjugado un rbNTP a un enlazante ramificado. En realizaciones preferidas, dos de los cuatro rbNTP comprenden marcadores fluorescentes unidos mediante enlazantes. Los enlazantes pueden comprender uno o más sitios de escisión. Por ejemplo, un enlazante que asocia uno o más rbNTP a un fluoróforo puede comprender una azida y/o un grupo alcoxi, por ejemplo sobre el mismo carbono, de tal manera que los enlazantes puedan escindirse después de cada ciclo de incorporación por medio de un reactivo de fosfina como se describe anteriormente, liberando de este modo el resto fluorescente para la prolongación posterior de la secuencia.

En la tercera estrategia ejemplar, los complejos rbATP y rbCTP comprenden enlazantes ramificados. Por ejemplo, rbATP comprende un enlazante ramificado en el que una rama termina con un resto fluorescente y una segunda rama termina en una biotina. En este ejemplo, el rbCTP está también complejado con un enlazante ramificado y cada una de dos ramas termina en una biotina. El rbCTP en este ejemplo está inicialmente sin marcar. Un rbTTP marcado con fluorescencia y un rbGTP sin marcar u oscuro completan la mezcla de nucleótidos. Un tratamiento subsiguiente con una estreptavidina marcada fluorescentemente provoca una unión muy fuerte del estreptavidina-colorante a los restos de biotina en los nucleótidos C y A y después de tal tratamiento las posiciones donde rbCTP y rbATP se incorporaron fluorescen cuando se exponen a la longitud de onda apropiada de luz y la fluorescencia resultante de la interacción B-SA se registra durante la segunda etapa de formación de imágenes.

La estrategia ejemplar mencionada anteriormente como tal puede comprender:

- rbATP-enlazante FM y B ramificado
- rbTTP-FM
- rbCTP-enlazante ramificado-(B)²
- rbGTP-oscuro

Un esquema de detección ejemplar para un ciclo de secuenciación para el análisis en tiempo real de la secuencia por incorporación de nucleótidos de síntesis utilizando la estrategia antes mencionada comprende dos etapas de formación de imágenes y en realizaciones particulares no más de dos eventos de formación de imágenes. Todos los rbNTPs se añaden simultáneamente al comienzo de un ciclo de secuenciación. La luz de la longitud de onda de excitación para el resto fluorescente se aplica a la reacción de secuenciación y se registra una primera imagen (imagen 1). La primera imagen incluye fluorescencia (1) para las incorporaciones de rbATP y rbTTP, pero ninguna fluorescencia (0) para las incorporaciones de rbCTP y rbGTP. Después de la primera etapa de formación de imágenes, se añade a la reacción un fluoróforo SA-FM marcado con estreptavidina. El SA une las dos biotinas (B²) del conjugado de rbCTP, haciendo la transición de rbCTP de indetectable (0) a detectable (1) y la B en el enlazante bifurcado de rbATP, aumentando de este modo con eficacia la fluorescencia (2) de la incorporación de rbATP de la imagen 1 y permitiendo la detección de localizaciones en las que rbCTP fue incorporado y diferenciando la incorporación de rbATP, en la hebra de ácido nucleico en crecimiento. En este ejemplo, no hay cambios de transición para rbTTP o rbGTP. Después de la segunda imagen, se utilizan las transiciones de fluorescencia, o la falta de ellas, para determinar qué nucleótido se incorporó en qué localización en la secuencia por la reacción de síntesis y se identifica la secuencia de interés. Cada ciclo subsiguiente sigue el mismo patrón de extensión de polimerasa-imagen 1-tratamiento-imagen 2-ciclo siguiente hasta que se completa la secuenciación del objetivo deseado. El ciclo puede incluir opcionalmente una etapa de determinación de nucleótidos. Adicionalmente o alternativamente, la determinación de nucleótidos o la secuencia de nucleótidos puede ocurrir después de que uno o más ciclos estén completos. Se pueden incluir también otras etapas por ciclo que incluyen, pero no se limitan a, desbloqueo, lavado y/u otras etapas utilizadas en métodos de secuencia por síntesis conocidos en la técnica.

En otra realización, una cuarta estrategia ejemplar para detectar y determinar la incorporación de nucleótidos en una reacción de secuenciación usa un colorante fluorescente (o dos colorantes de los mismos o similares espectros de excitación/emisión) y sólo una etapa de formación de imágenes como se ejemplifica mediante la siguiente tabla y cuadrícula de detección.

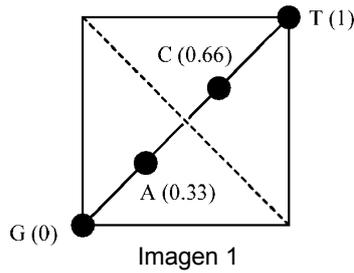


	Imagen 1
A	0,33
C	0,66
G	0
T	1,0

5 La realización ejemplar mencionada anteriormente puede comprender solamente un colorante, o dos colorantes del mismo o similares espectros de excitación/emisión, en los que la concentración de colorante cambia para cada uno de los tres rbNTP marcados. Un estado oscuro indica la incorporación de, en este caso, rbGTP basado en la interpretación de la medición de fluorescencia de los tres rbNTP marcados fluorescentemente.

La estrategia ejemplar mencionada anteriormente como tal puede comprender:

- rbATP-FM (concentración 0,33X)
- rbTTP-FM (concentración 1,0X)
- 10 rbCTP-FM (concentración 0,66X)
- rbGTP-oscuro

Un método alternativo descrito en la presente invención comprende un colorante (o dos colorantes del mismo espectro de excitación/emisión o similares) y un evento de imagen es el siguiente:

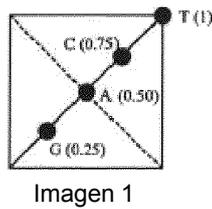


	Imagen 1
A	0,50
C	0,75
G	0,25
T	1,0

15

La estrategia ejemplar mencionada anteriormente como tal puede comprender:

- rbATP-FM (concentración 0,50X)

rbTTP-FM (concentración 1,0X)

rbCTP-FM (concentración 0,75X)

rbGTP-FM (concentración 0,25X)

5 El método mencionado anteriormente puede comprender un colorante o dos colorantes de espectros de excitación/emisión similares, de modo que cada uno de los cuatro rbNTP se marquen con diferentes concentraciones de colorante. En los métodos en los que cada uno de cuatro rbNTP diferentes están unidos a una concentración diferente de un colorante (o dos colorantes de espectros de excitación/emisión similares) sólo se toma una imagen por ciclo para determinar la incorporación de nucleótidos. Un ciclo de secuencia ejemplar que practica métodos de eventos de un colorante/una imagen sería la extensión de polimerasa-imagen 1-siguiente ciclo.

10 Un esquema de detección ejemplar para un ciclo de secuenciación para una secuencia de eventos de un colorante/una imagen por incorporación de nucleótidos de síntesis utilizando las estrategias antes mencionadas comprende una etapa de formación de imágenes. Todos los rbNTP se añaden simultáneamente al comienzo de un ciclo de secuenciación. La luz de la longitud de onda de excitación para el resto fluorescente se aplica a la reacción de secuenciación y se registra una primera imagen (imagen 1). Después de la primera etapa de formación de imágenes, se lleva a cabo el siguiente ciclo de adición de reactivo, extensión de polimerasa y adquisición de imagen hasta completar el número deseado de ciclos. Después de la primera imagen, la intensidad de fluorescencia puede correlacionarse con las diferentes concentraciones de colorante que se utilizan para determinar qué nucleótido se incorporó en qué posición en la secuencia por reacción de síntesis y se identifica la secuencia de interés. Cada ciclo subsiguiente sigue el mismo patrón de extensión de polimerasa-imagen 1-ciclo siguiente hasta que se completa la secuenciación del objetivo deseado. El ciclo puede incluir opcionalmente una etapa de determinación de nucleótidos. Adicionalmente o alternativamente, la determinación de nucleótidos o la secuencia de nucleótidos puede ocurrir después de que uno o más ciclos estén completos. Se pueden incluir también otras etapas por ciclo que incluyen, pero no se limitan a, desbloqueo, lavado y/u otras etapas utilizadas en métodos de secuencia por síntesis conocidos en la técnica.

25 En realizaciones que practican una secuenciación de eventos de un colorante/una imagen, se proporcionan concentraciones de colorantes que permiten la diferenciación de la incorporación de los nucleótidos marcados y/o no marcados. Además, cuando se practica una reacción de secuenciación de un colorante/una imagen como se ha ejemplificado anteriormente, no es necesario un tratamiento químico adicional como se describió anteriormente para realizaciones para una estrategia de secuenciación de un colorante/dos imágenes.

30 En otra realización ejemplar, una estrategia adicional para detectar y determinar la incorporación de nucleótidos en una reacción de secuenciación comprende el uso de dos colorantes fluorescentes de diferentes espectros de excitación y emisión y 1) un evento de formación de imágenes que comprende dos espectros de emisión o 2) dos eventos de formación de imágenes secuenciales.

35 Para los propósitos del ejemplo, la Figura 3A muestra dos colorantes ejemplares, un colorante que emite aproximadamente a 590 $\lambda_{\text{Máximo}}$ (DEG527) y un colorante que emite aproximadamente a 720 $\lambda_{\text{Máximo}}$ (Dy681). A modo de ejemplo, se realizan las siguientes conjugaciones de rbNTP-colorante:

rbATP-DEG527

rbCTP-Dy681

rbTTP-DEG527/Dy681

40 RbGTP-oscuro

Como tal, el porcentaje de cada nucleótido incorporado conjugado a un fluoróforo particular. En este ejemplo las conjugaciones son:

RbNTP	Dy681	DEG527	Oscuro
rbatp		100%	
rbCTP	100%		
rbTTP	50%	50%	
rbGTP			100%

Como ejemplo, siguiendo protocolos estándar de SBS, los cuatro nucleótidos se añaden simultáneamente a una reacción SBS. Los rbNTPs compiten por su incorporación en la hebra de ácido nucleico en crecimiento. Como se ha descrito anteriormente, los rbNTP poseen un terminador 3' que es removible por escisión para la elongación posterior. Después de la incubación que permite la incorporación del nucleótido apropiado en la hebra de ácido nucleico en crecimiento, la reacción se expone a la longitud de onda apropiada de luz dependiendo de qué imagen, simultánea o secuencial, se desea. Por ejemplo, la reacción puede exponerse simultáneamente a la longitud de onda de excitación de ambos colorantes fluorescentes (en este ejemplo, se excita el DEG527 a aproximadamente 532 nm y se excita el Dy681 a aproximadamente 660 nm) provocando de este modo la emisión simultánea de los dos colorantes fluorescentes, cuya emisión puede ser detectada simultáneamente por dos diferentes filtros de detección y óptica de imagen. En dicho sistema simultáneo en el que sólo se realiza 1 evento de formación de imagen para dos canales de detección diferentes simultáneamente, los estados de imagen para cada canal de detección serían:

	Imagen 1-verde	Imagen 1-rojo
A	1	0
C	0	1
T	0,5	0,5
G	0	0

Alternativamente, después de la incorporación del nucleótido marcado apropiado en la hebra de ácido nucleico en crecimiento, los dos colorantes fluorescentes pueden excitarse de una manera paso a paso, tal como primero excitar un fluoróforo seguido de un primer evento de formación de imágenes y luego excitar el segundo fluoróforo seguido de un segundo evento de formación de imágenes. En dicho sistema de formación de imágenes paso a paso se llevan a cabo dos eventos de formación de imágenes y la tabla de detección sería, por ejemplo, si DEG527 se excita primero seguido de Dy681 (por ejemplo, viceversa si la emisión de fluorescencia roja es la primera imagen seguida de la fluorescencia verde) los estados de imagen para cada evento de imagen serían:

	Imagen 1-verde	Imagen 2-rojo
A	1	0
C		1
T	0,5	0,5
G	0	0

En cualquier caso, la incorporación de una A sería detectada con una cierta intensidad sólo en el canal verde, la incorporación de una C sería detectada con una cierta intensidad en el canal rojo solamente, la incorporación de una T sería detectada en ambos canales verde y rojo a la mitad de la intensidad de A y C, y G sería mínimamente detectado o no detectado en canales verdes o rojos (Figura 3B). Después la formación de imágenes etapa por etapa, el colorante fluorescente y el terminador 3' se escinden y se lleva a cabo el siguiente ciclo de secuenciación.

Este ejemplo no se limita a dos colorantes o combinaciones de conjugados particulares y se podrían usar dos colorantes de diferentes espectros de fluorescencia en un sistema de secuenciación de dos colorantes, en cualquier combinación de combinación de conjugado de rbNTP-colorante. Por ejemplo, los colorantes representados en el ejemplo anterior emitidos en las longitudes de onda rojas y verdes. Sin embargo, los métodos y sistemas no están limitados por las longitudes de onda de excitación o emisión (por ejemplo, espectros de fluorescencia) de cualquier colorante particular, ya que cualquier colorante que difiera en los espectros de fluorescencia puede ser potencialmente útil. Además, el ejemplo describe ciertos conjugados de rbNTP-colorante; sin embargo los conjugados no están limitados a esas combinaciones particulares. Por ejemplo, cualquiera de los tres rbNTP podría estar potencialmente conjugado con cualquiera de los colorantes listados (un nucleótido que permanece no conjugado u oscuro). Ejemplos de colorantes y derivados de los mismos útiles en las realizaciones descritas en este documento incluyen, pero no se limitan a, los descritos a continuación.

Adicionalmente, uno o más conjugados de tipo nucleótido descritos en la estrategia anterior podrían comprender además uno o más enlazantes como se describe en realizaciones y estrategias alternativas. Como tal, una o más reacciones químicas o modificadoras podrían incorporarse en una reacción de secuenciación en combinación con la estrategia en la que dos colorantes de diferentes espectros de fluorescencia se conjugan a diferentes tipos de nucleótidos. Por lo tanto, los conjugados de tipo nucleótido en este ejemplo podrían modificarse adicionalmente de cualquier número de maneras como se describe en la presente memoria sin desvirtuar la realización en la que se pueden emplear dos colorantes de diferentes espectros de fluorescencia para determinar la secuencia de un ácido nucleico.

En otra realización ejemplar, una estrategia adicional para detectar y determinar la incorporación de nucleótidos en una reacción de secuenciación comprende usar dos conjuntos de colorantes fluorescentes en los que cada conjunto de colorantes comprende dos colorantes de espectros de emisión de fluorescencia similares o con emisión $\lambda_{\text{Máximo}}$ desplazado hasta, por ejemplo, 100 nm, en el que uno de los dos colorantes emite a una intensidad detectablemente mayor que el otro colorante en el conjunto y en el que los dos conjuntos de colorantes fluorescentes difieren en los espectros de emisión de fluorescencia. En realizaciones preferidas, el colorante en un conjunto de colorantes que es detectablemente de mayor intensidad que el otro colorante en el conjunto es al menos 0,5X, al menos 0,75X, al menos 1X, al menos 2X tan intenso como el colorante de intensidad más baja. Cuando se practican dos conjuntos de colorantes fluorescentes como se describe en el presente documento, la determinación de la secuencia puede ser a través de un evento de formación de imágenes o dos eventos de formación de imágenes.

Para los propósitos del ejemplo, la Figura 4A muestra dos conjuntos de colorantes ejemplares; tanto DEG527 como Atto532 pueden detectarse juntos (emisión de fluorescencia desde aproximadamente $\lambda_{\text{Máximo}}$ 555-595nm) y Dy681 y SO7181 pueden detectarse juntos (emisión de fluorescencia de aproximadamente $\lambda_{\text{Máximo}}$ 670-715 nm). A modo de ejemplo, se realizan las siguientes conjugaciones de rbNTP-colorante:

rbATP-DEG527

rbCTP-Dy681

rbTTP-Atto532/SO7181

RbGTP-oscurο

Como tal, el porcentaje de cada nucleótido incorporado conjugado a un fluoróforo particular en este ejemplo es:

rbNTP	S07181	Dy681	Atto532	DEG527	Oscuro
rbatp				100%	
rbCTP		100%			
rbTTP	50%		50%		
rbGTP					100%

Como ejemplo, siguiendo protocolos estándar de SBS, los cuatro nucleótidos se añaden simultáneamente a una reacción SBS. Los rbNTP compiten por su incorporación en la hebra de ácido nucleico en crecimiento. Como se ha descrito anteriormente, los rbNTP poseen un terminador 3' que es removible por escisión para la elongación posterior. Después de la incubación que permite la incorporación del nucleótido apropiado en la hebra de ácido nucleico en crecimiento, la reacción se expone a una primera longitud de onda de luz, se realiza un primer evento de formación de imágenes, entonces la reacción se expone a la segunda longitud de onda de luz y se realiza un segundo evento de formación de imágenes.

Por ejemplo, después de la incorporación del nucleótido marcado apropiado en la hebra de ácido nucleico en crecimiento, los dos conjuntos de colorantes fluorescentes pueden excitarse de una manera paso a paso, tal como primero excitar un conjunto de fluoróforos seguido de un primer evento de formación de imágenes y luego excitar el segundo conjunto de fluoróforos seguido por un segundo evento de formación de imágenes. Como ejemplo, si DEG527/Atto532 es excitado primero seguido por Dy681/SO7181 (por ejemplo, viceversa si la emisión de fluorescencia roja es la primera imagen seguida de fluorescencia verde) los estados de imagen para cada evento de imagen serían:

Imagen 1-verde	Imagen 2-rojo
----------------	---------------

	Imagen 1-verde	Imagen 2-rojo
A	1	0
C	0	1
T	>1	>1
G	0	0

Los estados de imagen para T se enumeran como > 1 para cada evento de imagen. La designación de > 1 supone que el colorante de intensidad más alta es al menos mayor en intensidad que el del colorante de intensidad más baja en el par de colorantes.

- 5 Alternativamente, la reacción puede exponerse simultáneamente a la longitud de onda de excitación de ambos colorantes fluorescentes, provocando de este modo la emisión simultánea de los dos colorantes fluorescentes, cuya emisión puede ser detectada simultáneamente por dos filtros de detección y óptica de imagen diferentes. En dicho sistema simultáneo en el que sólo se realiza 1 evento de formación de imagen para dos canales de detección diferentes simultáneamente, los estados de imagen para cada canal de detección serían:

	Imagen 1-verde	Imagen 1-rojo
A	1	0
C	0	1
T	>1	>1
G	0	0

10

En cualquier caso, la incorporación de una A sería detectada con una cierta intensidad sólo en el canal verde y la incorporación de una C se detectaría con una cierta intensidad sólo en el canal rojo. Sin embargo, debido al aumento de la intensidad de los colorantes que se conjugan con rbTTP en comparación con los colorantes de intensidad más baja conjugados con el rbATP y el rbCTP, se contempla que la incorporación de una T se detectará tanto en los canales verdes como en los rojos a igual o mayor intensidad de A y C. Una vez más, en este ejemplo la incorporación de G sería mínimamente detectada o no detectada en los canales verdes o rojos.

15

La Figura 4B muestra un mapa de nubes de calor que demuestra la detección de rbTTP incorporado en comparación con rbCTP y rbATP cuando se practican los dos conjuntos de colorantes descritos, en donde un colorante es de intensidad mayor que el otro colorante en el conjunto. Después de las etapas de formación de imágenes, el colorante fluorescente y el terminador 3' se escinden y se lleva a cabo el siguiente ciclo de secuenciación.

20

Adicionalmente, uno o más conjugados de tipo nucleótido descritos en este ejemplo podrían comprender además uno o más enlazantes como se describe en realizaciones alternativas. Como tal, se podrían incorporar una o más reacciones químicas o modificadoras en una reacción de secuenciación en combinación con la estrategia en la que dos conjuntos de colorantes de diferentes espectros de emisión se conjugan a diferentes tipos de nucleótidos. Por lo tanto, los conjugados de tipo nucleótido en este ejemplo podrían modificarse adicionalmente de cualquier número de formas como se describe en la presente memoria sin desvirtuar la realización en la que se pueden emplear dos conjuntos de colorantes de diferentes espectros de emisión para determinar la secuencia de un ácido nucleico.

25

Además, este ejemplo no se limita a dos conjuntos de colorantes o combinaciones de conjugados particulares y podrían usarse cualesquiera dos conjuntos de colorantes de diferentes espectros de emisión, en cualquier combinación de combinación de conjugado de rbNTP-colorante mientras se sigue la estrategia de conjugación como se describe en la presente memoria (por ejemplo, dos tipos de nucleótidos se conjugan a diferentes colorantes de menor intensidad y un tipo de nucleótido se conjuga con dos colorantes de intensidad más alta). El ejemplo que describe el uso de dos conjuntos de colorantes para métodos de secuenciación no está limitado a ningún conjunto particular de dos colorantes y se pueden usar conjuntos de colorantes de diferentes espectros de fluorescencia en el sistema de secuenciación como se describe en la presente memoria. Los conjuntos de colorantes adicionales comprenden aquellos que tienen un desplazamiento de $\lambda_{m\acute{a}x}$ de emisión de al menos 60 nm, al menos 70 nm, al

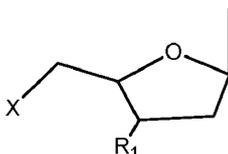
35

menos 80 nm, al menos 90 nm, al menos 100 nm, preferiblemente al menos 100 nm. Ejemplos de conjuntos de colorantes incluyen, pero no se limitan a, Atto465, 488, 495/Atto514, 520, 532, 550, 565; Atto 520, 532, 550/Atto565, 590, 594, Rho11, Rho 12, Rho 13; Atto 647, 655, 665/Atto 680, 700, 725; Alexa 647, 660, Cy5/Alexa 680, 700, Cy5.5; Alexa532, Cy3/Alexa555, 556, 578, 590, Cy3,5; Alexa 488/Alexa532, 555, 556, 578; Dy 647, 648, 649, 650, 651, 652, 654/Dy675, 676, 677, 678, 679, 680, 681, 682, 700, 701, 703, 704; Dy490, 495, 505/Dy530, 547, 548, 549, 550, 554, 555, 556, 560; Dy530, 547, 548, 549, 550, 554, 555, 556, 560/Dy590, 591, 594, 605, 610, 615.

Las estrategias anteriores son de naturaleza ejemplar, describen sólo varias de muchas estrategias potenciales y sirven para proporcionar una guía subyacente a los métodos y composiciones innovadores descritos en la presente memoria para utilizar un resto fluorescente, o una pluralidad de restos fluorescentes de los mismos o similares espectros de excitación/emisión, para la secuenciación de un ácido nucleico. Cualquier experto en la materia comprenderá que las diferentes estrategias proporcionan una guía para crear estrategias adicionales que utilizan un resto fluorescente o una pluralidad de restos fluorescentes del mismo espectro de excitación/emisión o similar, para secuenciar un ácido nucleico y estar todavía dentro del alcance de los métodos descritos en la presente memoria.

En una realización, un conjugado de rbNTP como se describe en la presente invención comprende

la detección del resto o restos  enlazante(s)  base



en el que un resto de detección es uno o más de un resto fluorescente, un hapteno o combinaciones de los mismos, en el que un enlazante es uno o más de un enlazante espaciador, un enlazante con uno o más sitios de escisión, o combinaciones de los mismos, en donde una base es una de tres nucleótidos modificados (por ejemplo, rbNTPs) en donde X es un monofosfato, difosfato o trifosfato y en donde R₁ es -H, -OH, -OCH₂N₃ o cualquier grupo que puede ser transformado en un -OH, incluyendo carbonilo unido covalentemente al carbono 3'.

En algunas realizaciones, un resto de detección es un resto fluorescente. En algunas realizaciones, un resto de detección es un hapteno que es detectable a través de un conjugado de pareja de unión-resto fluorescente. En algunas realizaciones, un conjugado de rbNTP comprende uno o ambos de un resto fluorescente y un hapteno unido a un rbNTP a través de uno o más enlazantes. En algunas realizaciones, un hapteno es una biotina, digoxigenina (DIG) o dinitrofenol (DNP). En algunas realizaciones, se detecta un hapteno mediante un conjugado de pareja de unión-resto fluorescente. En algunas realizaciones, la pareja de unión es una molécula pequeña o un anticuerpo o fragmento del mismo, por ejemplo estreptavidina, anti-DIG o anti DNP.

Ejemplos de restos fluorescentes, o derivados de los mismos, para uso como restos fluorescentes de acuerdo con las realizaciones descritas incluyen, pero no se limitan a, fluoresceína y derivados de fluoresceína tales como carboxifluoresceína, tetraclorofluoresceína, hexaclorofluoresceína, carboxinaftofluoresceína, isotiocianato de fluoresceína, NHS-fluoresceína, iodoacetamidofluoresceína, maleimida de fluoresceína, SAMSA-fluoresceína, fluoresceína tiosemicarbazida, carbohidrazinometiltioacetil-amino fluoresceína, rodamina y derivados de rodamina tales como TRITC, TMR, lisamina rodamina, Texas Red, rodamina B, rodamina 6G, rodamina 10, NHS-rodamina, TMR-iodoacetamida, cloruro de sulfonilo de lisamina rodamina B, sulfonil hidracina de lisamina rodamina B, cloruro de sulfonilo de Texas Red, hidrazida de Texas Red, coumarina y derivados cumarínicos tales como AMCA, AMCA-NHS, AMCA-sulfo-NHS, AMCA-HPDP, DCIA, AMCE-hidrazida, BODIPY y derivados tales como BODIPY FL C3-SE, BODIPY 530/550 C3, BODIPY 530/550 C3-SE, BODIPY 530/550 C3 hidrazida, BODIPY 493/503 C3 hidrazida, BODIPY FL C3 hidrazida, BODIPY FL IA, BODIPY 530/551 IA, Br-BODIPY 493/503, Cascade Blue y derivados tales como Cascade Blue acetil azida, Cascade Blue cadaverina, Cascade Blue etilendiamina, hidracida Cascade Blue, Lucifer Yellow y derivados tales como Lucifer Yellow yodoacetamida, Lucifer Yellow CH, cianina y derivados tales como colorantes de cianina basados en indolio, colorantes de cianina basados en benzo-indolio, colorantes de cianina basados en piridio, colorantes de cianina basados en tiozolio, colorantes de cianina basados en quinolinio, colorantes de cianina basados en imidazolio, Cy3, Cy5, quelatos de lantánidos y derivados tales como BCPDA, TBP, TMT, BHHCT, BCOT, quelatos de Europio, quelatos de Terbium, colorantes Alexa Fluor, colorantes DyLight, colorantes Atto, colorantes LightCycler Red, colorantes CAL Flour, JOE y sus derivados, colorantes de Oregon Green, colorantes WellRED, colorantes IRD, colorantes de ficoeritrina y ficobilina, estilbeno, colorantes de DEG (por ejemplo como los descritos en el documento de EE.UU. 2010/0009353), Colorantes NR, colorantes cercanos al infrarrojo y otros conocidos en la técnica tales como los descritos en Haugland, Molecular Probes Handbook, (Eugene, OR) 6^a Edición; el catálogo de Synthegen (Houston, TX.), Lakowicz, Principles of Fluorescence Spectroscopy, 2^a edición, Plenum Press New York (1999), Hermanson, Bioconjugate Techniques, 2^a edición, documento de EE.UU 2010/0009353 o el documento WO 98/59066.

En algunas realizaciones, un resto de detección se conjuga a un rbNTP a través de un enlazante. En algunas realizaciones, un conjugado de rbNTP comprende uno o más de un enlazante. En algunas realizaciones, un enlazante es un enlazante espaciador que está conjugado en un extremo a un rbNTP y en el otro a un resto de

detección. En algunas realizaciones, un enlazante espaciador comprende uno o más grupos de escisión. Por el contrario, en algunas realizaciones un conector espaciador no contiene ningún grupo de escisión. En una realización, un enlazador espaciador (por ejemplo, con o sin un grupo de escisión) es una molécula de polietilenglicol (PEG) o sus concatámeros. Por ejemplo, en algunas realizaciones, un enlazante espaciador comprende concatámeros de al menos dos, de al menos tres, de al menos cuatro, de al menos cinco, de al menos seis, de al menos siete, de al menos ocho, de al menos diez o de al menos doce moléculas de PEG.

En realizaciones preferidas, los enlazantes espaciadores utilizados para conjugar un rbNTP con un resto de detección, por ejemplo un resto fluorescente o un hapteno, comprenden al menos de cuatro a doce concatámeros de PEG (es decir, PEG4, PEG 8, PEG 12). En algunas realizaciones, un enlazante espaciador comprende ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético. En algunas realizaciones, el enlazante espaciador que comprende ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético comprende uno o más grupos de escisión. En algunas realizaciones, un rbNTP está unido a dos enlazantes espaciadores (por ejemplo, enlazantes separados de un constructo enlazante bifurcado), que pueden ser iguales o diferentes, cada uno de los cuales termina en un resto de detección. En algunas realizaciones, dos enlazantes espaciadores comprenden un PEG y un enlazador de ácido 2-{2-[3-(2-amino-etilcarbomil)-fenoxil-1-azido-etoxi]-etoxi-acético, uno o ambos de los cuales pueden o no comprender uno o más grupos de escisión, que terminan en un resto de detección. En algunas realizaciones, dos enlazantes espaciadores pueden ser dos enlazantes PEG que pueden ser de longitudes iguales o desiguales (por ejemplo, un PEG4 y el otro PEG12), cada uno de los enlazantes PEG terminando en un resto de detección, además con o sin un grupo de escisión.

Se pueden encontrar ejemplos de enlazantes en, por ejemplo, las patentes de Estados Unidos 7.816.503, 7.771.973, y la solicitud de patente 2010/00317531). Los métodos y composiciones como se describen en este documento no están limitados por ningún enlazante espaciador particular y las alternativas serán fácilmente evidentes para cualquier experto en la técnica y se consideran dentro del alcance de la presente descripción.

En algunas realizaciones, un enlazante espaciador comprende uno o más grupos de escisión. Los grupos de escisión para uso en los métodos como se describen en este documento pueden incluir, pero no se limitan a, grupos disulfuro, grupos de ácido lábil, grupos Sieber, grupos indol, grupos t-butilo Sieber, grupos escindibles por electrofilia, grupos escindibles por nucleófilo, grupos fotoescindibles, grupos de escisión que se escinden en condiciones reductoras, condiciones oxidativas, escisión mediante el uso de grupos de captura de seguridad, escisión por mecanismo de eliminación y grupos escindibles asistidos por metal. Tal como se utiliza en este documento, la expresión "enlazante escindible" se considera equivalente a un enlazante espaciador que comprende uno o más grupos de escisión. Se puede encontrar una discusión de los enlazantes en, por ejemplo, Guiller et al, 2000, Chem. Rev. 100: 2091-2157 y según lo dispuesto en la patente de EE.UU. 7.771.973. Los métodos y composiciones como se describen en la presente memoria descriptiva no están limitados por ningún grupo de división particular y las alternativas serán fácilmente evidentes para cualquier experto en la técnica y se consideran dentro del alcance de la presente descripción.

En algunas realizaciones, los nucleótidos modificados bloqueados reversiblemente como se describe en la presente memoria se unen a una molécula pequeña a través de un enlazante. En algunas realizaciones, el enlazante comprende uno o más grupos escindibles y puede ser referido como un enlazante escindible. Los grupos escindibles incluyen, pero no se limitan a, disulfuro, diol, diazo, éster, sulfona azida, éteres de alilo y sililéter, azida y alcoxi. En realizaciones preferidas, uno o más de un grupo azida, un grupo alcoxi y un grupo disulfuro está asociado con el nucleótido bloqueado reversiblemente (rbNTP) con otra molécula, por ejemplo un hapteno o un resto de detección, o ambos, para uso en métodos como se describen en la presente memoria. La incorporación de un enlace disulfuro en un enlazante como se describe en el presente documento se puede llevar a cabo de varias maneras, por ejemplo como se proporciona en este documento, tal como se encuentra en la patente de EE.UU. 7.771.973, o como se describe en Hermanson, Bioconjugate Techniques, Segunda Edición, Academic Press.

En algunas realizaciones, una composición que comprende un agente de escisión se añade a una reacción de secuenciación para escindir un grupo de escisión en un conector espaciador de un conjugado rbNTP. El agente de escisión añadido depende del grupo de escisión presente. Por ejemplo, la escisión de enlaces disulfuro u otros grupos de escisión reductores se lleva a cabo mediante un agente reductor. La reducción de un enlace disulfuro da como resultado la liberación del rbNTP a partir de la molécula unida, por ejemplo un hapteno, un conjugado de hapteno y/o un resto de detección tal como un resto fluorescente. Los agentes reductores útiles en la práctica de las realizaciones como se describen en este documento incluyen, pero no se limitan a, compuestos de fosfina, fosfinas solubles en agua, fosfinas que contienen nitrógeno y sales y derivados de las mismas, ditioneitol (DTE), ditioneitol (DTT) (isómeros cis y trans, respectivamente, de 2,3-dihidroxi-1,4-ditolbutano), 2-mercaptoetanol o β -mercaptoetanol (BME), 2-mercaptoetanol o aminoetanotiol, glutatión, tioglicolato o ácido tioglicólico, 2,3-dimercaptopropanol y tris(2-carboxietil)fosfina (TCEP), tris(hidroxiometil)fosfina (THP) y ácido β -[tris(hidroxiometil)fosfino]propiónico (THPP). En algunas realizaciones, un agente reductor usado para escindir un enlace disulfuro en un enlazante como se describe en este documento es DTT. En algunas realizaciones, la concentración de un reactivo reductor, por ejemplo DTT, utilizado para escindir un enlace disulfuro es de al menos 1 a 1000 mM, al menos 20 a 800 mM, al menos 40 a 500 mM y preferiblemente al menos 50 a 200 mM. En algunas realizaciones, un agente reductor usado para escindir un enlace disulfuro en un enlazante como se describe en este documento es un reactivo de fosfina, un reactivo de fosfina soluble en agua, un reactivo de fosfina que contiene

5 nitrógeno y sales y derivados de los mismos. Los reactivos de fosfina ejemplares incluyen, pero no se limitan a, TCEP, THP y los descritos en la publicación de patente de EE.UU. 2009/0325172 tales como triaril fosfinas, trialquil fosfinas, fosfinas que contienen sulfonato y carboxilato y fosfinas solubles en agua derivatizadas. En algunas realizaciones, la concentración de una fosfina utilizada para escindir un enlace disulfuro es al menos 0,5-500 mM, al menos de 5 a 50 mM, y preferiblemente al menos de 10 a 40 mM. Los métodos y composiciones como se describen en la presente memoria descriptiva no están limitados por ningún grupo de división particular y las alternativas serán fácilmente evidentes para cualquier experto en la técnica y se consideran dentro del alcance de la presente descripción.

10 En algunas realizaciones, un enlazante como se describe en este documento, que puede o no comprender un sitio de escisión, enlaza un rbNTP con un resto fluorescente y un patrón de transición de fluorescencia para detectar la incorporación del nucleótido en una reacción de SBS se realiza mediante la adición de un colorante inhibidor en un ciclo de secuenciación. Por ejemplo, un rbNTP conjugado a un resto fluorescente a través de un enlazante (en el que el enlazante puede o no comprender un sitio de escisión) se añade a una reacción de secuenciación. Se registra una primera imagen con lo que se establece un primer patrón de detección. Durante una etapa de reacción intermedia, se añade un colorante de extinción a la reacción (por ejemplo, en lugar de una pareja de FRET eliminada de la reacción a través de una etapa de escisión), en el que el colorante de inactivación inactiva suficientemente la fluorescencia del resto fluorescente anteriormente mencionado dando como resultado un patrón de cambio de fluorescencia detectable (por ejemplo, de fluorescencia a ninguna fluorescencia o mínima) en una etapa de formación de imágenes posterior para ese nucleótido. Esta realización es una alternativa a un sistema donador/aceptor de FRET como se describe en este documento, en el que la combinación de dos colorantes da como resultado fluorescencia y la eliminación de uno de los colorantes, por ejemplo mediante una reacción de escisión, da como resultado la pérdida de fluorescencia.

15 Los colorantes de inactivación según se contemplan en este documento incluyen, pero no se limitan a, aquellas sustancias que absorben la energía de excitación de un fluoróforo, inactivando eficazmente la fluorescencia del fluoróforo diana, sin embargo no son típicamente fluorescentes por sí mismos. Ejemplos de colorantes inactivadores incluyen, pero no se limitan a, inactivadores oscuros tales como DABCYL (absorbe en el espectro verde), Iowa negro FQ (absorbe en el espectro verde-amarillo), Iowa negro RQ (absorbe en el espectro rojo anaranjado) IRDye QC-1 (absorbe en el rango de 500-900 nm) y colorantes Black Hole Quencher™ (absorbe en el rango de 500-700 nm). Por ejemplo, el DABCYL se utiliza frecuentemente para inactivar la fluorescencia de la fluoresceína y se utilizan los colorantes Black Hole Quencher™ para inactivar la fluorescencia de los colorantes FAM, TET, HEX, JOE, TAMRA, ROX y CY dependiendo de las características (por ejemplo, absorbancia máxima) del Black Hole Quencher™ particular. En realizaciones adicionales, tales inactivadores oscuros pueden utilizarse en un sistema FRET, en el que la escisión del inactivador oscuro durante un paso intermedio da lugar a un cambio de estado de fluorescencia desde fluorescencia inactivada a fluorescencia, estableciendo así un patrón de detección para la incorporación de un nucleótido en un ciclo de reacción de SBS.

25 El uso de realizaciones de inactivación de colorantes como se describe en este documento se contempla para su uso en permutaciones y combinaciones para detectar la incorporación de un nucleótido en un ciclo de SBS como es reconocido por cualquier experto en la materia. Por ejemplo, un rbNTP puede estar enlazado a un resto fluorescente en el que se utiliza un colorante inactivador para determinar la incorporación de nucleótidos, un segundo rbNTP puede estar enlazado a una biotina en la que se utiliza la adición de un resto fluorescente SA para determinar la incorporación de nucleótidos y un tercer colorante puede estar unido a un resto fluorescente en el que se utiliza una reacción de escisión para determinar la incorporación de nucleótidos. Los métodos como se describen en este documento no están limitados por qué nucleótido se conjuga a qué sistema de detección particular, aparte de que su combinación permite la determinación de la incorporación de nucleótidos en una reacción de secuenciación.

30 En algunas realizaciones, el resto de detección fluorescente se modifica para proporcionar una diferencia de fluorescencia detectable entre la imagen 1 y la imagen 2. Por ejemplo, un resto fluorescente que está unido, directa o indirectamente, a un rbNTP, puede ser visualizado durante un primer evento de imagen. Entre el primer y segundo evento de imagen se puede añadir una molécula química, pequeña, etc. a la reacción de secuenciación de tal manera que la estructura del fluoróforo se modifique haciendo de este modo que el resto fluorescente sea indetectable o mínimamente detectable durante el segundo evento de formación de imágenes. Por ejemplo, se puede añadir un agente de escisión que se dirige a uno o más enlaces y/o entidades estructurales del resto fluorescente que puede destruir la naturaleza fluorescente del resto fluorescente permitiendo de este modo la detección de estados de imagen indicativos de la incorporación del rbNTP unido. Como tal, las modificaciones del propio resto fluorescente pueden proporcionar cambios detectables en estados de formación de imágenes que pueden ser ventajosos en métodos como se describen en la presente memoria.

35 En algunas realizaciones de la presente descripción, un tipo de nucleótido para uso en una reacción de secuenciación es un conjugado de rbNTP que comprende una base, por ejemplo una base natural o modificada. En realizaciones preferidas, una base es una base modificada. En realizaciones preferidas, una base modificada comprende tres grupos fosfato fuera del esqueleto de azúcar, como tal es un trifosfato, como se indica por NTP. En realizaciones preferidas, la base modificada se bloquea de forma reversible en la que el NTP comprende un grupo de bloqueo 3' terminador reversible que, una vez eliminado, permite la prolongación continua en una secuencia por la reacción de secuenciación de síntesis. En algunas realizaciones, el grupo de bloqueo 3' comprende un grupo

azido y/o alcoxi y es removible por escisión con un reactivo de fosfina. Dichos nucleótidos se denominan "bloqueados reversiblemente" o "rb", un tipo del cual es un NTP "completamente funcional" o NTP "ff" (comercialmente disponible en Illumina, Inc.). La discusión adicional de rbNTPs se encuentra en, por ejemplo, las patentes de los Estados Unidos 7.816.503 y 7.771.903 y la publicación de la solicitud de patente de los Estados Unidos US2010/00317531.

Los métodos descritos para la detección de ácidos nucleicos encuentran utilidad particular cuando se usan en secuenciación, por ejemplo en la secuenciación por tecnologías de síntesis (SBS). La secuenciación por síntesis comprende generalmente la adición secuencial de uno o más nucleótidos marcados fluorescentemente a una cadena polinucleotídica en crecimiento en la dirección de 5' a 3' usando una polimerasa. La cadena polinucleotídica extendida es complementaria a la plantilla de ácido nucleico fijada sobre el sustrato (por ejemplo, célula de flujo, chip, lámina, etc.), la secuencia diana. En algunas realizaciones, la identidad de un nucleótido incorporado en la cadena extendida se determina después de dos etapas de formación de imágenes proporcionando de este modo datos de secuencia de incorporación en tiempo real.

El método descrito para la detección de ácidos nucleicos también es útil cuando se usa en secuenciación mediante ligación, secuenciación por hibridación y otras tecnologías de secuenciación en las que se emplean esquemas de modificación de nucleótidos "oscuros" y/o nucleótidos ortogonales.

La secuencia por ligación es un método de secuenciación en el que un cebador de secuenciación es alargado sobre una secuencia diana mediante la ligación de una sonda que comprende un tipo de nucleótido (por ejemplo, A, T, C o G), en donde la sonda ligada es indicativa de la secuencia del subsiguiente nucleótido en una cadena de nucleótidos diana. La secuenciación por sondas de ligación puede comprender sitios de escisión que pueden escindir después de un evento de ligación de manera que se puede realizar otra determinación de la adición, ligación e incorporación del nucleótido de la sonda redonda. Una secuencia ejemplar por metodología de ligación es la codificación de di-bases (por ejemplo, secuenciación de espacio de color) utilizada por el sistema de secuenciación AppliedByosistemas SOLiD™. La codificación de di-bases o la secuenciación de "espacios de color" utiliza sondas problema que comprenden 2 bases específicas de sonda (por ejemplo, constituidas por todas las combinaciones posibles de los cuatro tipos de nucleótidos diferentes) seguidas por tres bases degeneradas y seis bases universales, en donde cada una de las sondas problema está unida a uno de los cuatro colorantes fluorescentes diferentes. Las sondas se añaden a una reacción de secuenciación que comprende una diana y un cebador de secuenciación y una ligasa termoestable liga la sonda de di-bases complementaria a las secuencias adyacentes al cebador de secuenciación tal como se encuentra en la plantilla. La fluorescencia se detecta mediante cuatro imágenes en color, las sondas ligadas se escinden para eliminar el colorante fluorescente y regenerar el fosfato 5' para rondas adicionales de ligación y detección. Cada base de plantilla se ensaya dos veces. Después de varias rondas de ligación y detección de un cebador de secuenciación, la hebra sintética se desnaturaliza, se añade un nuevo cebador de secuenciación y se inicia de nuevo el proceso de detección de ligación. Los bits de datos del espacio de color fluorescente dicodificados se alinean, se aplican a una rejilla de genoma de referencia de espacio de color y se determina la secuencia (Voelkerding et al., 2009, Clin Chem 55: 641-658).

Los nucleótidos modificados descritos en este documento podrían utilizarse en secuencia mediante tecnologías de ligación. Por ejemplo, se pueden asociar sondas de un esquema de codificación de dos bases en el que cuatro secuencias de dinucleótidos están asociadas con un color, por ejemplo AA, CC, GG y TT, con un colorante fluorescente azul, otras cuatro secuencias de dinucleótidos están asociadas con un colorante rojo, otras cuatro un colorante verde son la detección es a través de un sistema de imágenes de cuatro colores podría ser modificado como se describe en el presente documento. La incorporación de menos de cuatro colorantes, por ejemplo un colorante o dos o más colorantes de excitación/emisión similares mientras se practican manipulaciones químicas y/o enzimáticas permitiría menos eventos de formación de imágenes, por lo tanto, una óptica instrumental más simplificada. Por ejemplo, una sonda que comprende cuatro secuencias de dinucleótidos tales como AA, CC, GG y TT, que comprende además un número o nucleótidos degenerados y/o universales (opcionalmente), podría comprender además un enlazante que contiene un sitio de escisión (por ejemplo, un sitio de escisión de azida o alcoxi) que une el dinucleótido con un resto fluorescente. Una sonda que comprende un segundo conjunto de cuatro dinucleótidos, por ejemplo TA, GC, CG y AT, que comprende además un número o nucleótidos degenerados y/o universales (opcionalmente), podría comprender además un enlazante que contiene dos sitios de escisión (el segundo sitio de escisión diferente desde el primero, por ejemplo un enlace SS) que une el dinucleótido con un resto fluorescente. Un conjunto de sondas que comprende un tercer conjunto de cuatro dinucleótidos, por ejemplo CA, AC, GT y TG, que comprende además un número o nucleótidos degenerados y/o universales (opcionalmente), podría comprender además un enlazante que contiene un sitio de escisión que une el dinucleótido con un resto hapteno (por ejemplo, biotina). El cuarto conjunto de sondas de cuatro dinucleótidos podría comprender nucleótidos adicionales, enlazantes, etc., sin embargo carecería de un resto fluorescente. Las sondas se podrían añadir a la secuenciación mediante reacción de ligación, unirse a la plantilla y se podría registrar una primera imagen para capturar un primer estado de señal. Se podría añadir un reactivo de escisión a la reacción para escindir el segundo sitio de escisión (por ejemplo, enlace SS) liberando de este modo el resto fluorescente, pudiendo añadirse una pareja de unión de hapteno (por ejemplo, estreptavidina) conjugada a un resto fluorescente y una segunda imagen podría registrarse para capturar un segundo estado de señal. Se podría añadir un agente de escisión al primer sitio de escisión (por ejemplo, azida/alcoxi) a la reacción para liberar todos los restos fluorescentes y podría llevarse a

cabo la siguiente ronda de secuenciación por ligación. Se podrían alinear los estados de la señal y determinar las secuencias.

La secuencia por hibridación comprende el uso de una serie de secuencias cortas de sondas de nucleótidos a las que se añade ADN diana marcado y fragmentado (Drmanac et al., 2002, *Adv Biochem Eng Biotechnol* 77: 75-101; Lizardi et al., 2008, *Nat Biotech* 26: 649-650). Los fragmentos se hibridan con su sonda complementaria en el conjunto y la hibridación se captura mediante el marcador unido tal como un colorante fluorescente, determinando de este modo la secuencia de la diana. Algunas aplicaciones de secuencia mediante hibridación utilizan sondas que comprenden análogos universales (por ejemplo, análogos de nucleótidos) y nucleótidos designados y se denominan sondas interrumpidas, cuyo uso se sabe que aumenta la sensibilidad de hibridación y, por tanto, la detección del ensayo de secuenciación (Patente de EE.UU. 7.071.324). Se pueden encontrar mejoras adicionales en la secuencia por hibridación en, por ejemplo, las publicaciones de solicitud de patente de EE.UU. 2007/0178516, 2010/0063264 y 2006/028783. Sin embargo, independientemente del método, son necesarios sistemas de óptica a menudo complejos para capturar eventos de hibridación.

Los nucleótidos modificados descritos en este documento podrían utilizarse en secuencia mediante tecnologías de hibridación. Las sondas de ácido nucleico de múltiples muestras diferentes para la determinación de secuencias que se hibridan con sondas agrupadas podrían modificarse para comprender los atributos descritos en este documento para su uso en una secuenciación de colorante mínima, permitiendo así una óptica menos compleja con la determinación concurrente de secuencias de múltiples muestras de ensayo diferentes. Por ejemplo, una sonda de muestra de ensayo (por ejemplo, ácidos nucleicos de ensayo fragmentados) podría modificarse para comprender un enlazante que contiene un sitio de escisión (por ejemplo, un sitio de escisión de azida o alcoxi) que une la sonda con un resto fluorescente. Un segundo conjunto de sondas podría ser modificado para comprender un enlazante que contiene dos sitios de escisión (el segundo sitio de escisión diferente del primero, por ejemplo un enlace SS) que une la segunda sonda con un resto fluorescente. Un tercer conjunto de sondas podría comprender un enlazante que contiene un sitio de escisión que une la sonda de ácido nucleico con un resto hapteno (por ejemplo, biotina). Las sondas podrían añadirse a una secuencia por tipo de hibridación de matriz, reacciones de hibridación de las sondas de ensayo modificadas a las sondas inmovilizadas en la matriz llevada a cabo, y una primera imagen grabada para capturar un primer estado de señal. Se podría añadir un reactivo de escisión a la reacción para escindir el segundo sitio de escisión (por ejemplo, enlace SS) liberando de este modo el resto fluorescente, pudiendo añadirse una pareja de unión a hapteno (por ejemplo, estreptavidina) conjugada a un resto fluorescente y registrándose una segunda imagen para capturar el segundo estado de señal. Se podrían determinar estados de señal, en los que la rejilla del estado de señal de imagen dos podría utilizarse para determinar la localización y por lo tanto la secuencia de las múltiples sondas de ensayo hibridadas diferentes.

Se han desarrollado y comercializado enfoques de secuenciación que combinan las bioquímicas de hibridación y ligación, tal como la tecnología de secuenciación genómica practicada por Complete Genomics, Mountain View, CA). Por ejemplo, la ligación de sonda combinatoria-anclaje, o cPAL™ (Drmanac et al., 2010, *Science* 327 (5961): 78-81) utiliza la bioquímica de ligación mientras explota las ventajas de la secuencia por hibridación. En resumen, la secuenciación de las nanobolas de ADN diana comprende detectar productos de ligación que están formados por un oligonucleótido de anclaje que se hibrida con una secuencia adaptadora que se liga subsiguientemente a una sonda de secuenciación degenerada marcada fluorescentemente que comprende uno de los cuatro nucleótidos especificados en la posición problema. La ligación se produce cuando el nucleótido en la posición problema es complementario al nucleótido en el sitio de detección dentro de la nanobola del ADN diana. El producto de ligación de ancla/sonda estable resultante se detecta fluorescentemente. Después de la lectura, se libera todo el complejo anclaje/sonda, se hibrida el siguiente anclaje a la diana de ADN y se repite el proceso. Como con muchas reacciones de secuenciación, se utilizan cuatro colorantes diferentes detectables, uno para cada nucleótido problema especificado A, C, G y T utilizando una óptica de detección múltiple.

Los nucleótidos modificados descritos en la presente memoria podrían utilizarse en tecnologías de secuenciación de ligación de anclaje-sonda combinatoria. La incorporación de menos de cuatro colorantes permitiría menos eventos de imagen. Por ejemplo, una sonda que comprende un número o nucleótidos degenerados podría comprender además un enlazante que contiene un sitio de escisión (por ejemplo, un sitio de escisión de azida o alcoxi) que une el nucleótido problema con un resto fluorescente. Una sonda que comprende un segundo conjunto de nucleótidos degenerados podría comprender además un enlazante que contiene dos sitios de escisión que unen el nucleótido problema con un resto fluorescente. Un conjunto de sondas que comprende un tercer conjunto de nucleótidos degenerados podría comprender además un enlazante que contiene un sitio de escisión que une el nucleótido problema con un resto hapteno (por ejemplo, biotina). El cuarto conjunto de sondas de nucleótidos degenerados podría comprender nucleótidos adicionales, enlazantes, etc., sin embargo carecería de un resto fluorescente. Las sondas se podrían añadir a la reacción cPAL™, ligadas al anclaje/adaptador y se podría registrar una primera imagen para capturar un primer estado de señal. Se podría añadir un reactivo de escisión a la reacción para escindir el segundo sitio de escisión (por ejemplo, el enlace SS) liberando de este modo el resto fluorescente, pudiendo añadirse una pareja de unión de hapteno (por ejemplo, estreptavidina) conjugada a un resto fluorescente y podría ser registrada una segunda imagen para capturar un segundo estado de señal. Se podría añadir un agente de escisión al primer sitio de escisión (por ejemplo, azida/alcoxi) a la reacción para liberar todos los restos fluorescentes y podría llevarse a cabo la siguiente ronda de cPAL™. Se podrían alinear los estados de la señal y determinar las secuencias.

Los ácidos nucleicos o polinucleótidos para la secuenciación incluyen, pero no se limitan a, ácidos nucleicos tales como ADN, ARN o PNA (ácido nucleico peptídico), variantes o fragmentos de los mismos y/o concatámeros de los mismos. Los polinucleótidos pueden ser de una secuencia conocida o desconocida, de naturaleza natural o artificial y pueden ser de cualquier fuente (por ejemplo, eucariotas o procariotas). Los polinucleótidos pueden derivarse naturalmente, producirse recombinantemente o sintetizarse químicamente. Los polinucleótidos concatamerizados pueden contener subunidades o análogos de los mismos que pueden o no pueden producirse en la naturaleza, o subunidades modificadas. Pueden usarse métodos como se describe en este documento para determinar una secuencia de un polinucleótido. La longitud del ácido nucleico diana para la secuenciación puede variar. Por ejemplo, el ácido nucleico para la secuenciación puede incluir al menos 10, al menos 20, al menos 30, al menos 40, al menos 50, al menos 100, al menos 200, al menos 500, al menos 1.000, al menos 10.000, al menos 100.000, al menos 1.000.000, al menos 10.000.000 nucleótidos. El polinucleótido para la secuenciación puede ser de origen genómico o fragmentos o variantes de los mismos. La cadena de ácido nucleico para la secuenciación puede ser de cadena sencilla y puede o no derivarse de una molécula de ácido nucleico bicatenario. Las moléculas de cadena sencilla también pueden producirse, por ejemplo, mediante métodos y tecnologías de síntesis química o in vitro. Las realizaciones como se describen en la presente memoria descriptiva no están limitadas por los métodos preparatorios del ácido nucleico y cualquier número de métodos puede ser practicado por los expertos en la técnica para proporcionar una composición para uso en los métodos descritos. Por ejemplo, en la secuencia mediante metodologías de síntesis a menudo se genera una biblioteca que comprende los ácidos nucleicos diana, y después se secuencia una parte de la biblioteca de ADN.

El ADN aislado de muestras, por ejemplo muestras que contienen ADN genómico, se modifica típicamente antes de la caracterización, por ejemplo mediante secuenciación utilizando métodos como los que se describen en la presente memoria. Se crean bibliotecas de ADN genómico que pueden secuenciarse mediante la práctica de los métodos descritos en la presente memoria. Se produce una biblioteca, por ejemplo, realizando los métodos descritos en el kit de preparación de muestras de ADN Nextera™ (Epicentre® Biotechnologies, Madison WI), kits de preparación de bibliotecas SOLiD™ (Applied Biosystems™ Life Technologies, Carlsbad, CA) y similares. Una muestra de biblioteca de ADN puede amplificarse adicionalmente para la secuenciación mediante, por ejemplo, múltiples técnicas de amplificación de desplazamiento de soporte (MDA).

Para la secuenciación después de MDA, se prepara, por ejemplo, una biblioteca de muestras amplificada creando una biblioteca de ADN como se describe en el kit de Mate Pair Library Prep, kits de Preparación de Muestras de ADN Genómico o kits de Preparación de Muestras TruSeq™ y Enriquecimiento Exome (Illumina®, Inc., San Diego CA). Las bibliotecas de ADN pueden inmovilizarse en una célula de flujo y una amplificación de puente realizada sobre los polinucleótidos inmovilizados antes de la secuenciación, por ejemplo secuenciación por metodologías de síntesis. En la amplificación de puente, se hibrida un polinucleótido inmovilizado (por ejemplo, de una biblioteca de ADN) a un cebador de oligonucleótido inmovilizado. El extremo 3' de la molécula de polinucleótido inmovilizado proporciona la plantilla para una reacción de elongación dirigida por plantilla catalizada por polimerasa (por ejemplo, extensión de cebador) que se extiende desde el cebador oligonucleotídico inmovilizado. El producto de doble hebra resultante "puentea" los dos cebadores y ambas hebras están unidas covalentemente al soporte. En el ciclo siguiente, después de la desnaturalización que produce un par de hebras simples (la plantilla inmovilizada y el producto de cebador extendido) inmovilizadas al soporte sólido, ambas cadenas inmovilizadas pueden servir como plantillas para la extensión de un nuevo cebador. De este modo, la primera y segunda partes pueden amplificarse para producir una pluralidad de agrupaciones. Los grupos y colonias se usan indistintamente y se refieren a una pluralidad de copias de una secuencia de ácido nucleico y/o complementos de las mismas unidas a una superficie. Típicamente, el grupo comprende una pluralidad de copias de una secuencia de ácido nucleico y/o complementos de las mismas, unidas a través de sus extremos 5' a la superficie. La metodología de amplificación y agrupamiento de puentes ejemplares se describe, por ejemplo, en la Publicación de Patente PCT. No. WO00/18957 y WO98/44151, patente de EE.UU. No. 5.641.658; Publ. de patente de EE.UU. No. 2002/0055100; patente de EE.UU. No. 7.115.400; Publ. de patente de EE.UU. 2004/0096853; Publ. de patente de EE.UU. Nº 2005/0100900, Publ. de patente de EE.UU. 2004/0002090; Publ. de patente de No. 2007/0128624; y Publ. de patente de No. 2008/0009420. Las composiciones y métodos descritos en la presente memoria son particularmente útiles en secuencia mediante metodologías de síntesis que utilizan una célula de flujo que comprende clústeres.

También pueden usarse métodos de PCR en emulsión para amplificar ácidos nucleicos antes de la secuenciación en combinación con métodos y composiciones como se describen en la presente memoria. La PCR en emulsión comprende la amplificación por PCR de una biblioteca de ADN al azar flanqueada con adaptador en una emulsión de agua en aceite. La PCR es una PCR multi-plantilla; sólo se utiliza un solo par de cebadores. Uno de los cebadores de PCR está atado a la superficie (unida por 5') de perlas de microescala. Una concentración de plantilla baja da como resultado que la mayoría de las microvesículas de emulsión que contienen perlas tengan presentes cero o una moléculas de plantilla. En las microvesículas de emulsión productiva (una microvesícula de emulsión en la que están presentes tanto una perla como una molécula plantilla), los amplicones de PCR pueden capturarse en la superficie de la perla. Después de romper la emulsión, las perlas que llevan productos de amplificación pueden enriquecerse selectivamente. Cada perla clonalmente amplificada soportará en su superficie productos de PCR correspondientes a la amplificación de una única molécula de la biblioteca de plantillas. Se establecen diversas realizaciones de métodos de PCR en emulsión en Dressman et al., Proc. Natl. Acad. Sci. USA 100: 8817-8822

(2003), Publicación de patente PCT. No. WO 05/010145, Publ. de patente de EE.UU. Nos. 2005/0130173, 2005/0064460, y US2005/0042648.

Las nanobolas de ADN también se pueden usar en combinación con métodos y composiciones como se describen en la presente memoria. Los métodos para crear y utilizar nanobolas de ADN para la secuenciación genómica pueden encontrarse en, por ejemplo, las patentes de EE.UU. y publicaciones 7.910.354, 2009/0264299, 2009/0011943, 2009/0005252, 2009/0155781, 2009/0118488 y como se describe en Drmanac et al., 2010, Science 327 (5961): 78-81. Brevemente, después de la fragmentación del ADN genómico, las rondas consecutivas de ligación adaptadora, amplificación y digestión dan lugar a concatámeros de cabeza a cola de múltiples copias de las secuencias plantilla/adaptador de ADN genómico circular que se circularizan en ADN monocatenario por ligación con un ligasa en círculo y círculo rodante amplificado (como se describe en Lizardi et al., Nat. Genet. 19: 225-232 (1998) y el documento de EE.UU. 2007/0099208 A1). La estructura de adaptador de los concatámeros promueve el enrollamiento del ADN de una sola hebra creando así nanobolas de ADN compacto. Las nanobolas de ADN pueden ser capturadas sobre sustratos, preferiblemente para crear una matriz ordenada o modelada de manera que se mantenga la distancia entre cada nanobola permitiendo de este modo la secuenciación de las nanobolas de ADN separadas.

Cualquier experto en la técnica reconocerá métodos y tecnologías adicionales para amplificar ácidos nucleicos que también podrían usarse en combinación con los métodos y composiciones descritos en la presente memoria. Las realizaciones descritas en la presente memoria descriptiva no están limitadas a ningún método de amplificación de ADN.

Los métodos que se describen en la presente memoria descriptiva no están limitados por ningún método de preparación de muestras de secuenciación particular y las alternativas serán fácilmente evidentes para cualquier experto en la técnica y se consideran dentro del alcance de la presente descripción. Sin embargo, se encuentra una utilidad particular cuando se aplican los métodos de la presente invención a dispositivos de secuenciación tales como células o matrices de flujo para practicar la secuenciación mediante metodologías de síntesis u otras tecnologías de secuenciación relacionadas tales como las practicadas por una o más de la tecnología de secuenciación Polony (Dover Systems), secuenciación por plataformas de hibridación fluorescente (Complete Genomics), tecnología sTOP (Instituto de Investigación de Tecnología Industrial) y secuenciación por síntesis (Illumina, Life Technologies).

En algunas realizaciones, los métodos expuestos en la presente memoria pueden usarse en una versión modificada de los protocolos del fabricante en un sistema tal como los proporcionados por Illumina®, Inc. (sistemas HiSeq 1000, HiSeq 1000, Genome Analyzers, MiSeq, HiScan, iScan, BeadExpress), Applied Biosystems™ Life Technologies (sistemas de detección de secuencias ABI PRISM®, SOLiD™ System), u otros instrumentos de secuenciación basados en fluorescencia, además de los descritos en, por ejemplo, las patentes de EE.UU. y las solicitudes de patentes 5.888.737, 6.175.002, 5.695.934, 6.140.489, 5.863.722, 2007/007991, 2009/0247414, 2010/0111768 y la solicitud de patente PCT WO2007/123744, y WO2012/096703. Las modificaciones a los métodos comerciales pueden incluir, pero no se limitan a, la alteración de los marcadores utilizados y la adición de etapas para cambiar los estados del marcador como se expone en este documento.

La salida de un instrumento de secuenciación puede ser de cualquier tipo. Por ejemplo, la tecnología actual utiliza típicamente una salida legible generadora de luz, tal como fluorescencia o luminiscencia, sin embargo los presentes métodos no están limitados al tipo de salida legible mientras las diferencias en la señal de salida para una secuencia particular de interés sean potencialmente determinables. Ejemplos de software de análisis que pueden usarse para caracterizar la salida derivada de la práctica de métodos tal como se describen en este documento incluyen, pero no se limitan a, el software de análisis de datos Pipeline, CASAVA y GenomeStudio (Illumina®, Inc.), SOLiD™, DNASTAR® SeqMan® NGen® y el software de análisis de datos Partek® Genomics Suite™ (Life Technologies), el software de análisis de datos Feature Extraction and Agilent Workbench (Agilent Technologies), Genotyping Console™, el software de análisis de datos Chromosome Analysis Suite (Affymetrix®).

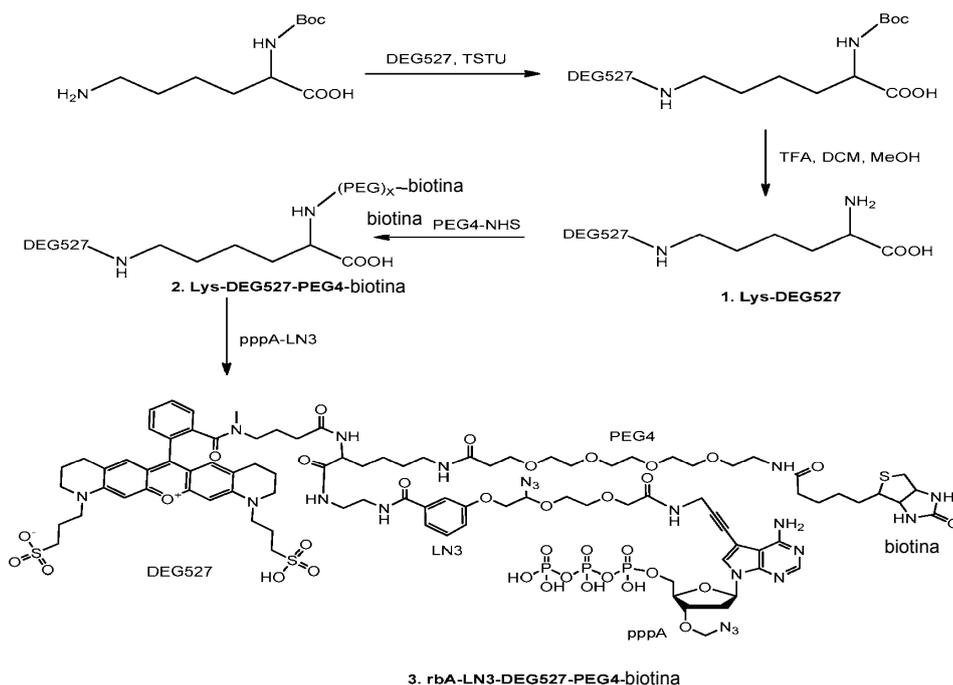
Cualquier experto en la materia conocerá otras numerosas alternativas de software comercial y académicamente disponibles para el análisis de datos para la producción generada por secuenciación. Las realizaciones descritas en la presente memoria descriptiva no están limitadas a ningún método de análisis de datos.

Ejemplos

Los siguientes ejemplos se proporcionan con el fin de demostrar e ilustrar adicionalmente ciertas realizaciones y aspectos de la presente descripción y no se deben interpretar como limitativos del alcance de la misma.

Ejemplo 1-Síntesis de rbATP-LN3-DEG527-PEG4-Biotina

Se sintetizó una construcción de adenina bloqueada reversiblemente, biotilada, ramificada y marcada con fluorescencia para su uso en SBS como sigue:



Lys-DEG527

5 A una solución de DEG527 (11 mg, 14,6 μmol) en DMA seco (2 ml) se añadió TSTU (5,3 mg, 17,5 μmol) y diisopropiletilamina (6,3 μl , 36,5 μmol). La mezcla se agitó durante 30 minutos a temperatura ambiente hasta la activación completa del ácido. Se añadió a la mezcla de reacción una solución de Boc-lisina (18 mg, 73 μmol) en TEAB 0,1 M (0,2 ml). La mezcla se agitó durante 3 horas hasta que TCL mostró un consumo completo del éster activado. Los volátiles se evaporaron a presión reducida y el residuo se disolvió en ácido trifluoroacético (0,1 ml), DCM (0,9 ml) y MeOH (0,1 ml). La solución se agitó a temperatura ambiente durante 1 hora hasta que la TLC mostró el consumo total del material de partida. La solución se concentró hasta sequedad, se volvió a disolver en TEAB 0,1 M (5 ml) y se purificó mediante RP-HPLC.

Lys-DEG527-PEG4-Biotina

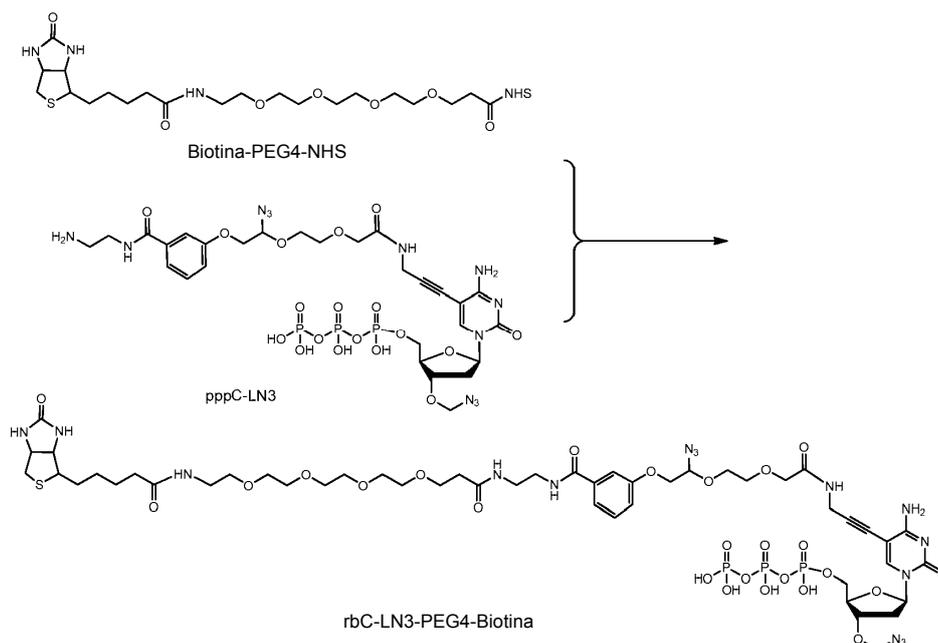
15 Se añadió PEG4-biotina-NHS (41 mg, 70 μmol) a una solución de Lys-DEG527 (14 μmol) y diisopropiletilamina (15 μl , 84 μmol) en DMA seco (5 ml). La mezcla se sonicó durante varios minutos y luego se agitó continuamente durante varias horas. La TCL mostró el consumo completo de lys-DEG527. Los volátiles se evaporaron bajo presión reducida. El residuo se disolvió de nuevo en TEAB 0,1 M (5 ml) y se purificó por RP-HPLC.

rbATP-LN3-DEG527-PEG4-Biotina

20 A una solución de Lys-DEG527-PEG4-biotina (9 μmol) en DMA seco (2 ml), se añadió TSTU (3,3 mg, 10,8 μmol) y diisopropiletilamina (4 μl , 22,5 μmol). La mezcla se agitó durante 30 minutos a temperatura ambiente hasta la activación completa del ácido. Se añadió una solución de LN3-pppA (18 μmol) en TEAB 0,1 M (0,2 ml) a la mezcla de reacción. La mezcla se agitó durante 5 horas hasta que la TCL mostró el consumo completo del éster activado. La reacción se inactivó con tampón TEAB (0,1 M, 10 ml) y se cargó en una columna de DEAE Sephadex (2 x 5 cm). La columna se eluyó con un gradiente de tampón TEAB de 0,1 M a 1 M en 30 min a 25 ml/min. Las fracciones que contenían el producto se combinaron, se evaporaron y se purificaron por HPLC.

Ejemplo 2-Síntesis de rbCTP-LN3-PEG4-Biotina

25 Se sintetizó un constructo de citosina biotinilado, bloqueado reversiblemente, para su uso en SBS como sigue:

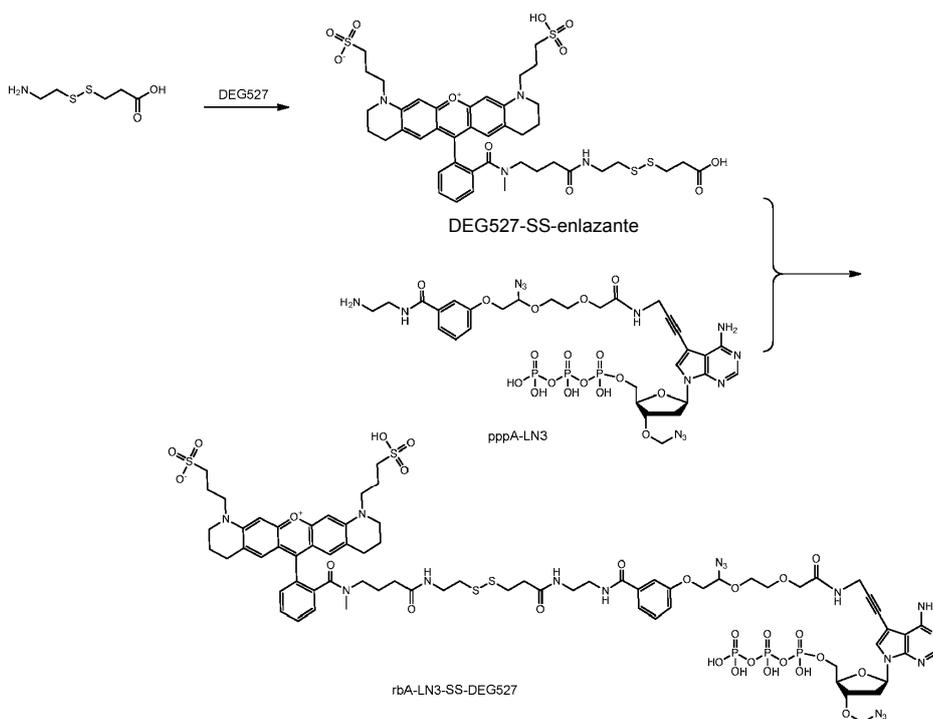


rbCTP-LN3-PEG4-Biotina

5 A una solución de PEG4-biotina-NHS (17,7 mg, 30 μ mol) y diisopropiletilamina (8,7 μ l, 50 μ mol) en DMA seco (3 ml), se añadió una solución de LN3-pppC (10 μ mol) en TEAB 0,1M (0,3 ml). La mezcla se agitó a temperatura ambiente durante 5 horas. El progreso de la reacción se controló mediante RP-HPLC hasta el consumo completo de LN3-pppC. La reacción se inactivó con tampón TEAB (0,1 M, 10 ml) y se cargó en una columna de Sephadex de DEAE (2 x 5 cm). La columna se eluyó con un gradiente de tampón TEAB de 0,1 M a 1 M en 30 min a 25 ml/min. Las fracciones que contenían el producto se combinaron, se evaporaron y se purificaron por HPLC.

Ejemplo 3-Síntesis de rbATP-LN3-SS-DEG527

10 Se sintetizó una construcción de adenina totalmente funcional, marcada fluorescentemente, que comprendía un enlazante escindible para uso en SBS, de la siguiente manera:



DEG527-SS-enlazante

5 A una solución de DEG527 (12,5 mg, 16 μ mol) en DMA seco (2 ml), se añadió TSTU (6 mg, 20 μ mol) y diisopropiletilamina (7 μ l, 40 μ mol). La mezcla se agitó durante 30 minutos a temperatura ambiente hasta la activación completa del ácido. Se añadió a la mezcla de reacción una solución de SS-enlazador (9 mg, 50 μ mol) en TEAB 0,1 M (0,2 ml). La mezcla se agitó durante 3 horas hasta que la TCL mostró un consumo completo del éster activado. Los volátiles se evaporaron a presión reducida y el residuo se disolvió en TEAB 0,1 M (5 ml) y se purificó por RP-HPLC.

rbATP-LN3-SS-DEG527

10 Se añadió TSTU (2,1 mg, 7,1 μ mol) y diisopropiletilamina (2,6 μ l, 14,8 μ mol) a una solución de DEG527-SS-enlazante (5,9 μ mol) en DMA seco (2 ml). La mezcla se agitó durante 30 minutos a temperatura ambiente hasta la activación completa del ácido. Se añadió una solución de LN3-pppA (17,7 μ mol) en TEAB 0,1 M (0,2 ml) a la mezcla de reacción. La mezcla se agitó durante 5 horas hasta que la TCL mostró el consumo completo del éster activado. La reacción se inactivó con tampón TEAB (0,1 M, 10 ml) y se cargó en una columna de Sephadex de DEAE (2 x 5 cm). La columna se eluyó con un gradiente de tampón TEAB de 0,1 M a 1 M en 30 min a 25 ml/min. Las fracciones que contenían el producto se combinaron, se evaporaron y se purificaron por HPLC.

15 Ejemplo 4-Detección de la incorporación de nucleótidos utilizando la construcción de nucleótidos conjugada con biotina

Se realizaron experimentos para demostrar el uso de un nucleótido conjugado con biotina en reacciones de secuenciación. La firma del espacio de tiempo de los experimentos siguió el patrón de formación de imágenes de espacio de tiempo

	Imagen 1	Imagen 2
A	1	0
C	0	1
G	0	0
T	1	1

20 Los experimentos se realizaron en un Genome Analyzer Iix configurado en modo de un solo carril. Se siguió una secuencia estándar mediante el programa de incorporación de enzimología de síntesis utilizando la mezcla de nucleótidos bloqueada de forma reversible que incluía un rbGTP no marcado, rbTTP-LN3-NR550 marcado fluorescentemente, rbCTP-LN3-PEG4-biotina biotinilada y rbATP con un enlazante disulfuro (SS) escindible rbATP-LN3-SS-DEG527. La adquisición y el análisis de datos difirieron de la química estándar de SBS de 4 colorantes. Brevemente, después de una etapa de incorporación de nucleótidos, los agrupamientos se excitaron con láser y se adquirió una imagen fluorescente. Se añadieron componentes de reacción adicionales a la reacción para escindir selectivamente el enlace SS de rbATP-LN3-SS-DEG527 y se añadió SA-NR555 para marcar selectivamente rbCTP-LN3-biotina para crear rbCTP-LN3-biotina-SA-NR555. Los agrupamientos se excitaron por láser una segunda vez que se registró una segunda imagen fluorescente. Por lo tanto, la incorporación de cada una de las cuatro bases es por cambios, o falta de los mismos, de estados de intensidad fluorescente usando colorantes que excitan y emiten en la misma longitud de onda.

35 Se creó una genoteca de ADN genómico para su uso en secuenciación de lectura única en un Genome Analyzer Iix (Illumina, Inc.). Después de la preparación de la biblioteca, se creó una célula de flujo secuencial con agrupamientos de secuenciación objetivo usando el TruSeq SR Cluster Kit v2 en el Illumina® cBot siguiendo el protocolo del fabricante para la secuencia de lectura única. Después de la generación del agrupamiento, la célula de flujo se colocó en un Genome Analyzer Iix y la muestra se secuenció usando reactivos de TruSeq SBS Reagent Kit v5 (Illumina®, Inc.).

40 Se prepararon soluciones madre de los nucleótidos bloqueados reversiblemente para su uso en la reacción de secuenciación; 100 μ M de soluciones madre de rbGTP oscuro o no marcado, rbATP-LN3-SS-DEG527, rbCTP-PEG4-biotina y rbTTP-LN3-NR550. Se preparó una solución madre de estreptavidina-NR555 (SA-NR555 a 1 mg/ml) en un tampón de unión y de lavado (Tris 5 mM, pH 7,5, EDTA 0,5 mM, NaCl 1M).

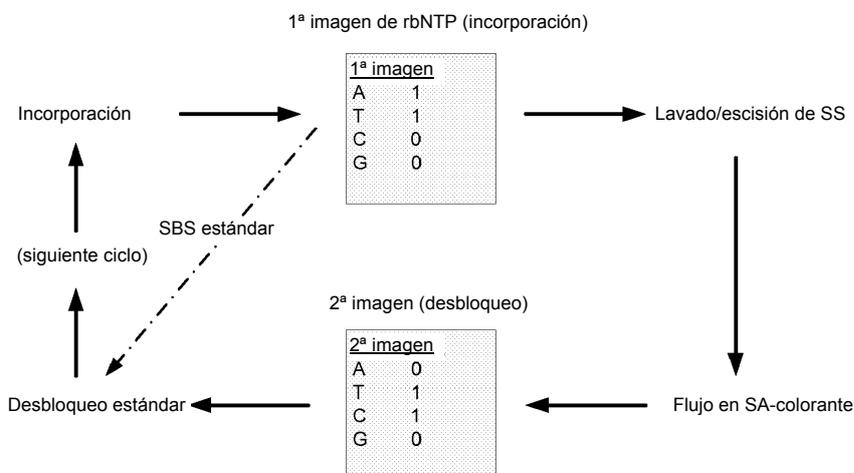
45 Para el Genome Analyzer Iix, las posiciones del reactivo del instrumento fueron reconfiguradas para una secuencia de colorante única por síntesis. Se seleccionó un carril para la secuenciación y los otros carriles se desconectaron, asegurando de este modo que los reactivos se extraían a través de un carril de secuenciación y que no pudiera ocurrir ningún cruce de ningún líquido de otro carril. Los reactivos se colocaron en el Genome Analyzer Iix (GAIix) como sigue:

Posición	Reactivo
1	Mezcla de incorporación (IMX)
2	Blanco
3	Mezcla Scan (SMX)
4	Blanco
5	Tampón de Incorporación (PR2)
6	Mezcla de escisión (CLM)
7	Blanco
8	SA-NR555

Los reactivos se prepararon para un ensayo de secuenciación de 150 ciclos. Del TruSeq SBS Reagent Kit v5, los reactivos CLM, SMX y PR2 se utilizaron según las instrucciones. Para el reactivo IMX que contenía los nucleótidos bloqueados reversiblemente, se añadieron a 20,1 ml de tampón IMX 1 ml de rbATP-SS-DEG527 (concentración final 4 μ M), 0,5 ml de rbGTP (concentración final 2 μ M), 2,5 ml de rbCTP-PEG4-biotina 10 μ M y 0,25 ml de rbTTP-LN3-NR550 (concentración final 1 μ M). La solución de rbNTP se filtró y se añadieron 0,6 ml de polimerasa de alta densidad (HDP, concentración final de 15 μ g/ml). Se hizo una dilución 1:200 de SA-NR555 en tampón de unión y lavado.

Los reactivos se cargaron en el Genome Analyzer Iix y se ejecutó el protocolo de secuenciación. Brevemente, se siguió una etapa de incorporación estándar (es decir, FirstBase) por formación de imágenes como se describe en el protocolo del fabricante. La obtención de imágenes se siguió inmediatamente por escisión de disulfuro (adición de CLM) y unión de SA-NR555 (adición de dilución 1:200 de SA-NR555) y una subsiguiente segunda formación de imágenes seguida por una etapa estándar de desbloqueo e incorporación (es decir, CompleteCycle). La escisión de las uniones disulfuro que daba como resultado un cambio en el estado de intensidad para rbATP de 1 a 0 fue selectiva y continuó a una velocidad rápida de <5 segundos a temperatura ambiente. La unión de biotina/estreptavidina también se produce rápidamente a una velocidad de <25 segundos a temperatura ambiente dando como resultado un cambio en el estado de intensidad para rbCTP de 0 a 1.

El tiempo total de ciclo excluyendo la formación de imágenes fue de alrededor de 9,3 minutos. Se repitió el ciclo para los ciclos restantes. El flujo general es el siguiente:



Se pueden encontrar resultados ejemplares en las Figuras 1 y 2 y en la Tabla 1. La Figura 1 ejemplifica un mapa de calor de estilo de nube registrado en diferentes ciclos a lo largo de la secuencia de secuenciación. Los mapas de nubes demuestran que la diferenciación de los cuatro nucleótidos fue exitosa (el fondo aislado y el mapa de nubes marcado con nucleótidos orientan las posiciones de los cuatro nucleótidos dentro del mapa de nubes). La Figura 2 reporta un seguimiento ejemplar de porcentajes de tasa de error de incorporación de nucleótidos sobre una

secuencia de secuenciación de 100 ciclos para el carril 4 seleccionado, losa 4. Se registró una tasa de error de 0,04% sobre 100 ciclos para la pista 4, la losa 4 en una célula de flujo, mientras que la Figura 2 demuestra que no había llamadas de base en blanco a lo largo del ciclo de 100 ciclos para ese carril y losa. La separación de fases se registró al 0,27% y la separación de fases previa al 0,43%. La Tabla 1 muestra los resultados del carril 4, losa 1-6.

5

Tabla 1

Carril	Losa	Agrupamientos (crudos)	% Agrupamientos PF	% Alineamiento (PF)	Av Puntuación de Alineamiento (PF)	% tasa de error (PF)
4	1	288360	77,64	96,08	121,87	0,51
4	2	285563	78,67	96,04	121,39	0,7
4	3	282653	79,5	96,12	121,97	0,48
4	4	280818	79,07	95,92	121,87	0,4
4	5	283422	78,36	96,05	121,97	0,43
4	6	282958	61,68	60,7	74,58	2,22

Ejemplo 5-Detección de la incorporación de nucleótidos usando un colorante

Se realizaron experimentos para demostrar que se puede usar un colorante para determinar la secuencia de un ácido nucleico.

10 Los nucleótidos utilizados en este experimento incluyeron:

rbATP-LN3-SS-NR550C4

rbTTP-LN3-NR550C4

rbCTP-(LN3)²-Biotina

RbGTP-sin marcador

15 Todas las concentraciones de la disolución madre de nucleótidos se almacenaron a 100 µM en tampón Tris 10 mM (pH 8,0). El resto fluorescente usado para marcar los nucleótidos fue NR550C4. En la Figura 5B se muestran dos espectros de emisión representativos para el colorante sobre rbATP y rbTTP. El rbGTP no se marcó y por lo tanto se consideró el nucleótido "oscuro". Para determinar la incorporación de citosinas en una hebra de ácido nucleico en crecimiento, se añadió a la reacción una mezcla maestra que incluía un conjugado de estreptavidina-NR550C4 como se detalla a continuación.

20 La síntesis de la composición enlazador-NR550C4-SS se realizó como se describió previamente para la composición de DEG527-SS-enlazador, excepto que se usó el resto fluorescente NR550C4 en lugar del fluoróforo DEG527. La síntesis de la composición de rbATP-LN3-SS-NR550C4 se realizó como se describió anteriormente para rbATP-LN3-SS-DEG527, sin embargo se usó la composición de NR550C4-SS-enlazador en lugar de la composición de DEG527-SS-enlazador. La síntesis de la composición de rbTTP-LN3-NR550C4 se realizó como se describió para rbATP-LN3-SS-550C4, sin embargo se usó rbTTP-LN3 en lugar de rbATP-LN3 y se usó NR550C4 en lugar de NR550C4-SS-enlazante. La síntesis de la rbCTP-(LN3)²-Biotina se realizó como se describió anteriormente para rbCTP-LN3-PEG4-Biotina, excepto que se usó LN3-Biotina en lugar de biotina durante la reacción de acoplamiento de amida.

30 La estreptavidina se conjugó con NR550C4 por métodos conocidos en la técnica y se preparó una solución madre de Strep-NR550C4 (SA-NR550C4) a 1 mg/ml en un tampón de Tris 5 mM, pH 7,5, EDTA 0,5 mM y NaCl 1M.

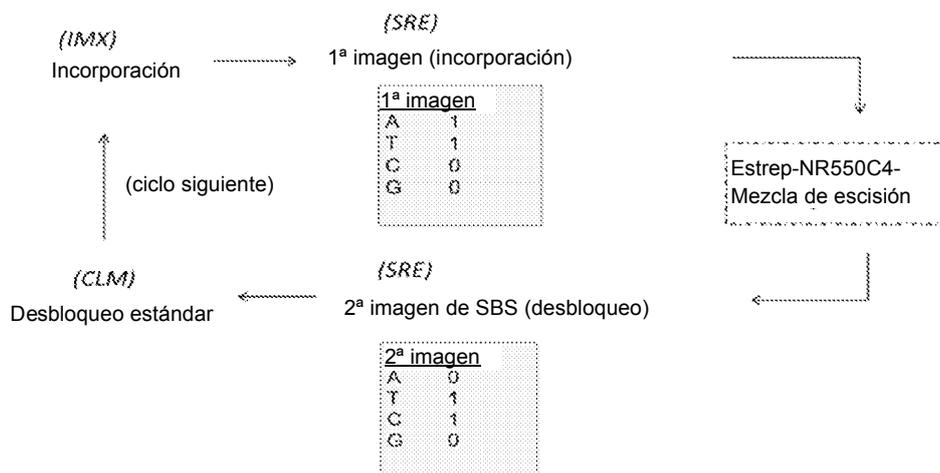
35 Al reactivo IMX, se añadieron soluciones madre de las composiciones de nucleótidos para producir las concentraciones finales de 2 µM de rbATP-LN3-SS-NR550C4, 10 µM de rbCTP-(LN3)²-Biotina, 1 µM de rbTTP-LN3-NR550C4 y 2 µM de rbGTP-oscuro. Adicionalmente, se añadió 15 µg/ml de una polimerasa de Alta Densidad al reactivo IMX/nucleótido. Los reactivos CLM, SMX y PR2 fueron como se ha descrito previamente. Se preparó una mezcla maestra, SA-NR550C4-Cleavage Mix, diluyendo SA-NR550C4 hasta una concentración final de 5 µg/ml en THP 2 mM, Tris 5 mM a pH 7,4, NaCl 1 M, EDTA 0,5 mM y Tween 0,005%.

Los experimentos de secuenciación de un colorante se realizaron en un instrumento de secuenciación MiSeq™ (Illumina, Inc.). La posición de los reactivos en el instrumento fue:

- 1-IMX
- 2-SRE (Scan Mix)
- 5 3-PR2
- 4-CLM
- 18-SA-NR550C4-Mezcla de escisión

El instrumento se ajustó a 60°C al comienzo de los experimentos de secuenciación y todas las etapas de secuenciación incluyendo las etapas de formación de imágenes se llevaron a cabo a esta temperatura para la secuenciación isotérmica. La secuenciación isotérmica se realizó comparativamente a la secuenciación realizada sobre el GAIIx como se describió anteriormente, donde la formación de imágenes se llevó a cabo a 22°C.

Para la secuenciación MiSeq™, el tiempo de ciclo de la química SBS total para un ciclo de incorporación (excluyendo los ciclos de formación de imágenes) fue de 3,37 minutos (1 colorante SBS fue 2,70 minutos y el marcado de SA y escisión fue 0,67 segundos). Los ciclos de secuenciación se repitieron básicamente como se describe a continuación:



Los resultados del experimento de secuenciación de un colorante se pueden encontrar en las Figuras 5A y D y la Tabla 2. La separación de fases se registró al 0,17% y la separación de fases previa al 0,36%. La Tabla 2 muestra los resultados del carril 1, losas 1-4.

20 Tabla 2

Carril	Losa	Agrupamientos (crudos)	% Agrupamientos PF	% Alineamiento (PF)	Puntuación de Alineamiento (PF)	% tasa de error (PF)
1	1	331576	71,85	90,85	673,37	2,13
1	2	331383	71,51	91,11	677,41	2,09
1	3	334956	72,19	90,87	676,4	2,07
1	4	333278	72,74	90,97	671,76	2,18

La Figura 5A muestra un ejemplo de seguimiento de los porcentajes de tasa de error de ciclos de base a lo largo de una secuencia de secuenciación de 150 ciclos para una losa. Se observó una tasa de error de aproximadamente 2,12% sobre 150 ciclos. Basándose en el diseño experimental, las Figuras 5B y C muestran ejemplos de eventos de formación de imágenes de patrones de detección que deberían resultar para los diferentes nucleótidos modificados para cada uno de los eventos de formación de imágenes. Por ejemplo, la Figura 5B Imagen 1 muestra que el primer

evento de imagen no debería capturar fluorescencia, o una mínima fluorescencia, para rbGTP o rbCTP-(LN3)²-Biotina ya que no están asociadas con ningún resto fluorescente antes del primer evento de formación de imágenes y fluorescencia para los nucleótidos marcados con rbATP y rbTTP ya que están asociados con un resto fluorescente antes del primer evento de formación de imágenes. La Figura 6C Imagen 2 muestra que después de la adición de SA-NR550C4-Mezcla de Separación no debería haber ninguna fluorescencia, o mínima, del nucleótido modificado con rbATP, ya que el disulfuro en la composición de rbATP-LN3-SS-NR550C4 debería ser escindido liberando así el fluoróforo unido y la incorporación de rbCTP en la cadena de ácido nucleico en crecimiento sería detectable debido a la unión de la composición de SA-NR550C4 a la biotina en el conjugado rbCTP-(LN3)²-Biotina. Los patrones fluorescentes de rbGTP y rbTTP-LN3-NR550C4 deben permanecer iguales de la Imagen 1 a la Imagen 2 cuando sigan el diseño experimental descrito en este Ejemplo.

La Figura 5D muestra un diagrama de nubes que demuestra que, sorprendentemente, el patrón de detección de fluorescencia siguió el patrón de imagen propuesto y que cada uno de los nucleótidos podía ser diferenciado entre sí cuando se incorporaba en una hebra de nucleótidos en crecimiento utilizando solamente un colorante y dos eventos de formación de imágenes en un experimento de secuenciación.

Estos resultados mostrados en esta descripción demuestran que la secuenciación de un ácido nucleico se puede lograr utilizando tan sólo un colorante fluorescente y menos de cuatro eventos de formación de imágenes para diferenciar la incorporación de los cuatro ácidos nucleicos diferentes en un ciclo de secuenciación.

Varias modificaciones y variaciones de los métodos y composiciones descritos serán evidentes para los expertos en la técnica sin apartarse del alcance de la descripción. Aunque la invención se ha descrito en conexión con realizaciones preferidas específicas, debe entenderse que la invención reivindicada no debe limitarse indebidamente a tales realizaciones específicas. De hecho, se pretende que varias modificaciones de las realizaciones descritas, que son evidentes para los expertos en los campos relevantes, estén dentro del alcance de las siguientes reivindicaciones.

REIVINDICACIONES

- 5 **1.** Un método de secuenciación por síntesis (SBS) basado en fluorescencia para determinar la secuencia de un polinucleótido que comprende detectar en una reacción de secuenciación la incorporación de tres tipos diferentes de conjugados de nucleótidos detectables en un polinucleótido y determinar la incorporación de un cuarto tipo de nucleótido basado en el patrón de detección de los tres tipos diferentes de nucleótidos detectables en el polinucleótido determinando con ello la secuencia de un polinucleótido, en el que la incorporación de tres tipos diferentes de conjugados de nucleótidos detectables se detecta a partir de un estado de señal; y en el que se detecta un primer nucleótido conjugado, unido a un primer fluoróforo, en un primer canal;
- se detecta un segundo nucleótido conjugado, unido a un segundo fluoróforo, en un segundo canal;
- 10 se detecta un tercer conjugado de nucleótidos en el primer canal y en el segundo canal, en el que el tercer conjugado de nucleótidos es una mezcla de conjugados que se diferencian por su unión a un tercer fluoróforo detectado en el primer canal o un cuarto fluoróforo detectado en el segundo canal;
- y en el que la incorporación del cuarto tipo de nucleótido se determina a partir de un estado oscuro.
- 2.** El método de la reivindicación 1, en el que los cuatro fluoróforos son diferentes restos fluorescentes.
- 15 **3.** El método de la reivindicación 1, en el que el primer fluoróforo es idéntico al tercer fluoróforo, y el segundo fluoróforo es idéntico al cuarto fluoróforo.
- 4.** El método de la reivindicación 2 ó 3, en el que la incorporación del cuarto nucleótido está determinada por el patrón de detección fluorescente de los restos fluorescentes, y en el que el patrón de detección de fluorescencia se determina mediante un primer y un segundo evento de formación de imágenes.
- 20 **5.** El método de la reivindicación 4, en el que el primer evento de formación de imágenes detecta un patrón de fluorescencia que es diferente del patrón de fluorescencia detectado por el segundo evento de formación de imágenes.
- 6.** El método de la reivindicación 5, en el que la incorporación de uno o más nucleótidos está determinada por la diferencia en el patrón de fluorescencia entre el primer y segundo eventos de formación de imágenes.
- 25 **7.** El método de la reivindicación 1, en el que uno o más conjugados de tipo nucleotídico comprenden además una o más secuencias enlazantes, y
- en el que una o más secuencias enlazantes comprenden uno o más de un enlazante escindible y un enlazante espaciador en el que el enlazante escindible comprende uno o más grupos de enlace escindibles seleccionados del grupo que consiste en un disulfuro, un diol, un diazo, un éster, una sulfona, una azida, un alilo y un éter silílico.
- 30 **8.** El método de la reivindicación 1, en el que los nucleótidos comprenden cuatro tipos de nucleótidos modificados en los que tres tipos de nucleótidos modificados están conjugados con uno o más restos de detección y uno o más enlazantes posicionados entre el nucleótido y uno o más restos de detección y en el que un cuarto tipo de nucleótido carece de un resto de detección, en el que la detección comprende:
- 35 a) detectar un patrón de incorporación de dichos nucleótidos modificados en una reacción de secuenciación capturando de este modo un primer patrón detectable,
- b) aplicar una o más composiciones a la reacción de secuenciación cambiando de este modo el primer patrón detectable, y
- c) detectar un segundo patrón detectable, y en el que
- d) se determina la secuencia de la muestra polinucleotídica en base a los patrones detectables.
- 40 **9.** El método de la reivindicación 8, en el que la incorporación del cuarto tipo de nucleótido está determinada por los patrones de detección de fluorescencia de los otros tres tipos de nucleótidos tanto en el primer como segundo patrones detectables.
- 10.** El método de la reivindicación 8, en el que uno o más enlazantes comprenden uno o más de un enlazante escindible y un enlazante espaciador.
- 45 **11.** El método de la reivindicación 8, en el que la aplicación de una o más composiciones añadidas a una reacción de secuenciación cambiando de este modo el primer patrón detectable comprende la adición de uno o más de un compuesto de escisión y un conjugado de pareja de unión-resto de detección.
- 12.** Una composición para uso en el método de la reivindicación 1, que comprende:
- a) un primer conjugado de nucleótido unido a un primer fluoróforo y detectable en un primer canal;

- b) un segundo conjugado de nucleótido unido a un segundo fluoróforo y detectable en un segundo canal;
- c) un tercer conjugado de nucleótidos, detectable en el primer canal y el segundo canal, en el que el tercer conjugado de nucleótidos es una mezcla de conjugados que se diferencian por estar unidos a un tercer fluoróforo, detectable en el primer canal, o un cuarto fluoróforo, detectable en el segundo canal; y

5 d) un cuarto conjugado de nucleótido que carece de un resto fluorescente.

13. La composición de la reivindicación 12, en la que los cuatro fluoróforos son restos fluorescentes diferentes.

14. La composición de la reivindicación 12, en la que el primer fluoróforo es idéntico al tercer fluoróforo, y el segundo fluoróforo es idéntico al cuarto fluoróforo.

FIGURA 1

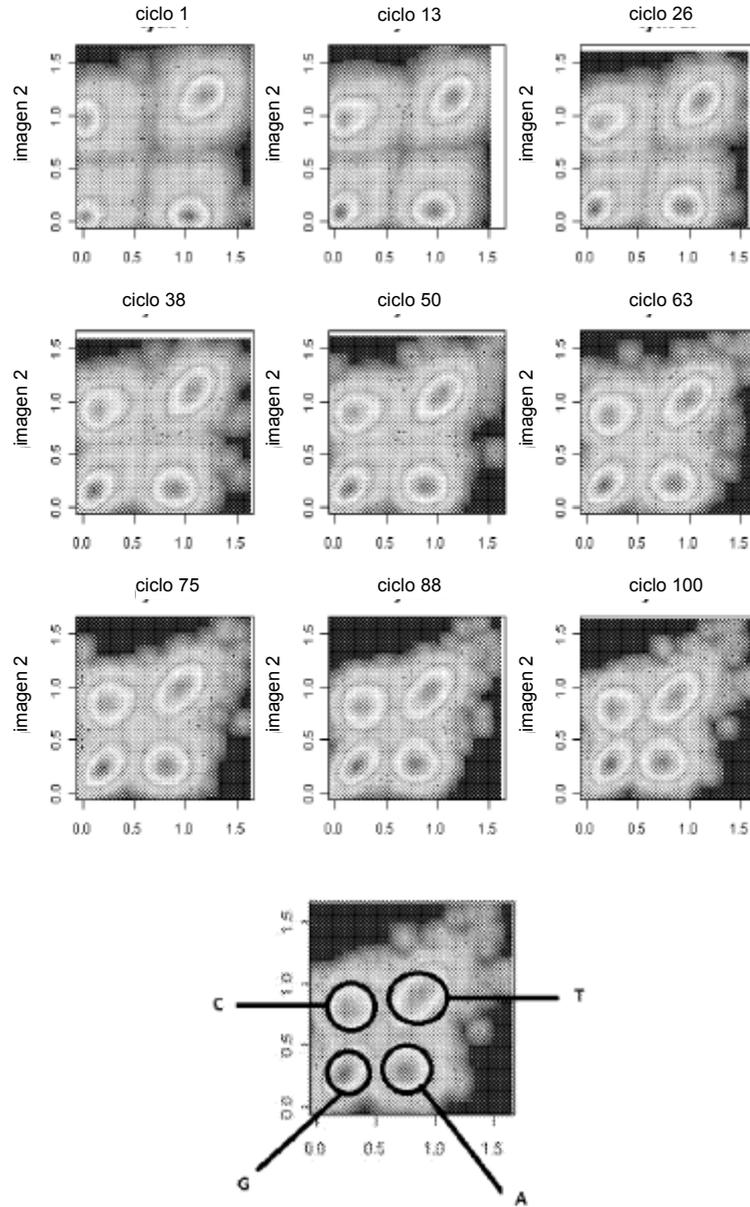


FIGURA 2

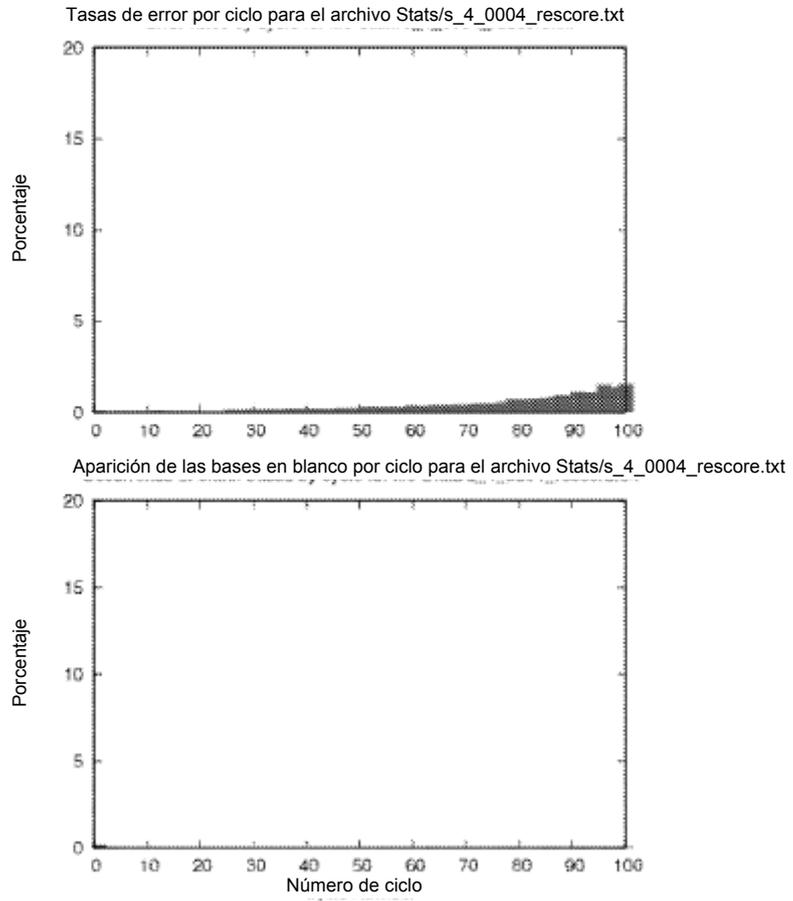
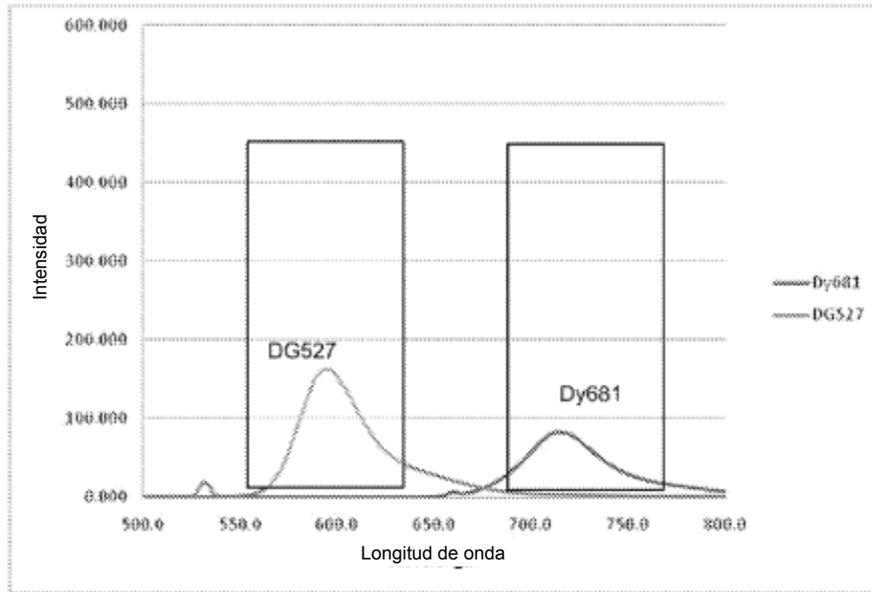


FIGURA 3

A



B

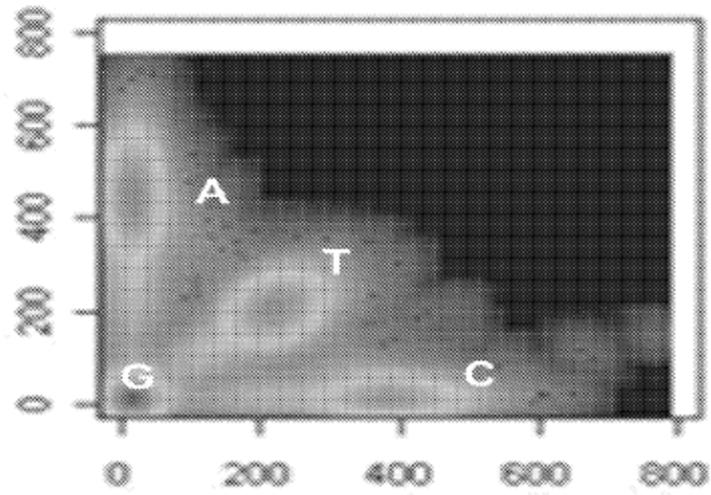
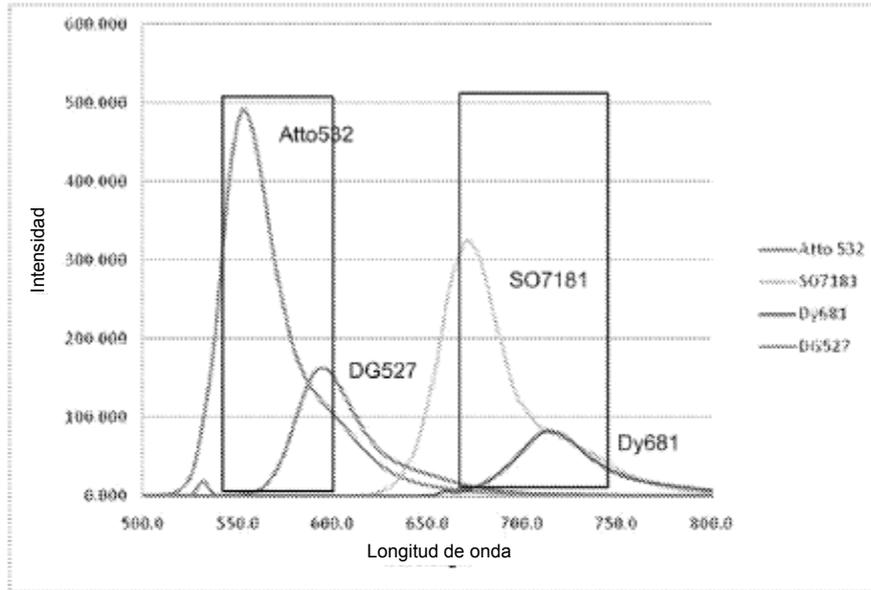


FIGURA 4

A



B

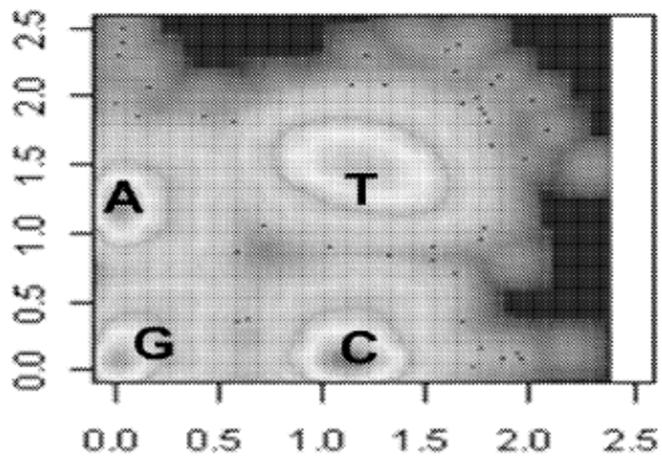
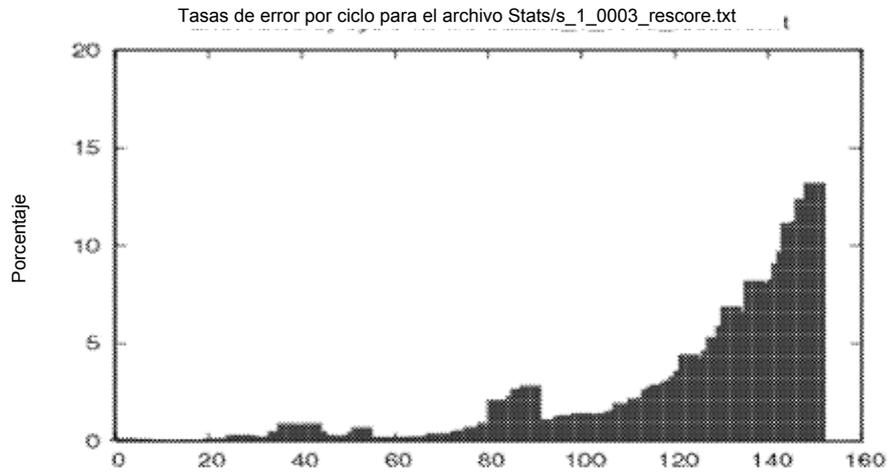


FIGURA 5

A



B

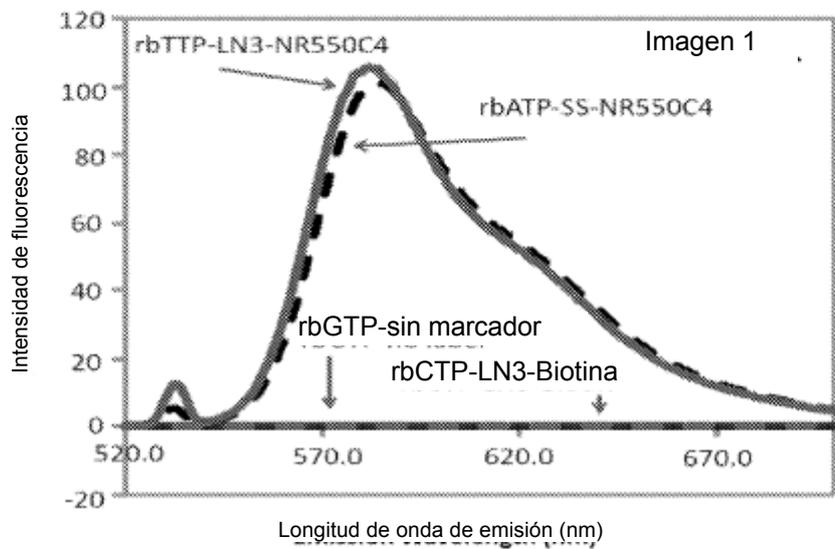
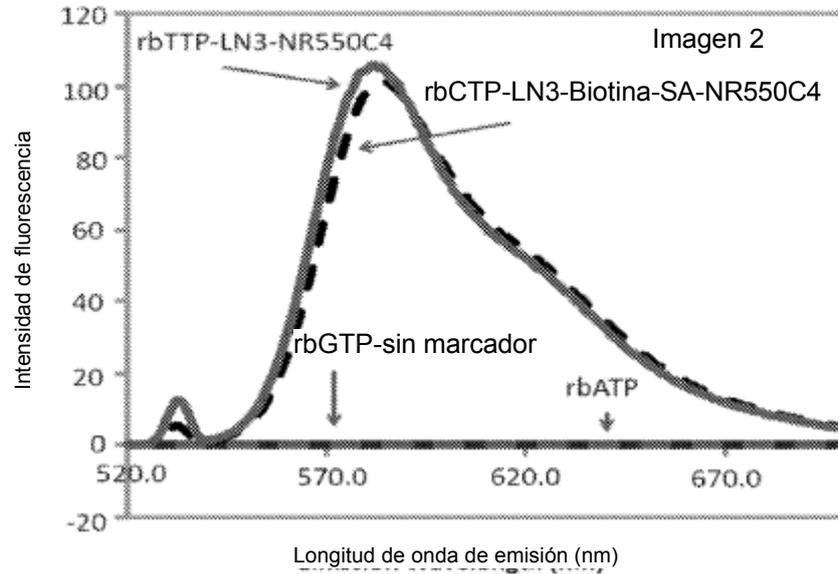


FIGURA 5 (cont)

C



D

