(19) OFICINA ESPAÑOLA DE PATENTES Y MARCAS

ESPAÑA

(11) Número de publicación: **2 876 432**

(51) Int. CI.:

C12Q 1/68 (2008.01)

(12) TRADUCCIÓN DE PATENTE EUROPEA T3

(86) Fecha de presentación y número de la solicitud internacional: **14.05.2015 PCT/US2015/030889**

(87) Fecha y número de publicación internacional: **19.11.2015 WO15175832**

(96) Fecha de presentación y número de la solicitud europea: **14.05.2015 E 15729968 (6)**

(97) Fecha y número de publicación de la concesión europea: **21.04.2021 EP 3143161**

(54) Título: **Técnicas de síntesis de ácidos nucleicos**

(30) Prioridad:

16.05.2014 US 201461994498 P

(45) Fecha de publicación y mención en BOPI de la traducción de la patente:
12.11.2021

(73) Titular/es:

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

(72) Inventor/es:

RONAGHI, MOSTAFA;
HE, MOLLY;
CHEN, CHENG-YAO;
PREVITE, MICHAEL y
BOWEN, SHANE

(74) Agente/Representante:

ELZABURU, S.L.P

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

Técnicas de síntesis de ácidos nucleicos

Antecedentes

La invención se refiere en general al campo de la síntesis de ácidos nucleicos y, más específicamente, a métodos y sistemas para sintetizar ácidos nucleicos y al control de calidad de dicha síntesis.

Las técnicas de ácidos nucleicos recombinantes han facilitado el estudio de genes aislados en una amplia gama de organismos. Por ejemplo, tales técnicas han permitido a los investigadores expresar un gen particular (p. ej., un gen de ratón) en un organismo hospedante (p. ej., una bacteria) para estudiar no solo los efectos del propio gen, sino también para estudiar sus productos de expresión. Si bien estas técnicas se han utilizado para estudiar ácidos nucleicos de origen natural, los investigadores también se han interesado en estudiar los efectos de mutaciones o cambios particulares en una secuencia de ácidos nucleicos. Por ejemplo, se pueden introducir mutaciones dirigidas en una secuencia de ácido nucleico que, a su vez, dan como resultado un cambio en una secuencia de proteína expresada. Luego, los investigadores pueden estudiar los efectos de tales mutaciones en las interacciones de proteínas de interés. Tales mutaciones pueden introducirse mediante técnicas basadas en PCR y las secuencias resultantes pueden expresarse en células hospedantes. Sin embargo, estas técnicas pueden implicar un análisis laborioso y una propagación de la célula hospedante que requiere mucho tiempo para determinar si una célula hospedante en particular ha incorporado la secuencia de interés. Además, aunque las técnicas recombinantes pueden ser adecuadas para generar simplemente mutaciones (p. ej., mutaciones puntuales, deleciones o sustituciones) de una secuencia de ácido nucleico preexistente, tales técnicas no son adecuadas para crear una molécula de ácido nucleico independiente del molde completamente sintética.

Algunas técnicas de síntesis de ácidos nucleicos son técnicas enzimáticas o de tipo químico, p. ej., síntesis química en fase sólida, que implican la adición de nucleótidos individuales entre sí en el orden deseado para formar una cadena de polinucleótidos, tal como un cebador de PCR de longitud limitada (p. ej., 30-50 nucleótidos). Sin embargo, las técnicas de síntesis de ácidos nucleicos adecuadas para formar secuencias relativamente cortas pueden ser lentas y/o propensas a errores cuando se usan para la síntesis de una cadena de polinucleótido larga (p. ej., más de 50 nucleótidos). Por ejemplo, las técnicas con una tasa de error de 1/500 nucleótidos generalmente pueden producir secuencias cortas sin errores de menos de 50 nucleótidos, pero pueden introducir varios errores cuando se utilizan para sintetizar secuencias de 500-2000 nucleótidos.

Las técnicas de síntesis de ácidos nucleicos se describen p. ej. en los documentos WO 2008/054543; US 2007/031942; US 2012/316086; WO 2012/050920; WO 2007/060456; WO 98/44151; Kim et al. (2012), *Nucleic Acid Research* 40: e140; y US 2010/216648.

Breve descripción

En el presente documento se proporcionan técnicas para sintetizar ácidos nucleicos con una eficacia y precisión mejoradas. Las técnicas descritas pueden incorporar retroalimentación de calidad en una o más fases de síntesis para facilitar la selección de una molécula de ácido nucleico con un nivel de calidad deseado (p. ej., sin errores en la secuencia de nucleótidos o errores con un intervalo de errores definido) de un grupo de moléculas sintetizadas producidas por un procedimiento de síntesis que se sospecha que genera errores. En ciertos aspectos de la descripción, las técnicas de síntesis de ácidos nucleicos pueden incorporar una pluralidad de fragmentos de ácidos nucleicos que se diseñan, se sintetizan por separado y luego se ensamblan. En implementaciones particulares, la retroalimentación de calidad puede basarse en los datos de secuenciación obtenidos para los fragmentos individuales y/o para la molécula de ácido nucleico ensamblada formada a partir de los fragmentos individuales. Por ejemplo, se pueden sintetizar y secuenciar una pluralidad de fragmentos individuales, recogiéndose únicamente los fragmentos que tienen la secuencia correcta para su posterior ensamblaje en una molécula de ácido nucleico. En los casos en los que se sintetizan por separado múltiples copias de cada fragmento, se puede recoger una copia correcta mientras se evitan o incluso se descartan las copias incorrectas. Las moléculas ensambladas también pueden secuenciarse para seleccionar moléculas con un perfil de errores de secuencia deseado. De nuevo, se pueden recoger solo moléculas de ácido nucleico ensambladas que tengan la secuencia deseada.

En otro aspecto de la descripción, los fragmentos se pueden sintetizar y ensamblar sin secuenciar los fragmentos individuales antes del ensamblaje. Aquí, el control de calidad puede proporcionarse mediante información de secuenciación para las moléculas ensambladas. Sin embargo, incluso en tal aspecto de la descripción, se pueden implementar otras etapas de control de calidad. Por ejemplo, cada procedimiento de síntesis de fragmentos individuales puede tener diferentes rendimientos, dependiendo del tamaño y/o secuencia del fragmento. La etapa de combinación de fragmentos puede explicar los diferentes rendimientos por combinación de los fragmentos en concentraciones aproximadamente iguales, p. ej., diluyendo ciertos fragmentos o sometiendo fragmentos de menor rendimiento a etapas de amplificación adicionales. De esta manera, la etapa de síntesis de fragmentos puede incorporar información de amplificación como parte de la retroalimentación de calidad. Además, la información del rendimiento y/o secuenciación de los fragmentos y/o moléculas ensambladas se puede proporcionar como retroalimentación a la fase de diseño. En un aspecto de la descripción, un conjunto de diseños de fragmentos

particulares puede tener un rendimiento pobre o tasas de error bajas a nivel de fragmento y/o molécula ensamblada. Dicha información puede proporcionarse de vuelta a la fase de diseño e implementarse como control de calidad para el diseño de fragmentos futuros.

Las técnicas descritas pueden incorporar la síntesis de ácido nucleico (p. ej., "escritura") junto con la secuenciación (p. ej., "lectura") en una única plataforma (o una pluralidad de plataformas acopladas) para facilitar el análisis. Por lo tanto, se proporcionan en el presente documento aparatos y métodos para leer y escribir ácidos nucleicos. Por consiguiente, las técnicas descritas pueden usarse junto con ciertas plataformas de secuenciación. En un aspecto de la descripción, ciertas plataformas de secuenciación que utilizan tecnologías que pueden realizarse de una manera dependiente del molde pueden incorporar una etapa de síntesis anterior para generar el molde en la plataforma in situ. Este molde sintetizado, correspondiente a una secuencia de fragmentos deseada, luego se puede amplificar y secuenciar para proporcionar retroalimentación con respecto al error, calidad, diseño o similares. Por consiguiente, a diferencia de ciertas técnicas de síntesis de ácido nucleico recombinante o PCR, las técnicas descritas pueden implementarse de una manera inicialmente sin molde con una secuencia sintética deseada diseñada in silico.

La invención se refiere a un método que se proporciona para sintetizar un ácido nucleico. El método puede incluir las etapas de (i) proporcionar una pluralidad de fragmentos de ácido nucleico que tienen secuencias que se solapan, en donde la pluralidad de fragmentos de ácido nucleico tiene secuencias complementarias de al menos otro fragmento de la pluralidad de fragmentos de ácido nucleico, en donde un primer fragmento de la pluralidad de fragmentos de ácido nucleico tiene una primera secuencia de adaptador escindible y un extremo 5' de la secuencia diana, en donde un extremo del primer fragmento comprende el primer adaptador escindible, y en donde un último fragmento de la pluralidad de fragmentos de ácido nucleico tiene un extremo 3' de la secuencia diana, en donde un extremo del último fragmento comprende un segundo adaptador escindible; en donde el proporcionar una pluralidad de fragmentos de ácido nucleico que se solapan comprende: (a) generar una pluralidad del primer fragmento, que comprende: proporcionar una pluralidad de oligonucleótidos inmovilizados sobre un sustrato de síntesis mediante hibridación con una pluralidad de primeros cebadores inmovilizados complementarios de los extremos 5' de los oligonucleótidos, en donde cada oligonucleótido respectivo comprende el primer adaptador escindible en un extremo 5', extender los oligonucleótidos para generar polinucleótidos extendidos, en donde la extensión comprende incorporar una pluralidad de nucleótidos o ácidos nucleicos individuales en los extremos 3' de los oligonucleótidos de la pluralidad de oligonucleótidos, unir un adaptador amplificador en los extremos 3' de los polinucleótidos extendidos que tienen la primera secuencia terminal, reasociar los extremos 3' de los polinucleótidos extendidos con el sustrato de síntesis mediante una pluralidad de segundos cebadores inmovilizados complementarios de los extremos 3' de los polinucleótidos extendidos para formar puentes, amplificar los puentes para formar una pluralidad de grupos amplificados de los puentes, secuenciar la pluralidad de grupos amplificados de los puentes para determinar la presencia de la primera secuencia terminal en uno o más de la pluralidad de grupos amplificados de los puentes y una tasa de error de síntesis, y recoger selectivamente uno o más de la pluralidad de grupos amplificados de los puentes basándose en la tasa de error de síntesis; y

(b) combinar los grupos amplificados recogidos que tienen la primera secuencia terminal con una segunda pluralidad de grupos amplificados, en donde al menos una parte de la segunda pluralidad de grupos amplificados comprende una secuencia complementaria de solo una parte de la primera secuencia terminal, obteniendo así una pluralidad de fragmentos de ácido nucleico; (ii) inmovilizar el primer fragmento sobre un sustrato con un primer cebador inmovilizado complementario de la primera secuencia de adaptador escindible; (iii) ensamblar la pluralidad de fragmentos de ácido nucleico en una molécula de polinucleótido ensamblada mediante hibridación de las secuencias complementarias, en donde la molécula de polinucleótido ensamblada es parcialmente bicatenaria y comprende huecos en ambas cadenas, y en donde la molécula de polinucleótido ensamblada está inmovilizada sobre el sustrato a través del primer cebador inmovilizado complementario del primer adaptador escindible; (iv) amplificar la molécula de polinucleótido ensamblada sobre el sustrato para generar un grupo amplificado que tiene amplicones de la molécula de polinucleótido ensamblada sobre el sustrato; (v) secuenciar el grupo amplificado para determinar una tasa de error de síntesis; y (vi) recoger selectivamente el grupo amplificado basándose en la tasa de error de síntesis.

También se describe un método para sintetizar fragmentos de ácido nucleico que incluye las etapas de (a) proporcionar una pluralidad de oligonucleótidos inmovilizados sobre un sustrato mediante hibridación con una pluralidad de primeros cebadores inmovilizados complementarios de los extremos 5' de los oligonucleótidos, en donde cada oligonucleótido respectivo comprende una primera secuencia de adaptador en un extremo 5'; (b) extender los oligonucleótidos para generar polinucleótidos extendidos que corresponden a una secuencia del fragmento, en donde la extensión incluye incorporar una pluralidad de nucleótidos o ácidos nucleicos individuales en un extremo 3' de los respectivos oligonucleótidos de la pluralidad de oligonucleótidos para generar los polinucleótidos extendidos que tienen la secuencia del fragmento; (c) unir una segunda secuencia de adaptador a los extremos 3' de los polinucleótidos extendidos que tienen la secuencia del fragmento; (d) reasociar los extremos 3' de los polinucleótidos extendidos con el sustrato mediante una pluralidad de segundos cebadores inmovilizados complementarios de los extremos 3' de los polinucleótidos extendidos para formar puentes; (e) amplificar los puentes para formar una pluralidad de grupos amplificados; (f) secuenciar la pluralidad de grupos amplificados para determinar si uno o más de la pluralidad de grupos amplificados incluye una secuencia del fragmento de ácido nucleico; (g) recoger el uno o más de la pluralidad de grupos amplificados si una secuencia del uno o más de la pluralidad de grupos amplificados incluye la secuencia del fragmento de ácido nucleico o es complementaria de la secuencia del fragmento de ácido nucleico para grupos amplificados recogidos generados que tienen la secuencia de fragmentos; y (h) combinar los grupos amplificados

recogidos que tienen la secuencia del fragmento con una segunda pluralidad de grupos amplificados, en donde al menos una parte de la segunda pluralidad de grupos amplificados comprende una secuencia complementaria de solo una parte de la secuencia del fragmento.

En algunos aspectos de la descripción, un método para sintetizar un ácido nucleico puede incluir las etapas de (a) proporcionar una pluralidad de secuencias diana basadas en una secuencia del ácido nucleico, en donde una combinación de la pluralidad de secuencias diana forma la secuencia del ácido nucleico; (b) proporcionar una pluralidad de oligonucleótidos cebadores inmovilizados sobre un sustrato; (c) extender los oligonucleótidos cebadores de una manera monocatenaria basándose en la pluralidad de secuencias diana en presencia de una primera polimerasa para generar una pluralidad de polinucleótidos de fragmentos; (d) proporcionar un primer reactivo de secuenciación a la pluralidad de polinucleótidos de fragmentos en presencia de una segunda polimerasa, en donde el primer reactivo de secuenciación tiene uno o más monómeros de nucleótidos, y en donde el uno o más monómeros de nucleótidos forman una pluralidad de polinucleótidos complementarios de al menos una parte de la pluralidad de polinucleótidos de fragmentos; (e) proporcionar un segundo reactivo de secuenciación a la pluralidad de polinucleótidos de fragmentos, en donde el segundo reactivo de secuenciación incluye al menos un monómero de nucleótido, en donde el al menos un monómero de nucleótido del segundo reactivo de secuenciación incluye un resto de terminación reversible, y en donde el segundo reactivo de secuenciación se proporciona después de proporcionar el primer reactivo de secuenciación, de modo que se obtiene una secuencia de cada uno de la pluralidad de polinucleótidos de fragmentos; (f) recoger la pluralidad de polinucleótidos de fragmentos basándose en la comparación de la secuencia de cada uno de la pluralidad de polinucleótidos de fragmentos con la secuencia diana respectiva; y (g) ensamblar la pluralidad de polinucleótidos de fragmentos para generar un polinucleótido ensamblado.

Esta descripción proporciona un sistema para sintetizar un ácido nucleico. El sistema puede incluir (a) un sustrato que tiene una pluralidad de canales de flujo, incluyendo cada uno de la pluralidad de canales de flujo una pluralidad de oligonucleótidos cebadores inmovilizados; (b) un dispositivo de tipo procesador que almacena instrucciones ejecutables y acoplado al sustrato, en donde las instrucciones ejecutables están configuradas para: (i) recibir una o más señales de síntesis para cada canal de flujo respectivo indicativo de una presencia de una primera polimerasa que genera una pluralidad respectiva de polinucleótidos de fragmentos monocatenarios basados en una de una pluralidad de secuencias diana; (ii) controlar la entrada de un primer reactivo de secuenciación, un segundo reactivo de secuenciación y una segunda polimerasa en los canales de flujo basándose en una o más señales de síntesis, en donde el primer reactivo de secuenciación incluye uno o más monómeros de nucleótidos, en donde el uno o más monómeros de nucleótidos forman una pluralidad de polinucleótidos complementarios de al menos una parte de la pluralidad de polinucleótidos de fragmentos, en donde el segundo reactivo de secuenciación incluye al menos un monómero de nucleótidos, y en donde el al menos un monómero de nucleótidos del segundo reactivo de secuenciación tiene un resto de terminación reversible; (iii) recibir una o más señales de secuenciación para cada canal de flujo respectivo indicativas de la presencia del primer reactivo de secuenciación y el segundo reactivo de secuenciación en presencia de la segunda polimerasa; (iv) determinar una secuencia de cada uno de la pluralidad de polinucleótidos de fragmentos basándose en una o más señales de secuenciación; y (v) proporcionar una indicación relacionada con una comparación de la secuencia de cada uno de la pluralidad de polinucleótidos de fragmentos con la secuencia diana respectiva; y (c) una celda de ensamblaje configurada para reunir cada uno de la pluralidad de polinucleótidos de fragmentos de cada uno de la pluralidad de canales de flujo y ensamblar cada uno de la pluralidad de polinucleótidos de fragmentaos para generar un polinucleótido ensamblado.

Se describe además un método para sintetizar un ácido nucleico, que incluye las etapas de: (a) proporcionar una secuencia de polinucleótido de partida que incluye una secuencia diana; (b) recibir o acceder a información representativa de secuencias de una pluralidad de fragmentos de ácido nucleico que se solapan basados en la secuencia diana, en donde los fragmentos están diseñados de tal manera que los fragmentos que incluyen la secuencia diana comprenden discontinuidades en una dirección de 5' a 3', y en donde las discontinuidades entre los fragmentos de ácido nucleico se producen cuando los fragmentos de ácido nucleico se hibridan con una cadena complementaria a la secuencia diana; (c) recibir información de secuenciación de grupos amplificados, en donde los grupos amplificados se amplifican a partir de puentes generados a partir de fragmentos sintetizados que tienen la secuencia de la pluralidad de fragmentos de ácido nucleico; y (d) determinar una calidad de uno o más de la pluralidad de grupos amplificados basado en la información de secuenciación.

Dibujos

Estas y otras características, aspectos y ventajas de la presente invención se entenderán mejor cuando se lea la siguiente descripción detallada con referencia a los dibujos adjuntos en los que los mismos caracteres representan partes iguales en todos los dibujos, en donde:

La fig. 1 es una descripción esquemática de una técnica para generar una molécula de polinucleótido sintética, que, en ciertos aspectos de la descripción, puede usarse a su vez para crear otros ácidos nucleicos (p. ej., por transcripción) y/o proteínas (p. ej., por traducción) según aspectos de la presente descripción;

La fig. 2 es un diagrama de bloques de un sistema que puede usarse junto con la técnica de la fig. 1;

La fig. 3 es una técnica de síntesis de fragmentos para generar un fragmento de polinucleótido monocatenario

independiente de molde de una molécula de polinucleótido sintética con retroalimentación de secuenciación según aspectos de la presente descripción;

La fig. 4 es una técnica de síntesis y ensamblaje de fragmentos para ensamblar múltiples fragmentos de polinucleótidos de una molécula de polinucleótido sintética con retroalimentación de secuenciación para los fragmentos y el polinucleótido sintético ensamblado según aspectos de la presente descripción;

La fig. 5 es una técnica de ensamblaje de fragmentos de polinucleótido con retroalimentación de secuenciación según aspectos de la presente descripción;

La fig. 6 es una celda de flujo para sintetizar diferentes secuencias de fragmentos de polinucleótido en paralelo para su posterior ensamblaje en una molécula de polinucleótido sintética según aspectos de la presente descripción;

La fig. 7 es un aspecto de la descripción de un casete de lectura y escritura de microfluidos para la síntesis, amplificación y secuenciación de plataforma única de una molécula de polinucleótido sintética según aspectos de la presente descripción;

La fig. 8 es un diagrama de flujo de una técnica para generar un fragmento de polinucleótido independiente de molde de una molécula de polinucleótido sintética con retroalimentación de secuenciación según aspectos de la presente descripción;

La fig. 9 es un diagrama de flujo de una técnica para generar una molécula de polinucleótido sintética ensamblada con retroalimentación de secuenciación según aspectos de la presente descripción; y

La fig. 10 es un diagrama de flujo de una técnica de retroalimentación de calidad para el diseño de polinucleótidos sintéticos basada en la retroalimentación de secuenciación del polinucleótido sintético ensamblado de acuerdo con aspectos de la presente descripción.

Descripción detallada

1. Definiciones

Como se usa en el presente documento, las formas singulares "un", "y" y "el", "la" incluyen referentes plurales a menos que el contexto indique claramente lo contrario. Así, por ejemplo, la referencia a "una celda" incluye una pluralidad de tales celdas, y así sucesivamente. Todos los términos técnicos y científicos usados en el presente documento tienen el mismo significado que el comúnmente entendido por un experto en la técnica a la que pertenece esta invención, a menos que se indique claramente lo contrario.

Como se usa en el presente documento, "amplificación" y/o sus variantes gramaticales se refieren a cualquier método para aumentar el número de copias de una secuencia de nucleótidos usando un ácido nucleico molde que tiene la secuencia o su complemento. La amplificación se puede realizar, por ejemplo, con una polimerasa o ligasa. El método puede llevarse a cabo in vitro, por ejemplo, en una celda de flujo u otro recipiente fluídico. En algunos aspectos de la descripción, la amplificación de ácido nucleico da como resultado la incorporación de nucleótidos en una molécula de ADN y/o ARN o cebador formando así una nueva molécula complementaria de un ácido nucleico diana. La molécula de ácido nucleico formada, su complemento o ambos pueden usarse como moldes para sintetizar moléculas de ácido nucleico adicionales. Como se usa en el presente documento, una reacción de amplificación puede consistir en al menos una o más rondas de replicación. Las reacciones de amplificación de ácidos nucleicos incluyen, por ejemplo, reacciones en cadena de la polimerasa (PCR), amplificación con cebado aleatorio, amplificación de puente, amplificación de círculo rodante (RCA), reacción en cadena de ligasa y otros métodos conocidos en la técnica. Se describen ejemplos de métodos de amplificación útiles en la publicación de solicitud de patente de EE.UU. Nº 2005/0037393 A1. Una reacción de PCR puede consistir en 5 a 100 "ciclos" de desnaturalización y síntesis de una molécula de ADN.

La amplificación se puede llevar a cabo en solución o en fase sólida. La amplificación de puente es un método particularmente útil de amplificación en fase sólida. Se describen ejemplos de amplificación de puente en Bentley et al., *Nature* 456:53-59 (2008); patentes de EE.UU. Nº 5.641.658 o 7.115.400; o en las Pub. de Pat. de EE.UU. Nº 2002/0055100 A1, 2004/0096853 A1, 2004/0002090 A1, 2007/0128624 A1 o 2008/0009420 A1. Otros métodos que pueden llevarse a cabo en solución o en fase sólida incluyen, por ejemplo, PCR, RCA, MDA y otros métodos de amplificación expuestos en el presente documento o conocidos en la técnica. Los ejemplos de RCA se describen en Lizardi y al., *Nat. Genet*. 19: 225-232 (1998) o Pub. de Pat. de EE.UU. Nº 2007/0099208 A1. También es útil la amplificación por desplazamiento múltiple (MDA), por ejemplo, usando un producto de RCA (es decir, un amplicón de RCA) como molde. Los métodos de ejemplo de MDA se describen en las patentes de EE.UU. Nº 6.124.120; 5.871.921; o EP 0868530 B1.

Como se usa en el presente documento, "complementario" o "complementariedad" y/o sus variantes gramaticales se refieren al grado de apareamiento de bases o hibridación entre ácidos nucleicos. Por ejemplo, como saben los expertos en la técnica, la adenina (A) puede formar enlaces de hidrógeno o pares de bases con timina (T) o uracilo (U) y la guanina (G) puede formar enlaces de hidrógeno o pares de bases con citosina (C). Por tanto, la A es complementaria

de T o U y la G es complementaria de C. Estos son los pares de bases estándar de "Watson-Crick" que se encuentran en la gran mayoría de híbridos de ADN y ARN in vivo. La complementariedad, cuando se usa en referencia a una región bicatenaria de ácido nucleico, puede ser completa cuando todas las bases en la región bicatenaria forman pares de bases. Alternativamente, la complementariedad puede ser "parcial", en la que solo algunas de las bases en la región bicatenaria están emparejadas de acuerdo con las reglas de emparejamiento de bases. El grado de complementariedad entre las cadenas de ácidos nucleicos tiene un efecto en la eficacia y la fuerza de la hibridación entre las cadenas de ácidos nucleicos. Las secuencias "complementarias" también pueden incluir, o estar formadas completamente a partir de pares de bases que no son de Watson-Crick y/o pares de bases formados a partir de nucleótidos modificados y no naturales, en la medida en que los requisitos anteriores con respecto a su capacidad de hibridación se cumplan.

Como se usa en el presente documento, "tasa de error" y/o sus variantes gramaticales se refieren al número relativo de nucleótidos erróneos en una secuencia de polinucleótido, en comparación con el número de nucleótidos en una secuencia del polinucleótido diana. El nucleótido erróneo se refiere al nucleótido en la secuencia de polinucleótido que es diferente del nucleótido correspondiente en la secuencia del polinucleótido diana cuando se alinean la secuencia de polinucleótido y la secuencia del polinucleótido diana. La tasa de error se puede presentar en una variedad de formatos que incluyen, pero no se limitan a, una proporción o fracción del número de nucleótidos erróneos en una secuencia al número de nucleótidos totales en una secuencia particular. Por ejemplo, cuando un polinucleótido diana tiene una secuencia de AAAAA, y un polinucleótido sintetizado basado en el polinucleótido diana tiene una secuencia de AAAAG, la tasa de error es 1/5 o 20%.

Como se usa en el presente documento, "hibridación" y/o sus variantes gramaticales se refiere a la interacción física de cadenas de polinucleótidos complementarias (incluidas parcialmente complementarias) por la formación de enlaces de hidrógeno entre nucleótidos complementarios cuando las cadenas están dispuestas antiparalelas entre sí. La hibridación y la fuerza de la hibridación (p. ej., la fuerza de la asociación entre polinucleótidos) se ve afectada por muchos factores bien conocidos en la técnica, que incluyen el grado de complementariedad entre los polinucleótidos y la rigurosidad de las condiciones implicadas, que se ve afectada por condiciones tales como la concentración de sales, la presencia de otros componentes (p. ej., la presencia o ausencia de polietilenglicol), la molaridad de las cadenas de hibridación y el contenido de G+C de las cadenas de polinucleótidos, todo lo cual da como resultado una temperatura de fusión característica (Tm) del híbrido formado. Los términos "hibridación (hibridar)" y "unión", cuando se usan en referencia a ácidos nucleicos, se pueden usar indistintamente y pueden referirse al proceso por el cual cadenas simples de secuencias de ácidos nucleicos forman segmentos de doble hélice mediante enlaces de hidrógeno entre nucleótidos complementarios. "Híbrido", "dúplex" y "complejo", cuando se usan en referencia a ácidos nucleicos, también se pueden usar indistintamente en el presente documento refiriéndose a una molécula de ácido nucleico bicatenario formada por hibridación (p. ej., especies de ADN-ADN, ADN-ARN y ARN-ARN).

Como se usa en el presente documento, "ácido nucleico" y/o sus variantes gramaticales incluye polímeros de desoxirribonucleótidos o ribonucleótidos en forma monocatenaria o bicatenaria. El término puede incluir polímeros que están compuestos de monómeros (nucleótidos) que contienen un azúcar, fosfato y una base que es una purina o pirimidina. A menos que se limite específicamente, el término abarca ácidos nucleicos que contienen análogos de nucleótidos naturales que tienen propiedades de unión similares a las del ácido nucleico de referencia y/o ácidos nucleicos que tienen cadenas principales no nativas tales como ácidos nucleicos proteicos. En variantes particulares, los análogos de nucleótidos también se metabolizan de una manera similar a los nucleótidos de origen natural. La referencia a una secuencia de ácido nucleico particular puede referirse implícitamente a sus secuencias complementarias así como a la secuencia de referencia indicada explícitamente. Por ejemplo, se entenderá que la enseñanza relacionada con la amplificación de una secuencia de ácido nucleico incluye variantes donde se producen una o más copias de la secuencia, su complemento o ambos, a menos que se indique explícitamente lo contrario.

Como se usa en el presente documento, "nucleótido" y/o sus variantes gramaticales se refiere a una combinación de base-azúcar-fosfato. Los nucleótidos son unidades monoméricas de una secuencia de ácido nucleico (ADN y ARN). El término nucleótido incluye trifosfato de ribonucleósido ATP, UTP, CTP, GTP y trifosfato de desoxirribonucleósido tales como dATP, dCTP, dITP, dUTP, dGTP, dTTP o derivados de los mismos. "Nucleótido", como se usa en el presente documento, también se refiere a trifosfatos de didesoxirribonucleósidos (ddNTP) y sus derivados que incluyen, pero no se limitan a ddATP, ddCTP, ddGTP, ddITP y ddTTP. En un aspecto de la presente descripción, un "nucleótido" puede estar sin marcar o marcado de forma detectable mediante técnicas bien conocidas. Los marcadores detectables incluyen, pero no se limitan a isótopos radiactivos, marcadores fluorescentes, marcadores quimioluminiscentes, marcadores bioluminiscentes y marcadores enzimáticos.

Como se usa en el presente documento, "fragmento de ácido nucleico" y/o sus variantes gramaticales se refiere a una molécula sintética o natural que comprende una secuencia de nucleótidos unidos covalentemente que es más corta que un ácido nucleico de referencia. Los nucleótidos pueden unirse mediante un enlace fosfodiéster entre la posición 3' de la desoxirribosa o ribosa de un nucleótido y la posición 5' de la desoxirribosa o ribosa del nucleótido adyacente. Los fragmentos de ácido nucleico pueden incluir bases naturales (p. ej., A, G, C, T o U) o modificadas (p. ej., 7-desaza-guanosina, inosina). Además, las bases en los fragmentos de ácido nucleico pueden unirse mediante un enlace distinto del enlace fosfodiéster, siempre que no interfiera con la hibridación del fragmento de ácido nucleico. Por tanto, los fragmentos de ácido nucleico pueden ser ácidos nucleicos peptídicos en los que una o más de las bases constituyentes están unidas por enlaces peptídicos en lugar de enlaces fosfodiéster. Los fragmentos de ácido nucleico pueden ser

monocatenarios o bicatenarios. Aunque se puede crear un fragmento de ácido nucleico a partir de un ácido nucleico de referencia más largo, se entenderá que el término "fragmento" pretende referirse a las longitudes comparativas de dos ácidos nucleicos y no necesariamente al origen o al método de creación de cualquiera de los dos ácidos nucleicos. Normalmente, un fragmento de ácido nucleico tendrá una secuencia que es idéntica o perfectamente complementaria a una parte de un ácido nucleico de referencia. Sin embargo, en algunas variantes, el fragmento puede ser menos que perfectamente homólogo o complementario de la parte de la secuencia de referencia, por ejemplo, siendo al menos 80%, 85%, 90%, 95%, 99% o 99,9% homólogo o complementario de la parte de la secuencia de referencia.

Como se usa en el presente documento, "cebador" y/o sus variantes gramaticales se refieren a un oligonucleótido monocatenario que se extiende por formación de enlace covalente de monómeros de nucleótidos durante la amplificación o polimerización de una molécula de ácido nucleico. Un cebador se puede producir de forma natural como en una digestión de restricción purificada o producir sintéticamente. Normalmente, un cebador puede actuar como punto de inicio de la síntesis de ácidos nucleicos dependiente de la muestra. El cebador puede ser monocatenario o bicatenario y, en aspectos particulares de la descripción, debe ser suficientemente largo para el cebado de la síntesis del producto de extensión deseado en presencia de la polimerasa elegida. La longitud exacta del cebador dependerá de muchos factores, incluidas las temperaturas de hibridación y polimerización, la fuente del cebador y el método utilizado. Por ejemplo, un cebador puede comprender aproximadamente al menos 15-75 nucleótidos, aunque puede contener menos o más nucleótidos. Los factores implicados en la determinación de la longitud adecuada del cebador para una aplicación particular serán fácilmente conocidos o determinables por un experto en la técnica basándose en las enseñanzas expuestas en el presente documento.

El término "inmovilizado", como se usa en el presente documento, pretende abarcar la unión directa o indirecta a un soporte sólido por enlace(s) covalente(s) o no covalente(s). En ciertos aspectos de la descripción, se puede usar unión covalente, pero generalmente todo lo que se requiere es que las moléculas (por ejemplo, ácidos nucleicos) permanezcan enlazadas o unidas a un soporte en las condiciones en las que se pretende usar el soporte, para ejemplo en aplicaciones que requieren amplificación y/o secuenciación de ácidos nucleicos. Normalmente, los oligonucleótidos que se van a utilizar como oligonucleótidos de captura u oligonucleótidos de amplificación se inmovilizan de manera que un extremo 3' esté disponible para la extensión enzimática y al menos una parte de la secuencia sea capaz de hibridar con una secuencia complementaria. La inmovilización puede producirse por hibridación con un oligonucleótido unido a la superficie, en cuyo caso el oligonucleótido o polinucleótido inmovilizado puede estar en la orientación 3'-5'. Alternativamente, la inmovilización puede producirse por medios distintos de la hibridación por apareamiento de bases, tal como la unión covalente expuesta anteriormente. En algunos aspectos de la descripción, las moléculas de ácido nucleico se pueden unir a un soporte sólido u otra superficie a través de una capa de gel como se describe, por ejemplo, en la solicitud de patente de EE.UU. Nº de Ser. 13/784.368 y publicación de solicitud de patente de EE.UU. Nº 2011/0059865 A1.

Como se usa en el presente documento, "soporte sólido", "fase sólida", "soporte" y "sustrato" y/o sus variantes gramaticales se refieren a cualquier material que proporcione una estructura sustancialmente rígida o una estructura que conserve su forma en lugar de adoptar la forma de un recipiente en el que se coloca. El material puede tener una superficie a la que se puede unir otro material, incluyendo, pero no limitado a soportes lisos (p. ej., superficies de metal, vidrio, plástico, silicio y cerámica), así como materiales texturizados y porosos. Los materiales del sustrato incluyen, pero no se limitan a acrílicos, carbono (p. ej., grafito, fibra de carbono), celulosa (p. ej., acetato de celulosa), cerámicos, vidrio de poro controlado, polisacáridos reticulados (p. ej., agarosa o SEPHAROSE®), geles, vidrio (p. ej., vidrio modificado o funcionalizado), oro (p. ej., Au (111) atómicamente liso), grafito, vidrios inorgánicos, polímeros inorgánicos, látex, óxidos metálicos (p. ej., SiO2, TiO2, acero inoxidable), metaloides, metales (p. ej., Au (111) atómicamente liso), mica, sulfuros de molibdeno, nanomateriales (p. ej., nanoláminas de grafito pirolítico altamente orientado (HOPG)), nitrocelulosa, NYLON®, haces de fibra óptica, polímeros orgánicos, papel, plásticos, polacriloilmorfolida, poli(4-metilbuteno), poli(tereftalato de etileno), poli(butirato de vinilo), polibutileno, polidimetilsiloxano (PDMS), polietileno, poliformaldehído, polimetacrilato, polipropileno, polisacáridos, poliestireno, poliuretanos, poli(difluoruro de vinilideno), cuarzo, rayón, resinas, cauchos, material semiconductor, sílice, silicio (p. ej., silicio de superficie oxidada), sulfuro y TEFLON®.

Los sustratos pueden tener una superficie plana pero no es necesario que sean planos y pueden incluir cualquier tipo de forma, incluidas formas esféricas (p. ej., perlas), formas porosas (p. ej., geles) o formas cilíndricas (p. ej., fibras). Los materiales unidos a soportes sólidos se pueden unir a cualquier parte del soporte sólido (p. ej., se pueden unir a una parte interior de un material de soporte sólido poroso).

Los sustratos pueden tener patrones, donde un patrón (p. ej., manchas, almohadillas, pocillos, postes, rayas, remolinos, líneas, triángulos, rectángulos, círculos, arcos, cruces, cuadriculados, diagonales, flechas, cuadrados o tramas cruzadas), se graba, imprime, trata, esboza, corta, talla, cincela, marca, fija, estampa, reviste, graba en relieve, incrusta o coloca en capas sobre un sustrato. El patrón puede comprender una o más regiones de escisión o regiones modificadas sobre el sustrato.

Un material biológico está "unido" a un sustrato cuando está asociado o ligado al sustrato sólido a través de una interacción química o física estable. En algunos aspectos preferidos de la descripción, la unión se realiza a través de un enlace covalente. Sin embargo, no es necesario que las uniones sean covalentes o permanentes. En un aspecto de la descripción, los materiales se unen a un sustrato mediante una "molécula espaciadora" o "grupo conector". Tales

moléculas espaciadoras son moléculas que tienen una primera parte que se une al material biológico y una segunda parte que se une al sustrato. Así, cuando se une al sustrato, la molécula espaciadora es intermedia entre el sustrato y los materiales biológicos, pero está unida a ambos. Los métodos para unir material biológico (p. ej., ácido nucleico, receptor de ligando de afinidad, enzima, generador químico de radicales hidroxilo) a un sustrato son bien conocidos en la técnica e incluyen, pero no se limitan a acoplamiento químico.

Como se usa en el presente documento, "reactivo de secuenciación" y sus equivalentes gramaticales se refiere a una composición, tal como una solución, que comprende uno o más reactivos, catalizadores o marcadores u otro analito que participa en una reacción para determinar el orden de los monómeros (p. ej., nucleótidos) en un polímero (p. ej., ácido nucleico). En algunos aspectos de la descripción, tales como un aspecto de secuenciación por síntesis de la descripción, un reactivo de secuenciación incluye uno o más monómeros de nucleótidos que tienen un resto marcador, un resto terminador o ambos. Dichos restos son grupos químicos que a menudo no son restos naturales de ácidos nucleicos, que se introducen por medios sintéticos para alterar las características naturales de los monómeros de nucleótidos con respecto a la detectabilidad en condiciones particulares o la reactividad enzimática en condiciones particulares. Alternativamente, un reactivo de secuenciación comprende uno o más monómeros de nucleótidos que carecen de un resto marcador no natural y/o un resto terminador no natural. En algunos aspectos de la descripción, el reactivo de secuenciación consiste en o consiste en no más de un tipo de monómero de nucleótido, dos tipos de monómeros de nucleótidos diferentes, tres tipos de monómeros de nucleótidos diferentes o cuatro tipos de monómeros de nucleótidos diferentes. Los tipos de monómeros de nucleótidos "diferentes" son monómeros de nucleótidos que tienen restos de bases que no son iguales entre sí. Dos o más tipos de monómeros de nucleótidos pueden tener otros restos, tales como los expuestos anteriormente, que son iguales entre sí o diferentes entre sí.

Para facilitar la ilustración, en el presente documento se describen varios métodos y composiciones con respecto a múltiples monómeros de nucleótidos. Se entenderá que los múltiples monómeros de nucleótidos de estos métodos o composiciones pueden ser del mismo tipo o de tipos diferentes a menos que se indique explícitamente lo contrario. Debe entenderse que cuando se proporciona un reactivo de secuenciación que comprende múltiples monómeros de nucleótidos a un ácido nucleico diana, los monómeros de nucleótidos no tienen que proporcionarse necesariamente al mismo tiempo. Sin embargo, en aspectos preferidos de los métodos descritos en el presente documento, se proporcionan múltiples monómeros de nucleótidos juntos (al mismo tiempo) al ácido nucleico diana. Independientemente de si los múltiples monómeros de nucleótidos se proporcionan al ácido nucleico diana por separado o juntos, el resultado puede ser que el reactivo de secuenciación, incluidos los monómeros de nucleótidos contenidos en el mismo, estén simultáneamente en presencia del ácido nucleico diana. Por ejemplo, se pueden suministrar dos monómeros de nucleótidos, juntos o por separado, a un ácido nucleico diana. En tales aspectos de la descripción, se habrá proporcionado al ácido nucleico diana un reactivo de secuenciación que comprende dos monómeros de nucleótidos. Alternativamente, se pueden administrar diferentes tipos de nucleótidos secuencialmente en un ciclo que incluye el suministro y la eliminación de cada tipo de nucleótido de manera que los diferentes tipos de nucleótidos no estén presentes simultáneamente con el ácido nucleico diana. En algunos aspectos de la descripción, se incorporarán cero, uno o dos de los monómeros de nucleótidos en un polinucleótido que es complementario del ácido nucleico diana. En algunos aspectos de la descripción, un reactivo de secuenciación puede comprender un oligonucleótido que puede incorporarse en un polímero. El oligonucleótido puede comprender un resto terminador y/o un resto marcador.

Como se usa en el presente documento, "un experimento de secuenciación" y/o sus variantes gramaticales se refiere a un procedimiento repetitivo de etapas físicas o químicas que se lleva a cabo para obtener señales indicativas del orden de los monómeros en un polímero. Las señales pueden ser indicativas de un orden de monómeros con una resolución de un solo monómero o una resolución más baja. En aspectos particulares de la descripción, las etapas se pueden iniciar en un ácido nucleico diana y llevar a cabo para obtener señales indicativas del orden de las bases en el ácido nucleico diana. El procedimiento puede llevarse a cabo hasta su finalización típica, que generalmente se define por el punto en el que las señales del procedimiento ya no pueden distinguir las bases del objetivo con un nivel razonable de certeza. Si se desea, la finalización puede ocurrir antes, por ejemplo, una vez que se haya obtenido la cantidad deseada de información de secuencia. En algunos aspectos de la descripción, un experimento de secuenciación se compone de varios ciclos, donde cada ciclo incluye una serie de dos o más etapas, y la serie de etapas se repite en cada ciclo. Por ejemplo, se pueden llevar a cabo 10 ciclos de un experimento de secuenciación por síntesis para identificar una secuencia de 10 nucleótidos. Cada uno de los 10 ciclos puede incluir etapas de extensión catalizada por polimerasa de un cebador para añadir un análogo de nucleótido que tiene un resto de bloqueo y un resto marcador; detectar el resto marcador en el cebador extendido; y eliminar el resto marcador y el resto de bloqueo del cebador extendido.

Puede llevarse a cabo un experimento de secuenciación en una única molécula de ácido nucleico diana o simultáneamente en una población de moléculas de ácido nucleico diana que tienen la misma secuencia, o simultáneamente en una población de ácidos nucleicos diana que tienen diferentes secuencias. En algunos aspectos de la descripción, una experimento de secuenciación finaliza cuando ya no se obtienen señales de una o más moléculas de ácido nucleico diana a partir de las cuales se inició la adquisición de señales. Por ejemplo, se puede iniciar un experimento de secuenciación para una o más moléculas de ácido nucleico diana que están presentes en un sustrato en fase sólida y terminar tras la eliminación de una o más moléculas de ácido nucleico diana del sustrato. La secuenciación se puede terminar interrumpiendo de otro modo la detección de los ácidos nucleicos diana que estaban presentes en el sustrato cuando se inició el experimento de secuenciación. La secuenciación se puede llevar

a cabo utilizando cualquier técnica de secuenciación adecuada, tal como las descritas en la publicación de solicitud de patente de EE.UU. Nº 2012/0122737 A1.

Como se usa en el presente documento, el término "superficie" y/o sus variantes gramaticales se refiere a una parte de una estructura de soporte (p. ej., sustrato) que es accesible para ponerse en contacto con reactivos, perlas o analitos. La superficie puede ser sustancialmente plana o llana. Alternativamente, la superficie puede ser redondeada o contorneada. Los contornos de ejemplo que se pueden incluir en una superficie son pocillos, depresiones, pilares, crestas, canales o similares. Los materiales de ejemplo que se pueden usar como estructura de soporte incluyen, pero no se limitan a acrílicos, carbono (p. ej., grafito, fibra de carbono), celulosa (p. ej., acetato de celulosa), cerámicos, vidrio de poro controlado, polisacáridos reticulados (p. ej., agarosa o SEPHAROSE®), geles, vidrio (p. ej., vidrio modificado o funcionalizado), oro (p. ej., Au (111) atómicamente liso), grafito, vidrios inorgánicos, polímeros inorgánicos, látex, óxidos metálicos (p. ej., SiO_2, TiO_2, acero inoxidable), metaloides, metales (p. ej., Au (111) atómicamente liso), mica, sulfuros de molibdeno, nanomateriales (p. ej., nanoláminas de grafito pirolítico altamente orientado (HOPG)), nitrocelulosa, NYLON®, haces de fibras ópticas, polímeros orgánicos, papel, plásticos, polacriloilmorfolida, poli(4-metilbuteno), poli(tereftalato de etileno), poli(butirato de vinilo), polibutileno, polidimetilsiloxano (PDMS), polietileno, poliformaldehído, polimetacrilato, polipropileno, polisacáridos, poliestireno poliuretanos, poli(difluoruro de vinilideno) (PVDF), cuarzo, rayón, resinas, cauchos, material semiconductor, sílice, silicio (p. ej., silicio de superficie oxidada), sulfuro y TEFLON®. Un solo material o una mezcla de varios materiales diferentes puede formar una superficie útil en los aspectos de la descripción del presente documento. Los términos "superficie" y "sustrato" se usan indistintamente en el presente documento.

2. Técnica(s) general(es) de la presente descripción

La presente descripción proporciona técnicas mejoradas para la síntesis de ácidos nucleicos. Las técnicas pueden incorporar secuenciación de ácidos nucleicos en una o más etapas de control de calidad de la síntesis. La secuenciación puede facilitar la recogida de ácidos nucleicos sintéticos sin errores seleccionados o recogidos de un campo de moléculas sintéticas. Además, las técnicas pueden permitir que se produzca la síntesis y la secuenciación (escritura y lectura) en un solo aparato de modo que la tasa de error de la síntesis pueda controlarse y cuantificarse para moléculas sintéticas individuales. En algunos casos, la escritura y la lectura pueden ocurrir simultáneamente o en rápida sucesión. Como tal, la presente descripción describe sistemas y métodos para desarrollar un dispositivo de síntesis y secuenciación (dispositivo de escritura y lectura) para llevar la biología sintética al trabajo biológico rutinario.

Además, las técnicas expuestas en el presente documento se pueden llevar a cabo de una manera independiente de molde. Más específicamente, un ácido nucleico puede sintetizarse de novo, por ejemplo, ensamblando fragmentos de ácidos nucleicos sin interacciones de los fragmentos con un ácido nucleico molde. Los propios fragmentos se pueden sintetizar sin el uso de un molde, por ejemplo, usando una técnica de síntesis química o usando una enzima independiente de molde. No obstante, se entenderá que se pueden usar métodos de síntesis dependientes de molde para generar fragmentos o ensamblar fragmentos según se desee para adaptarse a algunos aspectos de la descripción.

Volviendo ahora a los dibujos, y refiriéndonos primero a la Fig. 1, se ilustra una visión general esquemática de una técnica para generar una molécula de polinucleótido sintética, que, en ciertos aspectos de la descripción, se puede utilizar a su vez para crear otros ácidos nucleicos (p. ej., mediante transcripción) y/o proteínas (p. ej., mediante traducción) de acuerdo con los aspectos de la presente descripción. La síntesis de ácido nucleico puede comenzar con una fase de diseño de fragmentos 10. La fase de diseño de fragmentos 10 puede incluir el uso de información sobre una secuencia de ácido nucleico diana 12 para diseñar secuencias para uno o más fragmentos de ácido nucleico (p. ej., fragmentos 14a, 14b, 14c, 14d). Aunque se ilustran cuatro fragmentos de ácido nucleico (p. ej., fragmentos 14a, 14b, 14c, 14d) en la Fig. 1, se entiende que las secuencias para cualquier número (p. ej., al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más) de fragmentos de ácido nucleico pueden diseñarse basándose en la secuencia de ácido nucleico diana 12.

Como se ilustra, cada uno de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d incluye una secuencia complementaria de al menos otro fragmento de ácido nucleico 14a, 14b, 14c, 14d. Por ejemplo, el fragmento de ácido nucleico 14a tiene una secuencia complementaria (p. ej., una secuencia 16a) de una secuencia (p. ej., una secuencia 16b) del fragmento de ácido nucleico 14b, ilustrada por una región que se solapa. El fragmento de ácido nucleico 14b tiene una secuencia complementaria (p. ej., una secuencia 18a) de una secuencia (p. ej., una secuencia 18b) del fragmento de ácido nucleico 14c. El fragmento de ácido nucleico 14c tiene una secuencia complementaria (p. ej., una secuencia 20a) de una secuencia (p. ej., una secuencia 20b) del fragmento de ácido nucleico 14d. En ciertos aspectos de la descripción, la secuencia complementaria es complementaria de solo una parte de otro fragmento. Es decir, la secuencia complementaria 16a representa solo una parte del fragmento 14a. Además, las secuencias complementarias 16b y 18b representan solo una parte del fragmento 14b. En aspectos particulares de la descripción, las partes de secuencia complementaria se encuentran en los extremos de los fragmentos.

En algunos aspectos de la descripción, las secuencias de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d pueden diseñarse basándose en la secuencia de ácido nucleico diana 12 de manera que las secuencias de fragmentos de ácido nucleico 14a, 14b, 14c, 14d tengan una secuencia combinada (con las secuencias que se solapan contadas una vez) que es igual o complementaria de la secuencia de ácido nucleico diana 12. Por ejemplo, la secuencia general, que incluye la secuencia del fragmento de ácido nucleico 14a, la secuencia del fragmento de ácido nucleico 14b sin las secuencias 16b, 18a, la secuencia del fragmento de ácido nucleico 14c, y la secuencia del fragmento de ácido

nucleico 14d sin la secuencia 20b, es la misma o complementaria de la secuencia de ácido nucleico diana 12. En este aspecto de la descripción, no hay huecos en la secuencia combinada de los fragmentos en comparación con la secuencia de ácido nucleico diana. Como se expone en detalle a continuación, en aspectos alternativos de la descripción, pueden estar presentes huecos en la secuencia combinada de los fragmentos de ácido nucleico (p. ej., entre las secuencias de los fragmentos) en comparación con la secuencia de ácido nucleico diana.

La secuencia de ácido nucleico diana 12 generalmente tiene al menos dos nucleótidos. Por ejemplo, la secuencia de ácido nucleico diana 12 puede ser de al menos aproximadamente 10, 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 bases, así como de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 kb, así como aproximadamente 11, 12, 13, 14, 15, 16, 17 , 18, 19 o 20 kb, así como aproximadamente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 kb, así como aproximadamente 200, 300, 400, 500, 600, 700, 800, 900 o 1.000 kb y todos los incrementos en los mismos. Alternativa o adicionalmente, la secuencia de ácido nucleico diana 12 puede ser de no más de aproximadamente 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 bases, así como de aproximadamente 1, 2, 3, 4, 5, 6, 7, 8, 9 o 10 kb, así como aproximadamente 11, 12, 13, 14, 15, 16, 17, 18, 19 , o 20 kb, así como aproximadamente 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95 o 100 kb, así como aproximadamente 200, 300, 400, 500, 600, 700, 800, 900 o 1000 kb.

Las secuencias de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d tienen generalmente al menos dos nucleótidos. Por ejemplo, los fragmentos de ácido nucleico 14a, 14b, 14c, 14d pueden ser de al menos aproximadamente 10, 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 bases y todos sus incrementos. Alternativa o adicionalmente, los fragmentos de ácido nucleico 14a, 14b, 14c, 14d pueden ser de no más de aproximadamente 10, 25, 50, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900 o 950 bases.

Las secuencias complementarias que se solapan de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d (p. ej., secuencias 16a, 16b, 18a, 18b, 20a, 20b) generalmente tienen al menos un nucleótido. Por ejemplo, las secuencias que se solapan de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d pueden ser de más o menos de 2, 3, 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20, 21, 22, 23, 24, 25, 26, 27, 28, 29, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 150, 200, 250, 300, 350, 400, 450, 500, 550, 600, 650, 700, 750, 800, 850, 900, 950 bases y todos los incrementos de los mismos. En aspectos particulares de la descripción, las secuencias que se solapan de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d pueden tener una longitud de entre 1 y 50 bases, tal como entre 5 y 45 bases, entre 10 y 40 bases, entre 15 y 35 bases, entre 20 y 30 bases. En aspectos particulares de la descripción, las secuencias que se solapan o complementarias pueden ser más cortas que el tamaño de los fragmentos 14 de los que forman parte. Para los fragmentos con dos solapamientos complementarios, las secuencias complementarias son cada una menos de la mitad del tamaño total del fragmento.

La fase de diseño de fragmentos 10 es seguida después por una fase de síntesis de fragmentos 22. En un aspecto de la descripción, uno o más fragmentos de ácido nucleico 14a, 14b, 14c, 14d se sintetizan sobre un sustrato 24. La síntesis de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d pueden llevarse a cabo por extensión de cebadores inmovilizados sobre el sustrato 24 añadiendo sucesivamente nucleótidos a un grupo hidroxilo 3' libre, lo que da como resultado la síntesis en la dirección 5' a 3' de polinucleótidos que se dirigen a los fragmentos de ácido nucleico 14a, 14b, 14c, 14d. En un aspecto específico de la descripción, la síntesis del fragmento se puede llevar a cabo en presencia de una polimerasa, tal como la desoxinucleotidil transferasa terminal (TdT).

La síntesis ilustrada en la Fig. 1, se puede llevar a cabo en ausencia de un molde, puesto que la TdT puede añadir nucleótidos sin un molde. En aspectos particulares de la descripción, un cebador que está inmovilizado sobre el sustrato 24 se trata secuencialmente con diferentes tipos de análogos de nucleótidos en presencia de TdT. Los tipos de nucleótidos diferirán con respecto a qué base está presente en el nucleótido. El orden de los tipos de análogos de nucleótidos suministrados al cebador determinará la secuencia del fragmento producido. Además, los análogos de nucleótidos pueden incluir un resto de bloqueo reversible para evitar la adición de más de un nucleótido al cebador por ciclo. En este caso, un ciclo puede incluir (1) el suministro del tipo apropiado de análogo de nucleótido bloqueado reversiblemente al cebador inmovilizado en la superficie, en presencia de TdT, mediante el cual se añade un análogo de nucleótido único al extremo 3' del cebador para crear un cebador extendido; (2) eliminar por lavado el análogo de nucleótido que no ha reaccionado (y opcionalmente la TdT) del sustrato 24; y (3) eliminar el resto de bloqueo del análogo de nucleótido en el extremo 3' del cebador extendido. Los aspectos particulares de la descripción pueden usar una especie de TdT, o una variante modificada de una especie de TdT que tiene la actividad deseada para los nucleótidos que tienen un grupo de bloqueo particular. Dichas especies de TdT se pueden crear usando técnicas de mutagénesis conocidas y cribado de la actividad de polimerasa usando la especie de nucleótido particular en ensayos de polimerasa conocidos.

Alternativamente, la síntesis de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d puede llevarse a cabo por extensión de cebadores inmovilizados sobre el sustrato 24 añadiendo sucesivamente nucleótidos en la dirección 3' a 5'. Por ejemplo, la síntesis química se puede llevar a cabo en un procedimiento cíclico bien conocido que ensambla una cadena de nucleótidos. Los nucleótidos se añaden uno por uno a través de un ciclo de reacciones químicas, en el que se añade una molécula particular (p. ej., un nucleótido) a una molécula de ADN en crecimiento (p. ej., una cadena de ADN en crecimiento), a veces mediante catálisis, hasta que se completa la cadena deseada. Generalmente, cada ciclo de reacciones químicas incluye las etapas de destritilación, acoplamiento, rematado y oxidación. Durante

la etapa de destritilación o "desprotección", se elimina un grupo dimetoxitritilo (DMT) del último nucleótido de la cadena de ADN en crecimiento para permitir la adición del siguiente nucleótido. La cantidad de DMT liberado de cada ciclo se puede vigilar para determinar la eficiencia del acoplamiento. La liberación de DMT es evidente porque se emite un color naranja brillante cuando se libera DMT. Los métodos de síntesis química de ejemplo que se pueden usar se exponen en la patente de EE.UU. Nº 7.914.739 o publicación de solicitud de patente de EE.UU. Nº 2004/0219063 A1.

Otro método sintético útil es crear un fragmento de ácido nucleico añadiendo secuencialmente trinucleótidos (es decir, oligonucleótidos trímeros). En un aspecto particular de la descripción, se pueden sintetizar 64 trinucleótidos, por ejemplo, usando un método de síntesis química como los expuestos anteriormente. En este ejemplo, los 64 trinucleótidos corresponden a los codones de los 20 aminoácidos naturales. Los trinucleótidos se pueden ligar entre sí usando Circligase (Epicenter, WI). La ligadura se puede realizar en una dirección de 5' a 3' o alternativamente en una dirección de 3' a 5'. De cualquier manera, el fragmento resultante codificará una serie de aminoácidos que corresponden al número de trinucleótidos que se ligaron entre sí. Y la secuencia de los aminoácidos corresponderá a la secuencia de adición de trinucleótidos utilizada para sintetizar el fragmento de ácido nucleico.

En aspectos particulares de la descripción de la síntesis basada en ligación, es útil usar trinucleótidos que tienen un resto de bloqueo reversible en el extremo 5' o 3'. Por ejemplo, al sintetizar un fragmento de ácido nucleico en la dirección 5' a 3', es útil utilizar trinucleótidos que están bloqueados en la posición 3'. En este caso, cada ciclo de adición se puede llevar a cabo suministrando una solución de trinucleótidos para un primer codón a un fragmento naciente en presencia de Circligase. El fragmento se puede unir a través del extremo 5' a un soporte sólido, o el extremo 5' se puede bloquear de otro modo para evitar el ligado. Como resultado, el suceso de ligado deseado dará como resultado el ligado entre el extremo 3' del fragmento naciente y el extremo 5' del trinucleótido que entra. Se evita la adición de más de un trinucleótido por ciclo porque, una vez extendido por adición de un solo trunucleótido, el fragmento naciente ahora tendrá un grupo de bloqueo 3' que evita la extensión adicional. El exceso de trinucleótidos puede eliminarse, por ejemplo, mediante lavado, y luego el grupo de bloqueo 3' puede eliminarse o modificarse para generar un extremo 3' no bloqueado en el fragmento naciente extendido. A continuación, se puede llevar a cabo un segundo ciclo suministrando una solución de trinucleótidos para un segundo codón al fragmento naciente que se extendió y desbloqueó previamente, nuevamente, en presencia de Circligase.

El ligado de trinucleótidos también se puede llevar a cabo en la dirección 3' a 5'. Los trinucleótidos pueden opcionalmente bloquearse en la posición 5' y el fragmento naciente puede opcionalmente bloquearse o unirse a una superficie a través del extremo 3'. Cada ciclo de adición se puede llevar a cabo suministrando una solución de trinucleótidos para un primer codón a un fragmento naciente en presencia de Circligase, extendiendo así el fragmento naciente debido al ligado entre el extremo 5' del fragmento naciente y el extremo 3' del trinucleótido que entra. El exceso de trinucleótidos puede eliminarse, por ejemplo, mediante lavado, y luego el grupo de bloqueo 5' puede eliminarse o modificarse para generar un extremo 5' no bloqueado en el fragmento naciente extendido. A continuación, se puede llevar a cabo un segundo ciclo suministrando una solución de trinucleótidos para un segundo codón al fragmento naciente que se extendió y desbloqueó previamente, en presencia de Circligase. Se entenderá que, aunque se ilustra para trinucleótidos, se puede llevar a cabo una síntesis de ligado similar usando oligonucleótidos tetrámeros, oligonucleótidos pentámeros, oligonucleótidos hexámeros, u oligonucleótidos de mayor tamaño para el ligado a un fragmento naciente.

Como se ilustra, los polinucleótidos 26a, 26b, 26c, 26d, que se dirigen a las secuencias de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d, respectivamente, pueden sintetizarse sobre el sustrato 24. Además, se puede llevar a cabo una amplificación de los polinucleótidos sintetizados 26a, 26b, 26c, 26d sobre el sustrato 24 para generar grupos de polinucleótidos 26a, 26b, 26c, 26d. La amplificación sobre el sustrato 24 puede llevarse a cabo usando amplificación de puente, PCR en fase sólida, amplificación por círculo rodante (RCA) o cualquier otro método adecuado como se ha indicado anteriormente. En algunos aspectos de la descripción, la amplificación de los polinucleótidos 26a, 26b, 26c, 26d sintetizados es opcional.

La fase de síntesis de fragmentos 22 puede ir seguida de una fase de control de calidad de la síntesis 28 en la que se secuencian los polinucleótidos sintetizados 26a, 26b, 26c, 26d y se comparan las secuencias de los polinucleótidos sintetizados 26a, 26b, 26c, 26d con las secuencias de sus respectivos fragmentos de ácido nucleico 14a, 14b, 14c, 14d. Puede llevarse a cabo un experimento de secuenciación sucesivamente o en paralelo para los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación. El experimento de secuenciación se puede llevar a cabo como se describe en la publicación de solicitud de patente de EE.UU. Nº 2010/0279882 A1, o patente de EE.UU. Nº 8.637.242. Como se ilustra, el experimento de secuenciación puede incluir la utilización de un reactivo de secuenciación 30 que incluye uno o más nucleótidos que tienen un resto marcador, un resto terminador o ambos.

Una vez realizado el experimento de secuenciación, se puede obtener una secuencia de cada uno de los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación. Comparando luego la secuencia de cada uno de los polinucleótidos sintetizados 26a, 26b, 26c, 26d, y/o su secuencia complementaria en el caso de grupos de amplificación, con la secuencia de sus respectivos fragmentos de ácido nucleico 14a, 14b, 14c, 14d, se puede determinar una tasa de error de síntesis. En algunos aspectos de la descripción, también se puede determinar una puntuación de precisión de secuencia correspondiente (p. ej., restando la tasa de error de uno) para cada uno de los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación. Generalmente, el sustrato 24 incluirá una redundancia de sitios para cada fragmento deseado. Por ejemplo, cada fragmento puede sintetizarse de forma

redundante en al menos 2, 4, 5, 10, 25, 50, 100 o más sitios en un sustrato. Se puede determinar el número de sitios redundantes para producir un número deseado de sitios precisos basado en la tasa de error esperada para el procedimiento de síntesis que se está utilizando. Por tanto, aunque algunos sitios pueden contener fragmentos con errores en sus secuencias, pueden identificarse otros sitios con secuencias precisas y recoger los fragmentos de esos sitios como se expone con más detalle a continuación.

Al incorporar la secuenciación en la fase de control de calidad de la síntesis 28, los datos de secuenciación pueden usarse para el control de calidad de la síntesis de ácidos nucleicos. Por ejemplo, se puede usar una tasa de error de síntesis determinada mediante la secuenciación para determinar como de bueno era el diseño del fragmento. A modo de ejemplo adicional, una tasa de error de síntesis más baja de cada uno de los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación puede representar un mejor diseño de fragmentos de sus respectivos fragmentos de ácido nucleico 14a, 14b, 14c, 14d. A la inversa, una mayor tasa de error de síntesis de cada uno de los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación puede representar un diseño de fragmento peor de sus respectivos fragmentos de ácido nucleico 14a, 14b, 14c, 14d. Así pues, se puede establecer un umbral predeterminado de tasa de error de síntesis y usarlo para comparar con la tasa de error de síntesis de cada uno de los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación determinada por la secuenciación para determinar si sus respectivos fragmentos diseñados son deseables. Por ejemplo, la tasa de error umbral predeterminada puede ser como máximo aproximadamente 0,00%, 0,01%, 0,02%, 0,03%, 0,04%, 0,05%, 0,06%, 0,07%, 0,08%, 0,09%, 0,10%, 0,15%, 0,20 %, 0,25%, 0,30%, 0,35%, 0,40%, 0,45%, 0,50%, 0,55%, 0,60%, 0,65%, 0,70%, 0,75%, 0,80%, 0,85%, 0,90%, 0,95%, 1,00%, 1,10%, 1,20%, 1,30%, 1,40%, 1,50%, 1,60%, 1,70%, 1,80%, 1,90%, 2,00%, 2,10%, 2,20%, 2,30%, 2,40%, 2,50%, 2,60%, 2,70%, 2,80%, 2,90%, 3,00%, 4,00%, 5,00%, 6,00%, 7,00%, 8,00%, 9,00%, 10,00%, 15,00%, 20,00%, 25,00%, 30,00%, 35,00%, 40,00%, 45,00%, 50,00%, 55,00%, 60,00%, 65,00%, 70,00%, 75,00%, 80,00%, 85,00%, 90,00%, 95,00%, 100%. Esta información se puede utilizar para diseñar un conjunto diferente de secuencias de fragmentos. El procedimiento de síntesis y de QC se puede repetir varias veces para permitir una mejora iterativa en la calidad de la secuencia.

Además, la tasa de error de síntesis determinada mediante la secuenciación en la fase de control de calidad de la síntesis 28 puede usarse para determinar cómo recoger los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación. Por ejemplo, los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación se pueden recoger selectivamente cuando su tasa de error respectiva es igual, menor o mayor que una tasa de error umbral predeterminada. La tasa de error de umbral predeterminada puede ser como máximo de aproximadamente 0,00%, 0,01%, 0,02%, 0,03%, 0,04%, 0,05%, 0,06%, 0,07%, 0,08%, 0,09%, 0,10%, 0,15%, 0,20%, 0,25 %, 0,30%, 0,35%, 0,40%, 0,45%, 0,50%, 0,55%, 0,60%, 0,65%, 0,70%, 0,75%, 0,80%, 0,85%, 0,90%, 0,95%, 1,00%, 1,10%, 1,20%, 1,30%, 1,40%, 1,50%, 1,60%, 1,70%, 1,80%, 1,90%, 2,00%, 2,10%, 2,20%, 2,30%, 2,40%, 2,50%, 2,60%, 2,70%, 2,80% , 2,90%, 3,00%, 4,00%, 5,00%, 6,00%, 7,00%, 8,00%, 9,00%, 10,00%, 15,00%, 20,00%, 25,00%, 30,00%, 35,00%, 40,00%, 45,00%, 50,00%, 55,00%, 60,00%, 65,00%, 70,00%, 75,00%, 80,00%, 85,00%, 90,00%, 95,00%, 100%. En algunos aspectos de la descripción, se pueden comparar las tasas de error de los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación, y se pueden recoger selectivamente los polinucleótidos sintetizados que tienen la más baja, la más alta o igual a, menor que o mayor que una tasa de error de umbral predeterminada.

Como la secuencia de cada uno de los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación se puede obtener por secuenciación, también se puede obtener un rendimiento de una secuencia sintetizada predeterminada de uno de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d. Por ejemplo, una secuencia sintetizada predeterminada del fragmento de ácido nucleico 14a puede ser una secuencia exenta de errores (p. ej., con una tasa de error de síntesis de aproximadamente 0,0%), y si la secuencia de todos los polinucleótidos sintetizados 26a (p. ej., una sola copia o múltiples copias en el grupo de amplificación) se determina que es la secuencia exenta de errores, entonces el rendimiento de una secuencia exenta de errores del fragmento de ácido nucleico 14a a partir de la síntesis es de 100%. En algunos aspectos de la descripción, a partir de la secuenciación se puede determinar un rendimiento de secuencias sintetizadas con una tasa de error predeterminada, independientemente de las secuencias reales, de uno de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d. Por ejemplo, en los 100 polinucleótidos sintetizados 26a en el grupo de amplificación que se secuencian, hay 90 polinucleótidos 26a (p. ej., con o sin la misma secuencia) que tienen una tasa de error de síntesis de aproximadamente 0,05%. Por tanto, el rendimiento de los polinucleótidos sintetizados 26a que tienen una tasa de error de síntesis de aproximadamente 0,05% es de 90%.

Así pues, además de la tasa de error de síntesis, o alternativamente, el rendimiento de una secuencia sintetizada y/o secuencias sintetizadas predeterminadas con una tasa de error predeterminada puede usarse para el control de calidad de la síntesis de ácidos nucleicos en la fase de control de calidad de la síntesis 28. De manera similar a la tasa de error de síntesis, dicho rendimiento puede usarse para determinar cómo de bueno era el diseño del fragmento y/o para determinar cómo recoger los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación.

Como se ha expuesto antes, los fragmentos se pueden amplificar sobre una superficie antes de la secuenciación. La amplificación se puede llevar a cabo para crear una colonia (también denominada un "grupo") de copias para formar una característica en la superficie. Se pueden sintetizar y amplificar múltiples fragmentos en una superficie para crear una matriz de características. Es posible que el procedimiento de amplificación pueda introducir errores para producir un grupo mixto que contiene algunas copias del fragmento sintetizado originalmente y otras copias que contienen el error. La presencia de un error de amplificación en un grupo o característica puede reconocerse por la presencia de una señal mixta en una o más posiciones de nucleótidos para los datos de secuencia obtenidos en el grupo o

característica. Por lo tanto, el control de calidad puede incluir la determinación de errores de amplificación, por ejemplo, en función de la presencia o ausencia de señales mixtas en un grupo, un valor umbral para la variación estadística observada en una o más posiciones en los datos de secuencia obtenidos para un grupo, un valor umbral de la relación señal a ruido observada en una o más posiciones en los datos de secuencia obtenidos para un grupo, una tasa de disminución en la señal a ruido en una característica, o similares. Nuevamente, estos datos de control de calidad pueden usarse para determinar cómo de bueno era el diseño del fragmento y/o para determinar cómo recoger los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación.

Continuando con los aspectos ilustrativos de la descripción de la Fig. 1, después de que algunos o todos los polinucleótidos sintetizados 26a, 26b, 26c, 26d o sus grupos de amplificación son recogidos del sustrato 24, sigue una fase de ensamblaje de fragmentos 32. En la fase de ensamblaje de fragmentos 32, los polinucleótidos sintetizados se ensamblan para generar uno o más polinucleótidos ensamblados que se dirigen a la secuencia de ácido nucleico diana 12. En los aspectos ilustrados de la descripción, los polinucleótidos sintetizados 26b, 26c, 26d se recogen del sustrato 24 mientras que los polinucleótidos sintetizados 26a permanecen en el sustrato 24. Durante la fase de ensamblaje de fragmentos 32, los polinucleótidos sintetizados 26b, 26c, 26d se ensamblan con los polinucleótidos sintetizados 26a para generar los polinucleótidos ensamblados 34. Como se analiza en detalle a continuación, el ensamblaje se puede llevar a cabo en la presencia de una o más enzimas, que incluyen, pero no se limitan a, ADN recombinasa, proteína beta, ADN polimerasa, ADN ligasa, Circligase o cualquier combinación de las mismas, para el ensamblaje de fragmentos y la reparación de huecos.

A la fase de ensamblaje de fragmentos 32 le puede seguir una fase de control de calidad del ensamblaje 36 en la que los polinucleótidos 34 ensamblados se secuencian y las secuencias de los polinucleótidos ensamblados 34 se comparan con la secuencia del ácido nucleico diana 12. En algunos aspectos de la descripción, se puede llevar a cabo una amplificación de los polinucleótidos ensamblados 34 antes de la secuenciación, y se pueden llevar a cabo los experimentos de secuenciación en los grupos de amplificación. En otros aspectos de la descripción, la amplificación de los polinucleótidos ensamblados 34 es opcional. Los experimentos de secuenciación en las fases de control de calidad del ensamblaje 36 se pueden llevar a cabo de manera similar a los experimentos de secuenciación en la fase de control de calidad de la síntesis 28. Como se ilustra, los experimentos de secuenciación pueden incluir la utilización de un reactivo de secuenciación 38 similar al reactivo de secuenciación 30. Se pueden llevar a cabo etapas de control de calidad para el polinucleótido ensamblado usando técnicas expuestas anteriormente para fragmentos de ácido nucleico.

Después de que se lleven a cabo los experimentos de secuenciación, se puede obtener una secuencia de los polinucleótidos ensamblados 34 o sus grupos de amplificación. Luego, comparando la secuencia de los polinucleótidos ensamblados 34, y/o su secuencia complementaria en el caso de grupos de amplificación, con la secuencia de la secuencia del ácido nucleico diana 12, se puede determinar una tasa de error de síntesis. En algunos aspectos de la descripción, también se puede determinar una puntuación de precisión de secuencia correspondiente (p. ej., restando la tasa de error de uno).

De manera similar a la fase de control de calidad de la síntesis 28, la fase de control de calidad del ensamblaje 36 puede usar la tasa de error de síntesis para el control de calidad de la síntesis del ácido nucleico. Por ejemplo, la tasa de error de síntesis determinada por la secuenciación en la fase de control de calidad del ensamblaje 36 puede usarse para determinar cómo de bueno era el diseño del fragmento. A modo de ejemplo adicional, una tasa de error de síntesis más baja de los polinucleótidos ensamblados 34 o sus grupos de amplificación puede representar un mejor diseño de fragmentos de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d. A la inversa, una tasa de error de síntesis más alta de los polinucleótidos ensamblados 34 o sus grupos de amplificación puede representar un diseño de fragmento peor de los fragmentos de ácido nucleico 14a, 14b, 14c, 14d. Así pues, se puede establecer un umbral predeterminado de tasa de error de síntesis y usar para comparar con la tasa de error de síntesis de los polinucleótidos ensamblados 34 o sus grupos de amplificación determinada por la secuenciación para determinar si los fragmentos diseñados son deseables. Por ejemplo, la tasa de error de umbral predeterminada puede ser como máximo aproximadamente 0,00%, 0,01%, 0,02%, 0,03%, 0,04%, 0,05%, 0,06%, 0,07%, 0,08%, 0,09%, 0,10%, 0,15%, 0,20 %, 0,25%, 0,30%, 0,35%, 0,40%, 0,45%, 0,50%, 0,55%, 0,60%, 0,65%, 0,70%, 0,75%, 0,80%, 0,85%, 0,90%, 0,95%, 1,00%, 1,10%, 1,20%, 1,30%, 1,40%, 1,50%, 1,60%, 1,70%, 1,80%, 1,90%, 2,00%, 2,10%, 2,20%, 2,30%, 2,40%, 2,50%, 2,60%, 2,70%, 2,80%, 2,90%, 3,00%, 4,00%, 5,00%, 6,00%, 7,00%, 8,00%, 9,00%, 10,00%, 15,00%, 20,00%, 25,00%, 30,00%, 35,00%, 40,00%, 45,00 %, 50,00%, 55,00%, 60,00%, 65,00%, 70,00%, 75,00%, 80,00%, 85,00%, 90,00%, 95,00%, 100%. Esta información se puede usar para diseñar un conjunto diferente de secuencias de fragmentos o polinucleótidos ensamblados diferentes. El procedimiento de síntesis y de QC se puede repetir varias veces para permitir una mejora iterativa en la calidad de la secuencia.

Además, la tasa de error de síntesis determinada por la secuenciación en la fase de control de calidad del ensamblaje 36 puede usarse para determinar la recolección selectiva de los polinucleótidos ensamblados 34 o sus grupos de amplificación. Por ejemplo, los polinucleótidos ensamblados 34 o sus grupos de amplificación se pueden recoger selectivamente cuando la tasa de error es igual, menor o mayor que una tasa de error umbral predeterminada. La tasa de error de umbral predeterminado puede ser como máximo de aproximadamente 0,00%, 0,01%, 0,02%, 0,03%, 0,04%, 0,05%, 0,06%, 0,07%, 0,08%, 0,09%, 0,10%, 0,15%, 0,20%, 0,25 %, 0,30%, 0,35%, 0,40%, 0,45%, 0,50%, 0,55%, 0,60%, 0,65%, 0,70%, 0,75%, 0,80%, 0,85%, 0,90%, 0,95%, 1,00%, 1,10%, 1,20%, 1,30%, 1,40%, 1,50%, 1,60%, 1,70%, 1,80%, 1,90%, 2,00%, 2,10%, 2,20%, 2,30%, 2,40%, 2,50%, 2,60%, 2,70%, 2,80%, 2,90%, 3,00%, 4,00%, 5,00%, 6,00%, 7,00%, 8,00%, 9,00%, 10,00%, 15,00%, 20,00%, 25,00%, 30,00%, 35,00%, 40,00%, 45,00%,

50,00 %, 55,00%, 60,00%, 65,00%, 70,00%, 75,00%, 80,00%, 85,00%, 90,00%, 95,00%, 100%. En algunos aspectos de la descripción, se pueden comparar las tasas de error de los polinucleótidos ensamblados 34 o sus grupos de amplificación, y los polinucleótidos ensamblados que tienen la más baja, la más alta o igual a, menor que o mayor que una tasa de error de umbral predeterminado se pueden recoger selectivamente. Los polinucleótidos ensamblados 34 se pueden recoger del sustrato 24 de forma óptica, química, magnética, eléctrica, electromagnética o cualquier combinación de los mismos.

En algunos aspectos de la descripción, los polinucleótidos ensamblados 34 pueden usarse a su vez para crear otros ácidos nucleicos, por ejemplo, mediante una fase de transcripción 40. Por ejemplo, los polinucleótidos ensamblados 34 pueden ser ADN 42, que pueden usarse como molde para sintetizar ARNm 44. El ADN puede ser monocatenario o bicatenario para adaptarse a usos particulares. En ciertos aspectos de la descripción, la fase de transcripción 40 puede ir seguida de una fase de traducción 46. Por ejemplo, los ARNm 44 sintetizados pueden usarse como moldes para sintetizar proteínas 48. La transcripción y traducción pueden llevarse a cabo usando métodos conocidos en la técnica tales como los descritos en Sambrook, et al., Molecular Cloning: A Laboratory Manual (3ª edición, 2001); Sambrook, et al., Molecular Cloning: A Laboratory Manual (2ª edición, 1989); Maniatis et al., Molecular Cloning: A Laboratory Manual (1982); y Ausubel et al., Current Protocols in Molecular Biology (John Wiley and Sons, actualizado en julio de 2008).

La Fig. 2 es un diagrama de bloques de un sistema de síntesis 50 que puede usarse junto con la técnica de la Fig. 1 u otros métodos expuestos en el presente documento. El sistema de síntesis 50 incluye un sintetizador 52 diseñado para sintetizar ácidos nucleicos. El sintetizador 52 se comunica con una interfaz de control/operador del sistema 54. La interfaz de control/operador del sistema 54 puede usarse para el diseño de fragmentos basándose en la secuencia de ácido nucleico diana 12. Por ejemplo, la interfaz de control/operador del sistema 54 puede ejecutar un programa almacenado local o remotamente para facilitar los diseños de uno o más fragmentos de ácido nucleico basados en la secuencia de ácido nucleico diana 12, como se indicó anteriormente. Alternativa o adicionalmente, un usuario puede introducir una o más secuencias de fragmentos de ácido nucleico diseñadas en la interfaz de control/operador del sistema 54.

La interfaz de control/operador del sistema 54 puede incluir un ordenador de uso general o específico de la aplicación diseñado para controlar los parámetros del procedimiento, los datos adquiridos, los ajustes del sistema, etc. La interfaz de control/operador del sistema 54 puede incluir uno o más procesadores y uno o más dispositivos de memoria que almacenan instrucciones ejecutables por uno o más procesadores. La interfaz del operador puede ser generada por un programa ejecutado localmente o por programas almacenados y/o ejecutados dentro del sintetizador 52 para proporcionar indicaciones visuales del estado de los sistemas o subsistemas del sintetizador 52, la calidad de los datos adquiridos, etc.

La interfaz de control/operador del sistema 54 también puede permitir que un usuario interactúe con el sistema para regular el funcionamiento, iniciar e interrumpir la síntesis, evaluar la información de control de calidad y cualquier otra interacción que pueda desearse con el hardware o software del sistema. Por ejemplo, la interfaz de control/operador del sistema 54 puede realizar y/o modificar automáticamente las etapas que hay que llevar a cabo en un procedimiento de síntesis, sin la participación de un usuario. Alternativa o adicionalmente, la interfaz de operador/control del sistema 54 puede generar recomendaciones con respecto a las etapas que hay que llevar a cabo en un procedimiento de síntesis y mostrar estas recomendaciones al usuario. Este modo permitiría la participación del operador humano antes de realizar y/o modificar etapas en el procedimiento de secuenciación. Además, la interfaz de control/operador del sistema 54 puede proporcionar una opción para el usuario que permite al usuario seleccionar ciertas etapas en un procedimiento de síntesis para que el sintetizador 52 las realice automáticamente a la vez que se requiere la participación del usuario antes de llevar a cabo y/o modificar otras etapas. En cualquier caso, permitir modos tanto automatizados como interactivos con el usuario puede proporcionar una mayor flexibilidad en la realización del procedimiento de síntesis. Además, la combinación de automatización e interacción controlada por ser humano puede permitir además que el sistema de síntesis 50 sea capaz de crear y modificar nuevos procedimientos y algoritmos de secuenciación a través del aprendizaje automático adaptativo basado en las entradas recopiladas del usuario y/o subsistemas del sintetizador 52.

La interfaz de control/operador del sistema 54 puede incluir un sistema de posprocesamiento que puede incluir uno o más ordenadores programados que reciben información detectada, que puede estar en forma de datos de imagen pixelados y derivar datos de secuencia a partir de los datos de imagen. El sistema de posprocesamiento puede incluir algoritmos de reconocimiento de imágenes que distinguen entre los colores de los colorantes unidos a los nucleótidos que se unen en sitios individuales durante los procedimientos de secuenciación (p. ej., por análisis de los datos de imagen que codifican colores o intensidades específicas) y registran la secuencia de los nucleótidos en las ubicaciones individuales del sitio. Entonces, progresivamente, el sistema de posprocesamiento puede construir listas de secuencias para los sitios individuales de la matriz de muestras que pueden procesarse adicionalmente para establecer información genética para longitudes extendidas de material mediante varios algoritmos bioinformáticos.

El sintetizador 52 incluye un sistema de sintetizador/secuenciador/ensamblador 56 configurado para llevar a cabo varias etapas de la técnica ilustrada en la Fig. 1, incluida la fase de síntesis de fragmentos 22, la secuenciación de ácidos nucleicos durante la fase de control de calidad de la síntesis 28 y la fase de control de calidad del ensamblaje 36, y las fases de ensamblaje de fragmentos 32. En ciertos aspectos de la descripción, el sistema de sintetizador/secuenciador/ensamblador 56 está configurado para llevar a cabo la fase de transcripción 40 y/o la fase

de traducción 46 ilustrada en la Fig. 1. Estas fases son ilustrativas y se entenderá que el sistema puede configurarse para realizar fases adicionales o alternativas. Como se analiza en detalle a continuación, se puede usar uno o más sistemas de soporte (p. ej., el sustrato 24), que incluyen, pero no se limitan a una perla, una perla magnética, un portaobjetos de vidrio, un microchip, una nanogotita, un cartucho de electrohumectación o cualquier combinación de los mismos, para la fase de síntesis de fragmentos 22, las fases de control de calidad de la síntesis 28, la fase de ensamblaje de fragmentos 32 y/o la fase de control de calidad del ensamblaje 36 ilustrada en la Fig. 1. El sistema de soporte puede usarse como contenedor de reacción o portador junto con el sistema de sintetizador/secuenciador/ensamblador 56. El sistema de sintetizador/secuenciador/ensamblador 56 puede ser una sola plataforma o sustrato o puede incluir múltiples plataformas o sustratos, según sea adecuado. Por ejemplo, la síntesis de fragmentos puede ocurrir en un primer sustrato mientras que el ensamblaje puede ocurrir en un segundo sustrato. Además, la síntesis de fragmentos y el ensamblaje de fragmentos pueden implicar sustratos con diferente resolución (p. ej., diferente resolución de densidad de molécula). En aspectos particulares de la descripción, para el sintetizador/secuenciador/ensamblador multiplataforma, el movimiento entre las plataformas puede ser automático o puede incluir ciertas etapas mediadas por el operador.

El sintetizador 52 también incluye un sistema de control/suministro de fluidos 58 y un sistema de detección 60. El sistema de control/suministro de fluidos 58 puede recibir una pluralidad de fluidos del procedimiento para suministrar al sistema de sintetizador/secuenciador/ensamblador 56. Como apreciarán los expertos en la técnica, los fluidos del procedimiento pueden variar dependiendo de las fases particulares de la síntesis de ácidos nucleicos y dependiendo del protocolo sintético utilizado. Por ejemplo, en la fase de síntesis de fragmentos 22, los fluidos del procedimiento pueden incluir una polimerasa (p. ej., TdT) y nucleótidos de los cuatro tipos de ADN comunes. En la secuenciación (p. ej., secuenciación por síntesis, o SBS) de la fase de control de calidad de la síntesis 28 y la fase de control de calidad del ensamblaje 36, los fluidos del procedimiento pueden incluir una polimerasa y nucleótidos marcados de los cuatro tipos comunes de ADN. Los nucleótidos usados en la fase de secuenciación pueden diferir de los usados en la fase de síntesis al incluir marcadores tales como marcadores fluorescentes únicos. Los marcadores fluorescentes permiten que el sistema de detección 60 detecte qué nucleótidos se añadieron los últimos a las sondas hibridadas con ácidos nucleicos molde en sitios individuales del sustrato 24. Los nucleótidos usados en una o ambas de las fases de síntesis y secuenciación pueden incluir restos de bloqueo reversibles. Los restos de bloqueo reversibles evitan la adición de más de un nucleótido por ciclo de síntesis o secuenciación, respectivamente. En otros métodos de secuenciación, tales como secuenciación por ligado, los fluidos del procedimiento pueden incluir sondas de oligonucleótidos de consulta con marcadores fluorescentes únicos unidos a las mismas. De manera similar, las sondas de consulta se unirán a los moldes en cada sitio en una configuración que permite el ligado de las sondas de consulta a un cebador de anclaje y pueden ser detectadas por el sistema de detección 60 para secuenciar los moldes en cada sitio.

En algunos aspectos de la descripción, el sistema de control/suministro de fluidos 58 puede suministrar al sistema de sintetizador/secuenciador/ensamblador 56 una pluralidad de fluidos de procedimiento en varios momentos. Por ejemplo, durante la secuenciación, el sistema de control/suministro de fluidos 58, después de suministrar los fluidos de procedimiento que incluyen una polimerasa y nucleótidos marcados de los cuatro tipos comunes que se encuentran naturalmente en el ADN, puede suministrar fluidos de procedimiento que incluyen reactivos para eliminar los restos de bloqueo reversibles de los nucleótidos, escindiendo conectores de nucleótidos, o para eliminar bases de oligonucleótidos ligados para liberar un nuevo extremo de sonda extensible. Configuraciones de detección y de fluidos de ejemplo que se pueden usar en los métodos y dispositivos expuestos en el presente documento se describen en los documentos WO 07/123744; publicaciones de solicitudes de patentes de EE.UU. Nº 2012/0270305 A1; 2013/0023422 A1; y 2013/0260372 A1; y patentes de EE.UU. Nº 5.528.050; 5.719.391; 8.158.926 y 8.241.573. En aspectos particulares de la descripción, las plataformas de secuenciación disponibles comercialmente pueden modificarse para llevar a cabo una función de "escritura" de ácido nucleico, como se expone en el presente documento, además de la función de secuenciación (es decir, "lectura"). Las plataformas de ejemplo incluyen, pero no se limitan a plataformas HiSeq®, MiSeq® y NextSeq™ (Illumina, Inc. San Diego, CA), plataformas SOLiD® e Ion Torrent® (Thermo Fisher, Waltham, MA); plataformas de pirosecuenciación (Roche, Basilea Suiza) y plataformas de nanoporos (Oxford Nanopore, Oxford Inglaterra). Como se proporciona en el presente documento, el sistema de control/suministro de fluidos 58 puede controlar el suministro de reactivos de secuenciación basándose en si el sistema 50 determina que la fase de síntesis está completa (p. ej., mediante señales generadas por la incorporación de nucleótidos y detectadas por el sistema de sintetizador/secuenciador/ensamblador 56) o si el sistema 50 determina que de la fase de síntesis ha resultado un producto de calidad suficiente. Alternativamente, el sistema de suministro/control de fluidos 58 puede reiniciar todo o parte de un protocolo de síntesis para crear nuevos fragmentos de ácido nucleico y/o nuevos polinucleótidos ensamblados.

Como se señaló anteriormente, en algunos aspectos de la descripción, los polinucleótidos sintetizados (p. ej., los polinucleótidos sintetizados 26a, 26b, 26c, 26d, los polinucleótidos ensamblados 34) pueden amplificarse sobre el sustrato 24 antes de la secuenciación. Este procedimiento puede incluir la amplificación de fragmentos de ADN o ARN sobre un soporte para crear una multitud de sitios de ADN o fragmentos de ARN cuya secuencia se determina mediante el procedimiento de secuenciación. Los métodos de ejemplo para producir sitios de ácidos nucleicos amplificados adecuados para la secuenciación incluyen, pero no se limitan a amplificación de círculo rodante (RCA) (Lizardi et al., *Nat. Genet.* 19: 225-232 (1998)), amplificación de puente (Adams y Kron, Method for Performing Amplification of Nucleic Acid with Two Primers Bound to a Single Solid Support, Mosaic Technologies, Inc. (Winter Hill, MA); Instituto Whitehead de Investigación Biomédica, Cambridge, MA, (1997); Adessi et al., *Nucl. Acids Res.* 28:E87

(2000); Pemov et al., *Nucl. Acids Res*. 33: e11(2005); o US 5.641.658), generación polony (Mitra et al., *Proc. Natl. Acad. Sci. USA* 100: 5926-5931 (2003); Mitra et al., *Anal. Biochem*. 320:55-65 (2003)), o amplificación clonal en perlas usando emulsiones (Dressman et al., *Proc. Natl. Acad. Sci. USA* 100:8817-8822 (2003)) o ligado a bibliotecas de adaptadores basadas en perlas (Brenner et al., *Nat. Biotechnol*. 18: 630-634 (2000); Brenner et al., *Proc. Natl. Acad. Sci. USA* 97:1665-1670 (2000)); Reinartz et al., *Brief Funct. Genomic Proteomic* 1: 95-104 (2002)).

El sintetizador 52 también incluye un sistema de control de calidad 62. La calidad de las muestras (p. ej., los polinucleótidos sintetizados 26a, 26b, 26c, 26d, los polinucleótidos ensamblados 34), la calidad de los datos obtenidos por el sistema de síntesis 50, y los diversos parámetros utilizados para procesar las muestras (p. ej., secuenciación) pueden ser evaluados y/o controlados por el sistema de control de calidad 26. Por ejemplo, como se indicó anteriormente, el sistema de control de calidad puede recibir información de secuenciación y puede determinar una tasa de error de síntesis y/o el rendimiento de una secuencia sintetizada predeterminada y/o el rendimiento de secuencias sintetizadas con una tasa de error predeterminada, que a su vez puede ser utilizada por el sistema de control de calidad 62 para determinar cómo de bien se diseñó el fragmento y/o para determinar la recogida de los polinucleótidos sintetizados y/o ensamblados (p. ej., qué grupo recoger). En ciertos aspectos de la descripción, el sistema de control de calidad 62 puede comunicarse con la interfaz de control/operador del sistema 54 para proporcionar retroalimentación a las fases de diseño de fragmentos 10 de modo que se pueda llevar a cabo un diseño de fragmentos adaptativo, ya sea automáticamente o por un usuario. El sistema de control de calidad 62 puede incluir uno o más procesadores programados, u ordenadores de propósito general o de aplicación específica que se comunican con sensores y otros sistemas de procesamiento dentro del sistema de control de fluidos/suministro de muestras 58, el sistema de sintetizador/secuenciador/ensamblador 56, y el sistema de detección 62.

La Fig. 3 ilustra una técnica de síntesis de fragmentos 70 para generar un fragmento de polinucleótido monocatenario independiente de molde de una molécula de polinucleótido sintética con retroalimentación de secuenciación según aspectos de la presente descripción. Cabe señalar que, aunque se ilustra un polinucleótido monocatenario en la Fig. 3, el polinucleótido bicatenario también está contemplado en la técnica de síntesis de fragmentos 70 descrita en el presente documento. La técnica de síntesis de fragmentos 70 comienza con un sustrato 72 (p. ej., el sustrato 24 de la Fig. 1) (fase 74).

El sustrato 72 puede modificarse de manera que se inmovilice una pluralidad de cebadores (p. ej., un cebador P5 76, un cebador P7 78) sobre una superficie del sustrato 72 (fase 79). Los cebadores P5 y P7 se describen en la patente de EE.UU. Nº 8.563.477 y Bentley et al., *Nature* 456: 53-59 (2008). Aunque se ilustran una copia del cebador P5 76 y una copia del cebador P7 78 en la Fig. 3, cabe señalar que pueden inmovilizarse múltiples copias del cebador P5 76 y el cebador P7 78, o múltiples copias de otros cebadores adecuados, sobre el sustrato 72. Como se analiza en detalle a continuación, la densidad de los cebadores sobre el sustrato 72 se puede controlar.

La síntesis de ácidos nucleicos se puede llevar a cabo en el extremo 3' del cebador P5. La síntesis de ácido nucleico de un fragmento de polinucleótido monocatenario se basa en la secuencia de un fragmento de un ácido nucleico diana. Así pues, una serie de nucleótidos individuales puede fluir secuencialmente sobre el sustrato 72 de modo que los nucleótidos individuales se incorporen sucesivamente en el extremo 3' del cebador P5 76 para generar un polinucleótido extendido 80 (fase 82). El orden de suministro para los tipos de nucleótidos individuales se basa en la secuencia del fragmento de ácido nucleico diana (p. ej., desde un extremo 5' a un extremo 3') o el complemento del mismo. Debido a que el polinucleótido extendido 80 se sintetiza sin una cadena molde, esta síntesis puede denominarse síntesis independiente del molde. Como se indicó anteriormente, la síntesis del polinucleótido extendido 80 puede llevarse a cabo en presencia de una polimerasa, tal como TdT. Además, los nucleótidos pueden ser análogos que tienen un resto de bloqueo reversible para evitar la incorporación de más de un nucleótido al cebador por ciclo. En este caso, un ciclo puede incluir una etapa de eliminación del resto de bloqueo del cebador extendido para regenerar un oxígeno 3' que sea capaz de incorporar un nuevo análogo de nucleótido en un ciclo posterior.

En algunos aspectos de la descripción, la síntesis del fragmento de polinucleótido monocatenario se puede llevar a cabo en el extremo 3' de un adaptador escindible en lugar del cebador P5 76. Por ejemplo, el adaptador escindible puede incluir una secuencia complementaria del cebador P5 76 (p. ej., un adaptador P5') de manera que el adaptador escindible puede hibridarse con el cebador P5 inmovilizado 76. Así pues, el adaptador escindible también se inmoviliza al sustrato 72. Similar a la síntesis ilustrada anteriormente con respecto a la extensión del extremo 3' del cebador P5 76, la síntesis se puede llevar a cabo en el extremo 3' del adaptador escindible para generar el polinucleótido extendido 80. Como se ilustró previamente en el presente documento, se pueden usar otros métodos de síntesis de ácidos nucleicos para crear fragmentos de ácido nucleico, incluyendo, pero no limitados a métodos de síntesis química conocidos.

Después de la síntesis del polinucleótido extendido 80 sobre el sustrato 72, se puede llevar a cabo una fase de control de calidad (p. ej., la fase de control de calidad de la síntesis 28 en la Fig. 1). Como se señaló anteriormente, en ciertos aspectos de la descripción, el polinucleótido extendido 80 puede amplificarse antes de la fase de control de calidad. Un ejemplo de tal amplificación 84 se ilustra en la Fig. 3 con un cuadro discontinuo, que representa una fase opcional. Otras fases pueden ser opcionales para algunos aspectos de la descripción.

Como se ilustra, la amplificación 82 del polinucleótido extendido 80 puede ser por amplificación de puente. Por ejemplo, se puede ligar un adaptador P7' 86 al extremo 3' del polinucleótido extendido 80 (fase 88). El adaptador P7' 86 comprende una secuencia complementaria de una secuencia del cebador P7 78. El polinucleótido extendido monocatenario 80 puede

formar puente de manera que el adaptador P7' 86 pueda hibridarse con el cebador P7 78 inmovilizado sobre el sustrato 72 (fase 90). Se puede sintetizar una segunda cadena 92 que se extiende desde un extremo 3' del cebador P7 78 usando el polinucleótido extendido 80 y el adaptador P7' 86 como molde (fase 98). Así pues, se forma un ácido nucleico bicatenario 96 que incluye el polinucleótido extendido 80 y la segunda cadena complementaria 92. El extremo 3' de la segunda cadena 92 puede incluir un adaptador P5' 94 que es complementario del cebador P5 76. El ácido nucleico bicatenario 96 después se puede desnaturalizar de modo que se pueden formar dos polinucleótidos monocatenarios (p. ej., el polinucleótido 80 extendido y la segunda cadena 92) con un extremo de cada uno de los dos polinucleótidos monocatenarios unido al sustrato 72 (fase 100). Las etapas de amplificación se pueden repetir y se puede formar un grupo de amplificación 102 del polinucleótido extendido 80 sobre el sustrato 72 (fase 104).

Independientemente de si se lleva a cabo la amplificación 84 del polinucleótido extendido 80, el polinucleótido extendido 80 puede secuenciarse (fase 106). En los aspectos ilustrados de la descripción, la secuenciación puede ser una secuenciación por síntesis, o SBS, técnica en la que los nucleótidos marcados 108 de los cuatro tipos comunes de ADN/ARN pueden fluir sobre el sustrato. Cada uno de los nucleótidos marcados 108 puede tener un marcador fluorescente único y un resto de bloqueo reversible unido a él. El marcador fluorescente permite que un sistema de detección (p. ej., el sistema de detección 60 de la Fig. 2) detecte qué nucleótidos se añadieron los últimos a las sondas hibridadas con el ácido nucleico molde (p. ej., el polinucleótido extendido 80) en sitios individuales sobre el sustrato 72. El resto de bloqueo reversible evita la adición de más de un nucleótido por ciclo en cada sitio. Como se analiza en detalle a continuación, la secuenciación del polinucleótido extendido 80 puede llevarse a cabo con cualquier otro método adecuado, incluida la secuenciación por ligación, pirosecuenciación, secuenciación por detección de iones de hidrógeno, secuenciación de nanoporos o similares.

Los ácidos nucleicos preparados mediante los métodos expuestos en el presente documento se pueden secuenciar proporcionando diferentes nucleótidos (u oligonucleótidos) a una matriz de características, cada una de las cuales tiene un ácido nucleico sintetizado para así producir diferentes señales en cada característica, correspondiendo cada señal a una especie específica de nucleótido (u oligonucleótido) que se ha añadido a la característica. Por ejemplo, en un caso en el que se utilizan cuatro marcadores diferentes, que corresponden a cuatro especies diferentes de nucleótidos (u oligonucleótidos), se pueden adquirir imágenes individuales, en donde cada imagen captura un solo color (u otro tipo de señal) entre los cuatro colores posibles (o tipos de señales). En este ejemplo, el color de la señal es diferente para cada una de las cuatro imágenes diferentes, produciendo así un ciclo de cuatro imágenes de colores que corresponde a los cuatro posibles nucleótidos presentes en una posición particular en el ácido nucleico. En ciertos aspectos, dichos métodos pueden incluir además proporcionar nucleótidos (u oligonucleótidos) marcados adicionales a la matriz de moléculas, produciendo así una pluralidad de ciclos de imágenes de colores. Algunas técnicas de secuenciación de ejemplo que producen imágenes de múltiples ciclos y, a menudo, múltiples imágenes por ciclo, se exponen a continuación con más detalle.

Las técnicas de SBS generalmente implican la extensión enzimática de una cadena de ácido nucleico naciente mediante la adición iterativa de nucleótidos contra una cadena molde. En los métodos tradicionales de SBS, se puede proporcionar un monómero de un solo nucleótido a los ácidos nucleicos diana en presencia de una polimerasa en cada etapa de suministro. Sin embargo, en los métodos descritos en el presente documento, se puede proporcionar más de un tipo de monómero de nucleótidos a un ácido nucleico diana en presencia de una polimerasa en cada una de las etapas de suministro.

La SBS puede utilizar monómeros de nucleótidos que tienen un resto terminador o aquellos que carecen de cualquier resto terminador. Los métodos que utilizan monómeros de nucleótidos que carecen de terminadores incluyen, por ejemplo, pirosecuenciación y secuenciación utilizando nucleótidos marcados con gamma-fosfato, como se expone con más detalle a continuación. En los métodos que utilizan monómeros de nucleótidos que carecen de terminadores, el número de nucleótidos añadidos a una cadena de ácido nucleico naciente por la polimerasa en cada ciclo es generalmente variable y depende de la secuencia del molde y del modo de suministro de nucleótidos. Para las técnicas de SBS que utilizan monómeros de nucleótidos que tienen un resto terminador, el terminador puede ser efectivamente irreversible en las condiciones de secuenciación utilizadas, como es el caso de la secuenciación tradicional de Sanger que utiliza didesoxinucleótidos, o el terminador puede ser reversible como es el caso de los métodos de secuenciación comercialmente disponibles en Illumina, Inc. (San Diego, CA) o descritos en Bentley et al., *Nature* 456:53-59 (2008), WO 04/018497; WO 91/06678; WO 07/123744; patentes de EE.UU. Nº 7.057.026; 7.329.492; 7.211.414; 7.315.019 o 7.405.281, y publicación de solicitud de EE.UU. Nº 2008/0108082 A1.

De forma alternativa o adicional al uso de restos terminadores, las técnicas de SBS pueden utilizar monómeros de nucleótidos que tienen un resto marcador o aquellos que carecen de un resto marcador. Por consiguiente, los sucesos de incorporación pueden detectarse basándose en una característica del marcador, tal como la fluorescencia del marcador; una característica del monómero de nucleótidos, tal como el peso molecular o la carga; un subproducto de la incorporación del nucleótido, tal como la liberación de hidrógeno o pirofosfato; o similar. En aspectos de la descripción, donde dos o más nucleótidos diferentes están presentes simultáneamente en un reactivo de secuenciación o ácido nucleico extendido, los diferentes nucleótidos pueden distinguirse entre sí. Por ejemplo, los diferentes nucleótidos presentes en un reactivo de secuenciación o producto de extensión pueden tener diferentes marcadores y se pueden distinguir usando técnicas ópticas apropiadas como se ilustra en las plataformas de secuenciación comercialmente disponibles de Illumina, Inc. (San Diego, CA) o descritas en las publicaciones de solicitudes de patentes de EE.UU. Nº 2012/0270305 A1; 2013/0023422 A1; y 2013/0260372 A1; y patentes de EE.UU.

Nº 5.528.050; 5.719.391; 8.158.926 y 8.241.573.

Preferiblemente, en los aspectos de secuenciación basados en terminadores reversibles de la descripción, los marcadores no inhiben sustancialmente la extensión en las condiciones de reacción de SBS. Sin embargo, los marcadores de detección pueden eliminarse, por ejemplo, por escisión o degradación. Se pueden capturar imágenes después de la incorporación de marcadores en características de matrices de ácidos nucleicos sintetizados (p. ej., fragmentos de ácido nucleico o ácidos nucleicos ensamblados). En aspectos particulares de la descripción, cada ciclo implica el suministro simultáneo de cuatro tipos de nucleótidos diferentes a la matriz y cada tipo de nucleótido tiene un marcador espectralmente distinto. A continuación, se pueden obtener cuatro imágenes, cada una de las cuales utiliza un canal de detección que es selectivo para uno de los cuatro marcadores diferentes. Alternativamente, se pueden añadir secuencialmente diferentes tipos de nucleótidos y se puede obtener una imagen de la matriz entre cada etapa de adición. En cualquiera de los aspectos de la descripción, cada imagen mostrará características de los ácidos nucleicos que han incorporado nucleótidos de un tipo particular. Diferentes características estarán presentes o ausentes en las diferentes imágenes debido al diferente contenido de secuencia de cada característica. Sin embargo, la posición relativa de las características permanecerá sin cambios en las imágenes. Las imágenes obtenidas a partir de dichos métodos de terminador reversible-SBS se pueden almacenar, procesar y analizar como se expone en el presente documento. Después de la etapa de captura de imágenes, los marcadores se pueden eliminar y los restos de terminadores reversibles se pueden eliminar para los ciclos posteriores de adición y detección de nucleótidos. La eliminación de los marcadores después de que se hayan detectado en un ciclo particular y antes de un ciclo posterior puede proporcionar la ventaja de reducir la señal de fondo y la diafonía entre ciclos.

Los sistemas y métodos de SBS de ejemplo adicionales que se pueden utilizar con los métodos y sistemas descritos en el presente documento se describen en la publicación de solicitud de patente de EE.UU. Nº 2007/0166705, publicaciones de solicitudes de patentes de EE.UU. Nº 2006/0188901, patente de EE.UU. Nº 7.057.026, publicación de solicitud de patente de EE.UU. Nº 2006/0240439, publicación de solicitud de patente de EE.UU. Nº 2006/0281109, Publicación PCT Nº WO 05/065814, publicación de solicitud de patente de EE.UU. Nº 2005/0100900, Publicación PCT Nº WO 06/064199 y Publicación PCT Nº WO 07/010251.

Aspectos de secuenciación particulares de la descripción pueden utilizar técnicas de pirosecuenciación. La pirosecuenciación detecta la liberación de pirofosfato inorgánico (PPi) a medida que se incorporan nucleótidos particulares en la cadena naciente (Ronaghi et al. (1996) *Analytical Biochemistry* 242(1), 84-9; Ronaghi, M. (2001) *Genome Res*. 11(1), 3-11; Ronaghi et al. (1998) *Science* 281(5375), 363; patente de EE.UU. Nº 6.210.891; patente de EE.UU. Nº 6.258.568 y patente de EE.UU. Nº 6.274.320). En la pirosecuenciación, el PPi liberado puede detectarse al ser convertido en trifosfato de adenosina (ATP) por la ATP sulfurilasa, y el nivel de ATP generado puede detectarse mediante fotones producidos por luciferasa. Los ácidos nucleicos sintetizados (p. ej., fragmentos o ácidos nucleicos ensamblados) se pueden unir en las características respectivas en una matriz y se pueden generar imágenes de la matriz para capturar las señales quimioluminiscentes que se producen debido a la incorporación de nucleótidos en las características de la matriz. Se puede obtener una imagen después de que la matriz se trate con un tipo de nucleótido particular (p. ej., A, T, C o G). Las imágenes obtenidas después de la adición de cada tipo de nucleótido diferirán con respecto a las características de la matriz que se detecten. Estas diferencias en la imagen reflejan el contenido de secuencia diferente de las características en la matriz. Sin embargo, al igual que con otros métodos de SBS, las ubicaciones relativas de cada característica permanecerán sin cambiar en las imágenes. Las imágenes se pueden analizar usando los sistemas y métodos expuestos en el presente documento. Por ejemplo, las imágenes obtenidas después del tratamiento de la matriz con cada tipo de nucleótido diferente se pueden manejar de la misma manera que se ilustra en el presente documento para las imágenes obtenidas para métodos de secuenciación basados en terminadores reversibles.

La detección de pirofosfato se puede utilizar para controlar la síntesis de ácidos nucleicos en tiempo real. Por ejemplo, se puede usar un método de síntesis basado en TdT para sintetizar fragmentos de ácido nucleico, como se ha expuesto anteriormente. A medida que se añade cada nucleótido, se libera un pirofosfato. La cantidad de pirofosfato liberado en cada ciclo de síntesis se puede detectar y cuantificar para determinar el rendimiento y la eficiencia de la síntesis. De esta manera, la detección de pirofosfato se puede evaluar de manera similar a cómo se evalúa el DMT en la síntesis química.

Algunos aspectos de la descripción implican la secuenciación mediante técnicas de ligación. Dichas técnicas utilizan ADN ligasa para incorporar oligonucleótidos y luego se pueden identificar los oligonucleótidos incorporados. Los oligonucleótidos tienen típicamente diferentes marcadores que se correlacionan con la identidad de un nucleótido particular en una secuencia con la que se hibridan los oligonucleótidos. Al igual que con otros métodos de SBS, se puede usar una matriz de características a las que se unen fragmentos de ácido nucleico o ácidos nucleicos ensamblados y se pueden obtener imágenes después del tratamiento de la matriz con los reactivos de secuenciación marcados. Cada imagen mostrará características de ácidos nucleicos que han incorporado marcadores de un tipo particular. Diferentes características estarán presentes o ausentes en las diferentes imágenes debido al diferente contenido de secuencia de cada característica, pero la posición relativa de las características permanecerá sin cambios en las imágenes. Las imágenes obtenidas a partir de métodos de secuenciación basados en ligación se pueden almacenar, procesar y analizar como se expone en el presente documento. Se describen ejemplos de secuenciación por sistemas de ligación y métodos que se pueden utilizar con los métodos y sistemas descritos en el presente documento en la patente de EE.UU. Nº 6.969.488, patente de EE.UU. Nº 6.172.218, y patente de EE.UU. Nº 6.306.597.

Se apreciará que cualquiera de los procedimientos de secuenciación descritos anteriormente se puede incorporar en los métodos y/o sistemas descritos en el presente documento. Además, se apreciará que otros procedimientos de secuenciación conocidos pueden modificarse fácilmente para su uso con los métodos y/o sistemas descritos en el presente documento. Después de la secuenciación del polinucleótido extendido 80, se evalúa una secuencia del polinucleótido extendido 80. Cuando se compara la secuencia del polinucleótido extendido 80 con la secuencia del ácido nucleico diana, se puede obtener una tasa de error de síntesis como se ha ilustrado anteriormente. También como se ha ilustrado anteriormente, se puede obtener un rendimiento de una secuencia sintetizada predeterminada (p. ej., una secuencia exenta de errores) y/o un rendimiento de secuencias sintetizadas con una tasa de error predeterminada (p. ej., cuando el grupo amplificado del polinucleótido extendido 80 está presente). Dicho rendimiento y/o la tasa de error de síntesis se pueden usar para el control de calidad de la síntesis de ácidos nucleicos (p. ej., la fase de control de calidad de la síntesis 28 de la Fig. 1), por ejemplo, para determinar cómo de bueno era el diseño del fragmento y/o para determinar cómo recoger el polinucleótido extendido 80. Por ejemplo, basado en la tasa de error y/o el rendimiento, el polinucleótido extendido 80 puede o no escindirse del sustrato 72. En los aspectos ilustrados de la descripción, el polinucleótido extendido 80 se escinde del cebador P5 76 (fase 110).

La Fig. 4 ilustra una técnica de síntesis y ensamblaje de fragmentos 120 para ensamblar múltiples fragmentos de polinucleótido de una molécula de polinucleótido sintética con retroalimentación de secuenciación para los fragmentos y el polinucleótido sintético ensamblado de acuerdo con aspectos de la presente descripción. La técnica de síntesis y ensamblaje de fragmentos 120 puede comenzar con una fase de diseño de fragmentos 122 (p. ej., la fase de diseño de fragmentos 10 ilustrada en la Fig. 1). La fase de diseño de fragmentos 122 puede incluir el uso de un ácido nucleico diana 124 para diseñar secuencias para uno o más fragmentos de ácido nucleico (p. ej., fragmentos 126a, 126b, 126c, 126d). Aunque se ilustran cuatro fragmentos de ácido nucleico (p. ej., fragmentos 126a, 126b, 126c, 126d) en la Fig. 4, se entiende que se puede diseñar cualquier número (p. ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más) de secuencias de fragmentos de ácido nucleico basándose en la secuencia de ácido nucleico diana 124. Los fragmentos de ácido nucleico 126a, 126b, 126c, 126d pueden diseñarse igual o de forma similar al procedimiento ilustrado por la fase de diseño de fragmentos 10 de la Fig. 1. Además, cada uno de los fragmentos de ácido nucleico 126a, 126b, 126c, 126d puede sufrir los mismos o similares procedimientos de secuenciación y síntesis como se ilustra en la Fig. 3 con respecto a un fragmento de ácido nucleico monocatenario individual.

Basándose en la secuencia de cada uno de los fragmentos de ácido nucleico 126a, 126b, 126c, 126d, se puede sintetizar un polinucleótido 128a, 128b, 128c, 128d correspondiente sobre un sustrato 130 (fase 132). Los polinucleótidos sintetizados 128a, 128b, 128c, 128d corresponden a los fragmentos de ácido nucleico 126a, 126b, 126c, 126d, respectivamente. La síntesis de los polinucleótidos sintetizados 128a, 128b, 128c, 128d puede ser igual o similar a la síntesis del polinucleótido extendido 80 de la Fig. 3. En algunos aspectos de la descripción, los polinucleótidos sintetizados 128a, 128b, 128c, 128d pueden amplificarse para generar los correspondientes grupos de amplificación 134a, 134b, 134c, 134d en el sustrato 130 (fase 136). En ciertos aspectos de la descripción, dicha amplificación es opcional. Las siguientes fases, aunque se ilustran para los grupos de amplificación 134a, 134b, 134c, 134d, pueden llevarse a cabo en los polinucleótidos sintetizados 128a, 128b, 128c, 128d.

Los grupos de amplificación 134a, 134b, 134c, 134d pueden secuenciarse en paralelo o secuencialmente en presencia de uno o más reactivos de secuenciación 139 (fase 138). Como se discutió anteriormente, las tasas de error de síntesis y/o los rendimientos (o las tasas de error de amplificación y/o los rendimientos) con respecto a los grupos de amplificación 134a, 134b, 134c, 134d pueden obtenerse a partir de la secuenciación. En algunos aspectos de la descripción, las tasas de error de síntesis (o amplificación) y/o los rendimientos pueden usarse para proporcionar retroalimentación a la fase de diseño de fragmentos 122 para determinar cómo de bueno era el diseño del fragmento. En ciertos aspectos de la descripción, se puede proporcionar una pluralidad de ciclos de dicha retroalimentación de modo que se pueda llevar a cabo una rutina de diseño de fragmentos adaptativa, automáticamente o por un usuario.

Cuando se determina que el diseño del fragmento es deseable, se pueden recoger uno o más de los grupos de amplificación 134a, 134b, 134c, 134d del sustrato 130. En algunos aspectos de la descripción, no se recoge el primer grupo de amplificación (p. ej., el grupo de amplificación 134a) del sustrato 130 mientras que se recogen otros grupos de amplificación (p. ej., los grupos de amplificación 134b, 134c, 134d) (fase 140). En tales casos, los grupos de amplificación 134a, 134b, 134c, 134d se pueden ensamblar en un sustrato 130 para generar polinucleótidos ensamblados 150 (fase 142). En otros aspectos de la descripción, se recogen todos los grupos de amplificación 134a, 134b, 134c, 134d del sustrato 130 (fase 144).). En tales casos, los grupos de amplificación 134a, 134b, 134c, 134d se pueden ensamblar en un sustrato 146 diferente o en el mismo sustrato 130 para generar polinucleótidos ensamblados 152 (fase 148).

Como se discutió anteriormente, las tasas de error de síntesis (o amplificación) y/o los rendimientos con respecto a los grupos de amplificación 134a, 134b, 134c, 134d pueden usarse para determinar cómo se recogen los grupos de amplificación 134a, 134b, 134c, 134d. Por ejemplo, la totalidad o una parte de cada uno de los grupos de amplificación 134a, 134b, 134c, 134d se puede recoger basándose en las tasas de error de síntesis y/o los rendimientos.

Independientemente de cómo se puedan recoger y ensamblar los grupos de amplificación 134a, 134b, 134c, 134d, los polinucleótidos ensamblados 150 o 152 se pueden procesar para la reparación de huecos para generar un grupo de polinucleótidos ensamblados 154 unido al sustrato 130 o 146 (fase 156). Una fase de control de calidad del ensamblaje 158 puede seguir al ensamblaje de fragmentos. Como se ilustra, después el grupo de polinucleótidos

ensamblados 154 puede secuenciarse en presencia de uno o más reactivos de secuenciación 157 durante una fase de control de calidad de ensamblaje 158. De manera similar, se pueden obtener las tasas de error de síntesis (o amplificación) y/o los rendimientos con respecto al grupo de polinucleótidos ensamblados 154 de la secuenciación. En algunos aspectos de la descripción, las tasas de error de síntesis (o amplificación) y/o los rendimientos pueden usarse para proporcionar retroalimentación a la fase de diseño de fragmentos 122 para determinar cómo de bueno era el diseño del fragmento. En ciertos aspectos de la descripción, se puede proporcionar una pluralidad de ciclos de dicha retroalimentación de modo que se pueda llevar a cabo una rutina de diseño de fragmentos adaptativa, automáticamente o por un usuario.

Cuando se determina que el diseño del fragmento es deseable, se puede recoger todo o una parte del grupo de polinucleótidos ensamblados 154 del sustrato 130 o 146 (fase 160). La recolección puede basarse en las tasas de error de síntesis (o amplificación) y/o los rendimientos con respecto al grupo de polinucleótidos ensamblados 154.

La Fig. 5 ilustra una técnica de ensamblaje de fragmentos de polinucleótidos 170 con retroalimentación de secuenciación según aspectos de la presente descripción. La técnica de ensamblaje de fragmentos de polinucleótidos 170 puede comenzar con una fase de diseño de fragmentos 172 (p. ej., la fase de diseño de fragmentos 10 ilustrada en la Fig. 1). La fase de diseño de fragmentos 172 puede incluir el uso de un ácido nucleico diana 174 para diseñar uno o más fragmentos de ácido nucleico (p. ej., fragmentos 176a, 176b, 176c, 176d). Aunque se ilustran cuatro fragmentos de ácido nucleico (p. ej., fragmentos 176a, 176b, 176c, 176d) en la Fig. 5, se entiende que se puede diseñar cualquier número (p. ej., al menos 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más) de fragmentos de ácido nucleico basándose en la secuencia del ácido nucleico diana 174. Los fragmentos de ácido nucleico 176a, 176b, 176c, 176d pueden diseñarse igual o de forma similar al procedimiento ilustrado por la fase de diseño de fragmentos 10 de la Fig. 1.

Basándose en los cuatro fragmentos de ácido nucleico diseñados 176a, 176b, 176c, 176d, se pueden sintetizar los correspondientes fragmentos de ácido nucleico 179a, 179b, 179c, 179d. Por ejemplo, los fragmentos de ácido nucleico sintetizados 179a, 179b, 179c, 179d se dirigen a los fragmentos de ácido nucleico diseñados 176a, 176b, 176c, 176d, respectivamente. La síntesis de los fragmentos de ácido nucleico 179a, 179b, 179c, 179d se puede llevar a cabo de cualquier manera adecuada, incluidos los métodos descritos en relación con la Fig. 4. Además, el primer fragmento de ácido nucleico 179a y el último fragmento de ácido nucleico 179d pueden modificarse adicionalmente (fase 178). Como se ilustra, se puede ligar un adaptador escindible 180 a un extremo 5' del primer fragmento de ácido nucleico 179a, y se puede ligar un adaptador escindible 182 a un extremo 3' del último fragmento de ácido nucleico 179d. Los adaptadores escindibles 180, 182 pueden ser cualquier adaptador adecuado que tenga secuencias complementarias de los cebadores que pueden inmovilizarse en un sustrato 184. Por ejemplo, como se indicó anteriormente, se pueden inmovilizar un cebador P5 186 y un cebador P7 188 al sustrato 184. Así pues, el adaptador escindible 180 puede tener una secuencia de P5' que es complementaria de la secuencia del cebador P5 186, y el adaptador escindible 182 puede tener una secuencia de P7' que es complementaria de la secuencia del cebador P7 188.

Los fragmentos de ácido nucleico sintetizados 179a, 179b, 179c, 179d después se pueden hacer fluir sobre el sustrato 184 en paralelo o secuencialmente (fase 192). Debido a que el adaptador escindible 180 en el extremo 5' del fragmento de ácido nucleico 179a es complementario del cebador P5 186 inmovilizado en el sustrato 184, el fragmento de ácido nucleico 179a puede hibridarse con el cebador P5 186 de modo que el fragmento de ácido nucleico 179a se una al sustrato 184. Otros fragmentos de ácido nucleico 179b, 179c, 179d pueden ensamblarse luego al fragmento de ácido nucleico 179a (fase 194). El ensamblaje de los fragmentos de ácido nucleico 179a, 179b, 179c, 179d puede llevarse a cabo en presencia de cualquier enzima adecuada, que incluye, pero no se limita a ADN recombinasa, proteína beta o cualquier combinación de las mismas.

Después de la fase de ensamblaje de fragmentos 194, los fragmentos de ácido nucleico ensamblados 179a, 179b, 179c, 179d pueden repararse para generar un polinucleótido bicatenario 196 (fase 198). Por ejemplo, los huecos entre los fragmentos de ácido nucleico en cada cadena única 200 o 202 pueden rellenarse utilizando la otra cadena única 202 o 200 como molde. La fase de reparación de fragmentos 198 se puede llevar a cabo en presencia de cualquier enzima adecuada, que incluye, pero no se limita a ADN polimerasa, ADN ligasa, Circligase o cualquier combinación de las mismas.

A continuación, el polinucleótido bicatenario 196 se puede desnaturalizar (fase 204). Debido a que la cadena 200 se hibrida con la otra cadena 202 que está inmovilizada sobre el sustrato 184, la cadena 200 se puede eliminar por lavado del sustrato 184 tras la desnaturalización, y la cadena 202 permanecerá sobre el sustrato 184. La cadena 202 puede entonces amplificarse. Por ejemplo, como se ilustra, la cadena 202 puede sufrir una amplificación de puente en la que el adaptador escindible 182 se hibrida con el cebador P7 188 (fase 206), y luego se puede sintetizar una segunda cadena utilizando la cadena 202 como molde para generar un polinucleótido bicatenario (p. ej., similar a la fase 98 ilustrada en la Fig. 3). La amplificación de la cadena 202 también se puede llevar a cabo usando amplificación por círculo rodante (RCA), o cualquier otro método adecuado de amplificación en fase sólida como se indicó anteriormente. Como resultado de la amplificación, se puede formar un grupo de amplificación 208 de la cadena 202, que incluye copias de la cadena 202 y su secuencia complementaria, sobre el sustrato 184 (fase 210). En algunos aspectos de la descripción, dicha fase de amplificación (p. ej., las fases 206, 210) es opcional.

Después el grupo de amplificación 208 puede secuenciarse en presencia de uno o más reactivos de secuenciación 212 durante una fase de control de calidad del ensamblaje 214. De manera similar a la anterior, las tasas de error de

síntesis (o amplificación) y/o los rendimientos con respecto al grupo de amplificación 208 se pueden obtener a partir de la secuenciación. En algunos aspectos de la descripción, las tasas de error de síntesis (o amplificación) y/o los rendimientos pueden usarse para proporcionar retroalimentación a la fase de diseño del fragmentos 172 para determinar cómo de bueno era el diseño del fragmento. En ciertos aspectos de la descripción, se puede proporcionar una pluralidad de ciclos de dicha retroalimentación de modo que se pueda llevar a cabo una rutina de diseño de fragmentos adaptativa, automáticamente o por un usuario.

Cuando se determina que el diseño del fragmento es deseable, se puede recoger todo o una parte del grupo de amplificación 208 del sustrato 184 (fase 216). La recolección puede basarse en las tasas de error de síntesis y/o rendimientos con respecto al grupo de polinucleótidos ensamblados 154. Como se ilustra, el grupo de amplificación 208 puede escindirse de los cebadores P5 y P5 en el sustrato 184.

La Fig. 6 ilustra una celda de flujo 220 para sintetizar diferentes secuencias de fragmentos de polinucleótidos en paralelo para su posterior ensamblaje en una molécula de polinucleótido sintética según aspectos de la presente descripción. La celda de flujo 220 incluye un sustrato 221 y uno o más canales de flujo (p. ej., canales de flujo 222a, 222b, 222c, 222d) en los que se pueden introducir diversos fluidos vehículos, reactivos, etc. Aunque se ilustran cuatro canales de flujo en la Fig. 6, se entiende que la celda de flujo 220 puede incluir cualquier número de canales de flujo (p. ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más).

Cada uno de los canales de flujo 222a, 222b, 222c, 222d puede incluir en el sustrato 221 uno o más sitios de reacción en los que puede tener lugar la síntesis, amplificación y/o secuenciación de ácidos nucleicos, como se ha descrito antes. Por ejemplo, cada uno de los canales de flujo 222a, 222b, 222c, 222d incluye tres sitios de reacción 224 como se ilustra. Pueden contemplarse muchas distribuciones diferentes de los sitios de reacción, que incluyen patrones regulares, que se repiten y no regulares. Por ejemplo, las distribuciones de los sitios de reacción pueden incluir distribuciones rectilíneos (es decir, rectangulares), distribuciones triangulares, distribuciones hexagonales, etc. Las distribuciones particulares pueden seguir las enseñanzas de la patente de EE.UU. Nº 7.813.013, y/o de la solicitud de patente de EE.UU. Nº 13/267.565, presentada el 6 de octubre de 2011. Cabe señalar que el sustrato con patrón también se puede usar para controlar la densidad de las características que se pueden consultar (p. ej., a través de la formación de imágenes).

Los cebadores (p. ej., cebadores P5, P7) se pueden hacer fluir hacia cada uno de los canales de flujo 222a, 222b, 222c, 222d de manera que los cebadores se depositen y se unan a los sitios de reacción 224 de cada uno de los canales de flujo 222a, 222b, 222c, 222d. En cada sitio de reacción 224 de cada uno de los canales de flujo 222a, 222b, 222c, 222d, se puede sintetizar un fragmento de ácido nucleico 226 por extensión de los cebadores o unirse a los cebadores, como se ha descrito antes (fase 228). Cada uno de los canales de flujo 222a, 222b, 222c, 222d generalmente incluye un fragmento de ácido nucleico diferente 226. Por ejemplo, los canales de flujo 222a, 222b, 222c, 222d pueden incluir los fragmentos de ácido nucleico 128a, 128b, 128c, 128d de la Fig. 4. En algunos aspectos de la descripción, cualquiera de los canales de flujo 222a, 222b, 222c, 222d puede incluir el mismo fragmento de ácido nucleico que uno o más de los otros canales de flujo 222a, 222b, 222c, 222d.

El fragmento de ácido nucleico 226 en cada sitio de reacción 224 de cada uno de los canales de flujo 222a, 222b, 222c, 222d puede secuenciarse en presencia de uno o más reactivos de secuenciación 230 (fase 232). En algunos aspectos de la descripción, la amplificación de los fragmentos de ácido nucleico 226 puede llevarse a cabo en uno o más de los canales de flujo 222a, 222b, 222c, 222d antes de la secuenciación. De manera similar a la anterior, se pueden obtener las tasas de error de síntesis y/o rendimientos con respecto a los fragmentos de ácido nucleico 226 a partir de la secuenciación. En algunos aspectos de la descripción, las tasas de error de síntesis y/o los rendimientos pueden usarse para proporcionar retroalimentación al diseño de fragmentos para determinar cómo de bueno era el diseño del fragmento. En ciertos aspectos de la descripción, se puede proporcionar una pluralidad de ciclos de dicha retroalimentación de modo que se pueda realizar una rutina de diseño de fragmentos adaptativa, automáticamente o por un usuario.

Se pueden recoger todos o una parte de los fragmentos de ácido nucleico 226 del sustrato 221 (fase 234) basándose en las tasas de error de síntesis (o amplificación) y/o los rendimientos. Como se ilustra, algunos de los fragmentos de ácido nucleico 226 (p. ej., un fragmento de ácido nucleico 236) se recogen selectivamente y posteriormente se separan por lavado de los canales de flujo 222a, 222b, 222c, 222d. Otros fragmentos nucleicos 226 (p. ej., un fragmento de ácido nucleico 238) no se recogen y permanecen unidos al sustrato 221 en los sitios de reacción 224. La recolección puede basarse en las tasas de error de síntesis (o amplificación) y/o rendimientos con respecto a los fragmentos de ácido nucleico 226. Los fragmentos de ácido nucleico recogidos 226 se separan por lavado de los canales de flujo 222a, 222b, 222c, 222d y luego pueden recogerse para su posterior ensamblaje en la molécula de polinucleótido sintético.

La Fig. 7 ilustra un aspecto de la descripción de un casete de escritura y lectura de microfluidos 240 para la síntesis, amplificación y secuenciación de plataforma única de una molécula de polinucleótido sintética según realizaciones de la presente descripción. El casete 240 tiene una forma generalmente redonda como se ilustra, sin embargo, se pueden contemplar muchas otras formas, que incluyen la cuadrada, rectangular, triangular, hexagonal, etc.

El casete 240 incluye un sustrato 242 y uno o más canales de flujo 244. Los canales de flujo 244 están uno al lado del

otro y centrados radialmente hacia una región central 246. Se pueden introducir varios fluidos vehículo, reactivos, etc. (p. ej., a lo largo de una dirección hacia la región central 246 ilustrada por la referencia 248) en los canales de flujo 244, y pueden fluir hacia fuera de los canales de flujo 244 (p. ej., en una dirección generalmente opuesta a la dirección 248). Aunque se ilustran dieciséis canales de flujo en la Fig. 7, se entiende que el casete 240 puede incluir cualquier número de canales de flujo (p. ej., 1, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más).

La región central 246 puede estar separada de los canales de flujo 244 con un borde 250. El borde 250 puede estar abierto (p. ej., cuando se instala) o cerrado (p. ej., cuando se retira) para permitir o bloquear que los fluidos vehículo, reactivos, reaccionantes (p. ej., fragmentos de ácidos nucleicos), etc., fluyan desde los canales de flujo 244 a la región central 246. En algunos aspectos de la descripción, el borde 250 puede incluir una entrada entre la región central 246 y cada uno de los canales de flujo 244. La entrada se puede abrir y cerrar para proporcionar un flujo selectivo desde cualquiera de los canales de flujo 244 a la región central 246.

Cada uno de los canales de flujo 244, similar a los canales de flujo 222a, 222b, 222c, 222d ilustrados en la Fig. 6, puede incluir en el sustrato 242 uno o más sitios de reacción 252 en los que puede tener lugar la síntesis, amplificación y/o secuenciación de ácidos nucleicos, como se ha descrito antes. Por ejemplo, cada uno de los canales de flujo 224 incluye tres sitios de reacción 252 como se ilustra. Las distribuciones de los sitios de reacción 252 son similares a las distribuciones de los sitios de reacción 224 de la Fig. 6.

La síntesis, amplificación y/o secuenciación de ácidos nucleicos puede ocurrir en los canales de flujo 244, similar a los canales de flujo 222a, 222b, 222c, 222d como se describe en la Fig. 6. A diferencia de la celda de flujo 220 de la Fig. 6, sin embargo, el casete 240 incluye la región central 246 que puede usarse para ensamblar y reparar fragmentos de ácido nucleico, tales como los descritos en la fase 32 de la Fig. 1 y las fases 142, 148, 156 de la Fig. 4. Además, la región central 246 también puede usarse para el control de calidad del ensamblaje, tal como el descrito en la fase 36 de la Fig. 1 y la fase 158 de la Fig. 4. Además, el polinucleótido ensamblado (p. ej., a partir de los fragmentos de ácido nucleico) se puede recoger de la región central 246.

La Fig. 8 ilustra un diagrama de flujo de una técnica 260 para generar un fragmento de polinucleótido independiente de molde de una molécula de polinucleótido sintética con retroalimentación de secuenciación según realizaciones de la presente descripción. La técnica 260 puede comenzar proporcionando un cebador inmovilizado sobre un sustrato (bloque 262). La síntesis de un polinucleótido monocatenario se puede llevar a cabo por extensión del cebador (bloque 264). La síntesis del polinucleótido monocatenario puede basarse en una secuencia de ácido nucleico diana.

En algunos aspectos de la descripción, el polinucleótido monocatenario sintetizado puede amplificarse (bloque 266) usando cualquier método adecuado como se ha descrito antes. En otros aspectos de la descripción, dicha amplificación es opcional y, en consecuencia, el bloque 266 se ilustra con un cuadro discontinuo.

A continuación, se puede secuenciar el polinucleótido monocatenario sintetizado (bloque 268). La secuencia del polinucleótido monocatenario se puede comparar con la secuencia de ácido nucleico diana (bloque 270) para el control de la calidad de la síntesis. Como se ha descrito antes, las tasas de error de síntesis (o amplificación) y/o los rendimientos pueden usarse para el control de calidad. Por ejemplo, las tasas de error de síntesis (o amplificación) y/o los rendimientos pueden usarse para proporcionar retroalimentación al diseño de fragmentos para determinar la calidad del diseño. Basándose en la comparación de secuencias (p. ej., tasas de error de síntesis y/o rendimientos), se puede recoger el polinucleótido monocatenario sintetizado (bloque 272).

La Fig. 9 ilustra un diagrama de flujo de una técnica 280 para generar una molécula de polinucleótido sintética ensamblada con retroalimentación de secuenciación según aspectos de la presente descripción. La técnica 280 puede comenzar proporcionando cebadores inmovilizados sobre un sustrato (bloque 282). Basándose en una secuencia de ácido nucleico diana, se puede diseñar una pluralidad de fragmentos de ácido nucleico. Cada uno de los fragmentos de ácido nucleico diseñados se puede sintetizar, amplificar, secuenciar y/o recoger de acuerdo con la técnica 260 descrita en la Fig. 8.

Basándose en las secuencias de los fragmentos de ácido nucleico diseñados, se puede sintetizar una pluralidad de polinucleótidos de fragmentos por extensión de los cebadores sobre el sustrato (bloque 284). En algunos aspectos de la descripción, la pluralidad de polinucleótidos de fragmentos puede amplificarse (bloque 286) usando cualquier método adecuado como se ha descrito antes. En otros aspectos de la descripción, dicha amplificación es opcional y, en consecuencia, el bloque 286 se ilustra con un cuadro discontinuo.

A continuación, se puede secuenciar la pluralidad de polinucleótidos de fragmentos (bloque 288). La secuencia de cada uno de la pluralidad de polinucleótidos de fragmentos puede compararse con la secuencia de su respectivo fragmento de ácido nucleico diseñado (bloque 290) para el control de calidad de la síntesis. Como se ha descrito antes, las tasas de error de síntesis (o amplificación) y/o los rendimientos pueden usarse para el control de calidad. Por ejemplo, las tasas de error de síntesis (o amplificación) y/o los rendimientos pueden usarse para proporcionar retroalimentación al diseño de fragmentos para determinar la calidad del diseño.

Basándose en la comparación de secuencias (p. ej., tasas de error de síntesis (o amplificación) y/o rendimientos), después se puede recoger la pluralidad de polinucleótidos de fragmentos (bloque 292). La pluralidad de polinucleótidos de fragmentos recogida se puede ensamblar y/o reparar adicionalmente para generar un polinucleótido ensamblado

(bloque 294). Además, como se ha descrito antes, el control de calidad (p. ej., secuenciación) puede llevarse a cabo en el polinucleótido ensamblado. El control de calidad puede proporcionar retroalimentación al diseño de fragmentos y puede usarse para recoger el polinucleótido ensamblado.

La Fig. 10 ilustra un diagrama de flujo de una técnica de retroalimentación de calidad 300 para el diseño de polinucleótidos sintéticos basada en la retroalimentación de secuenciación del polinucleótido sintético ensamblado de acuerdo con aspectos de la presente descripción. La técnica 300 puede comenzar con el sistema (p. ej., el sistema 50) que recibe una pluralidad de secuencias de fragmentos que solapan basadas en una secuencia diana (bloque 302). Las secuencias de fragmentos diseñadas pueden ser diseñadas por un operador o pueden ser diseñadas automáticamente por el sistema de acuerdo con una técnica basada en reglas que puede actualizarse de acuerdo con la información de calidad recibida como se proporciona en el presente documento. Por ejemplo, una secuencia diana puede incluir una secuencia de nucleótidos que está asociada con una puntuación de fragmentación pobre. Es decir, los datos de secuenciación previos para otras secuencias diana pueden indicar que un fragmento diseñado para romperse dentro de esa secuencia puede estar asociado con una molécula ensamblada con una alta tasa de error (o un bajo rendimiento de fragmentos). Además, se pueden asociar otras secuencias de nucleótidos con puntuaciones de fragmentación elevadas. En tales aspectos de la descripción, un sistema automático basado en reglas puede preferir crear fragmentos que se rompan dentro de las regiones de puntuación de fragmentación más alta en relación con las regiones de puntuación de fragmentación más baja de la secuencia diana.

Basándose en las secuencias de fragmentos diseñadas, se puede sintetizar una pluralidad de polinucleótidos de fragmentos (bloque 303). Cada uno de la pluralidad de polinucleótidos de fragmentos puede sintetizarse con cualquier método adecuado, incluidos los métodos descritos en la Fig. 8. El primer y último fragmento de polinucleótidos pueden modificarse adicionalmente con adaptadores (bloque 304). Por ejemplo, se puede ligar un adaptador al extremo 5' del primer polinucleótido y se puede ligar otro adaptador al extremo 3' del último polinucleótido.

Los cebadores que son complementarios de los adaptadores pueden inmovilizarse sobre un sustrato (bloque 306). Así pues, los polinucleótidos del primer fragmento pueden ser capturados por los cebadores sobre el sustrato (bloque 308). Se pueden hacer fluir otros polinucleótidos de fragmentos sobre el sustrato de manera que los otros polinucleótidos de fragmentos se ensamblen junto con los primeros fragmentos de polinucleótidos, y los polinucleótidos ensamblados se reparen adicionalmente (bloque 310).

En algunos aspectos de la descripción, los polinucleótidos ensamblados pueden amplificarse (bloque 312) usando cualquier método adecuado como se ha descrito antes. En otros aspectos de la descripción, dicha amplificación es opcional y, en consecuencia, el bloque 312 se ilustra con un cuadro discontinuo.

A continuación, se pueden secuenciar los polinucleótidos ensamblados (bloque 314). La secuencia de los polinucleótidos ensamblados se puede comparar con la secuencia diana (bloque 316) para el control de calidad de la síntesis. Como se ha descrito antes, las tasas de error de síntesis y/o los rendimientos pueden usarse para el control de calidad. Por ejemplo, pueden usarse las tasas de error de síntesis y/o rendimientos para proporcionar retroalimentación al diseño de fragmentos (proporcionado en el bloque 302) para determinar la calidad del diseño del fragmento y/o la calidad general de la molécula de ácido nucleico sintética. Basándose en la comparación de secuencias (p. ej., tasas de error de síntesis (o amplificación) y/o rendimientos), después se pueden recoger los polinucleótidos ensamblados (bloque 318).

REIVINDICACIONES

1. Un método para sintetizar un ácido nucleico que comprende una secuencia diana, comprendiendo el método:

(i) proporcionar una pluralidad de fragmentos de ácido nucleico que solapan, en donde la pluralidad de fragmentos de ácido nucleico tiene secuencias complementarias de al menos otro fragmento de la pluralidad de fragmentos de ácido nucleico, en donde un primer fragmento de la pluralidad de fragmentos de ácido nucleico comprende un primer adaptador escindible y un extremo 5' de la secuencia diana, en donde un extremo del primer fragmento comprende el primer adaptador escindible, y en donde un último fragmento de la pluralidad de fragmentos de ácido nucleico comprende un extremo 3' de la secuencia diana, en donde un extremo del último fragmento comprende un segundo adaptador escindible, en donde proporcionar una pluralidad de fragmentos de ácido nucleico que solapan comprende:

(a) generar una pluralidad del primer fragmento, que comprende:

proporcionar una pluralidad de oligonucleótidos inmovilizados sobre un sustrato de síntesis mediante hibridación con una pluralidad de primeros cebadores inmovilizados complementarios de los extremos 5' de los oligonucleótidos, en donde cada oligonucleótido respectivo comprende el primer adaptador escindible en un extremo 5',

extender los oligonucleótidos para generar polinucleótidos extendidos, en donde la extensión comprende incorporar una pluralidad de nucleótidos o ácidos nucleicos individuales en los extremos 3' de los oligonucleótidos de la pluralidad de oligonucleótidos,

unir un adaptador de amplificador en los extremos 3' de los polinucleótidos extendidos que tienen la primera secuencia terminal,

reasociar los extremos 3' de los polinucleótidos extendidos con el sustrato de síntesis mediante una pluralidad de segundos cebadores inmovilizados complementarios de los extremos 3' de los polinucleótidos extendidos para formar puentes,

amplificar los puentes para formar una pluralidad de grupos amplificados de los puentes,

secuenciar la pluralidad de grupos amplificados de los puentes para determinar la presencia de la primera secuencia terminal en uno o más de la pluralidad de grupos amplificados de los puentes y una tasa de error de síntesis, y

recoger selectivamente el uno o más de la pluralidad de grupos amplificados de los puentes basándose en la tasa de error de síntesis; y

(b) combinar los grupos amplificados recogidos que tienen la primera secuencia terminal con una segunda pluralidad de grupos amplificados, en donde al menos una parte de la segunda pluralidad de grupos amplificados comprende una secuencia complementaria de solo una parte de la primera secuencia terminal, obteniendo así una pluralidad de fragmentos de ácido nucleico;

(ii) inmovilizar el primer fragmento sobre un sustrato con un primer cebador inmovilizado complementario del primer adaptador escindible;

(iii) ensamblar la pluralidad de fragmentos de ácido nucleico en una molécula de polinucleótido ensamblada mediante hibridación de las secuencias complementarias, en donde la molécula de polinucleótido ensamblada es parcialmente bicatenaria y comprende huecos en ambas cadenas, en donde el polinucleótido ensamblado se inmoviliza sobre el sustrato por el primer cebador inmovilizado complementario del primer adaptador escindible;

(iv) amplificar la molécula de polinucleótido ensamblada sobre el sustrato para generar un grupo amplificado que comprende amplicones de la molécula de polinucleótido ensamblada sobre el sustrato;

(v) secuenciar el grupo amplificado para determinar una tasa de error de síntesis; y

(vi) recoger selectivamente el grupo amplificado basándose en la tasa de error de síntesis.

2. El método de la reivindicación 1, en donde proporcionar la pluralidad de fragmentos de ácido nucleico comprende sintetizar *in vitro* los fragmentos de ácido nucleico; y, opcionalmente, en donde la pluralidad de fragmentos de ácido nucleico comprende partes conocidas de la secuencia diana.

3. El método de la reivindicación 1, que comprende además llenar los huecos en una cadena de la molécula de polinucleótido ensamblada por extensión de polimerasa.

4. El método de la reivindicación 1, en donde la amplificación comprende la amplificación de puente de la molécula de polinucleótido ensamblada sobre el sustrato.

5. El método de la reivindicación 1, en donde la pluralidad de fragmentos de ácido nucleico son ácidos nucleicos monocatenarios; y opcionalmente, en donde cada uno de la pluralidad de fragmentos de ácido nucleico tiene menos de 200 nucleótidos de longitud; y opcionalmente, en donde las secuencias que solapan de la pluralidad de fragmentos de ácido nucleico tienen entre 20 y 30 nucleótidos de longitud.

6. El método de la reivindicación 1, en donde una tasa de error de síntesis menor que un umbral predeterminado es indicativa del grupo amplificado que comprende la secuencia diana o un complemento de la misma; y opcionalmente, en donde la tasa de error de síntesis es cero.

7. El método de la reivindicación 1, en donde recoger el grupo amplificado comprende escindir el segundo adaptador escindible y/o el primer adaptador escindible; y, opcionalmente, en donde recoger el grupo amplificado comprende escindir el segundo adaptador escindible y/o el primer adaptador escindible óptica, química, magnética, eléctrica, electromagnéticamente o cualquier combinación de los mismos.

8. El método de la reivindicación 1, en donde el ensamblaje comprende ensamblar la pluralidad de fragmentos de ácido nucleico en presencia de ADN recombinasa, proteína beta, ADN polimerasa, ADN ligasa, Circligase o cualquier combinación de los mismos.

9. El método de la reivindicación 1, en donde la extensión se lleva a cabo o bien;

mediante una primera polimerasa que incorpora la pluralidad de nucleótidos individuales en los extremos 3' de los respectivos oligonucleótidos de la pluralidad de oligonucleótidos para generar los polinucleótidos extendidos, y opcionalmente en donde la primera polimerasa comprende desoxinucleotidil transferasa terminal (TdT); o

mediante una ligasa que incorpora la pluralidad de ácidos nucleicos individuales en el extremo 3' de los oligonucleótidos de la pluralidad de oligonucleótidos para generar los polinucleótidos extendidos que tienen la primera secuencia terminal y, opcionalmente, en donde la ligasa comprende Circunligase.

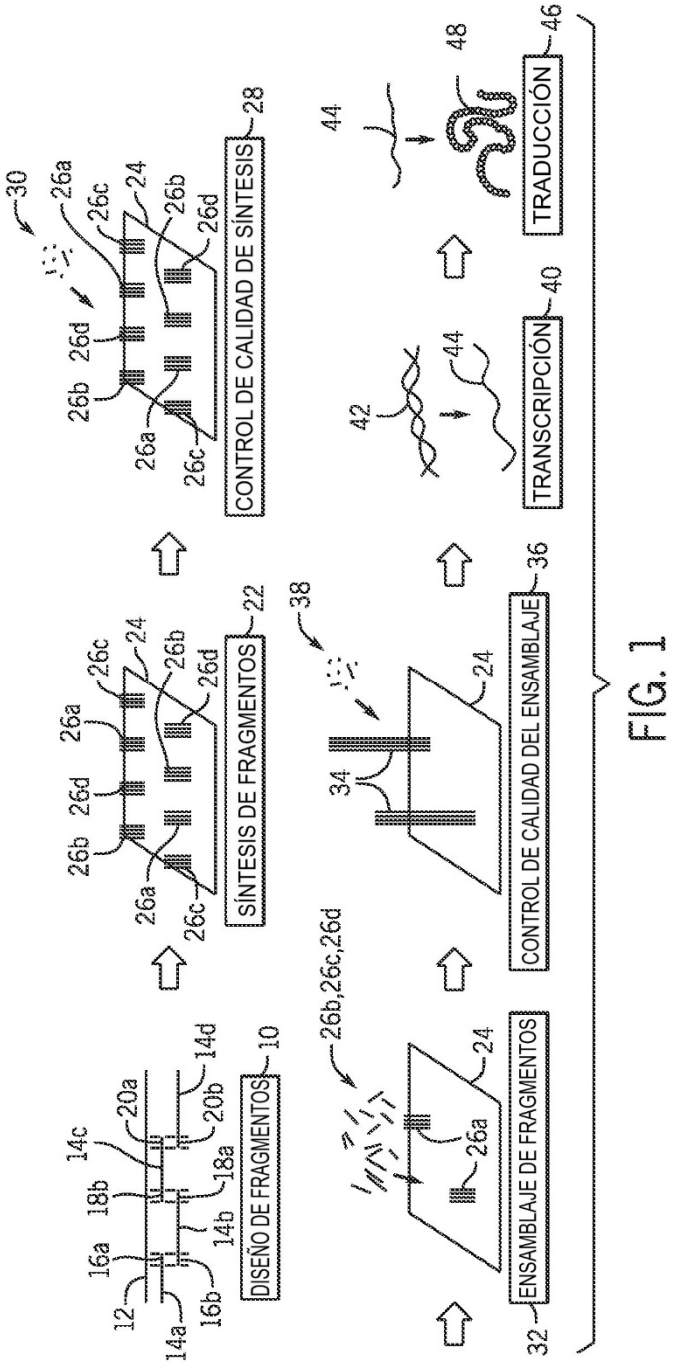
DISEÑO DE FRAGMENTOS
10
SÍNTESIS DE FRAGMENTOS
22
CONTROL DE CALIDAD DE SÍNTESIS
28
ENSAMBLAJE DE FRAGMENTOS
32
CONTROL DE CALIDAD DEL ENSAMBLAJE
36
TRANSCRIPCIÓN
40
TRADUCCIÓN
46
12
14a
14b
14c
14d
16a
16b
18a
18b
20a
20b
24
26a
26b
26c
26d
30
34
38
42
44
48

FIG. 1

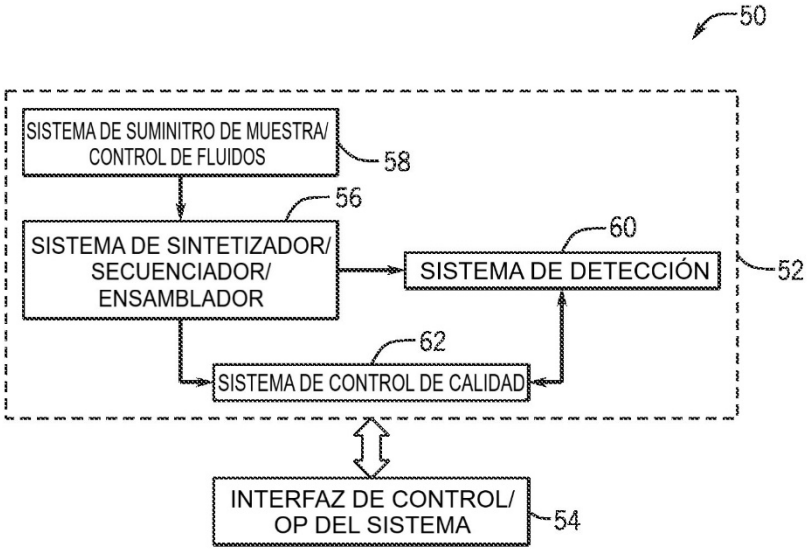
50
SISTEMA DE SUMINITRO DE MUESTRA/ CONTROL DE FLUIDOS
58
56
SISTEMA DE SINTETIZADOR/ SECUENCIADOR/ ENSAMBLADOR
60
SISTEMA DE DETECCIÓN
52
62
SISTEMA DE CONTROL DE CALIDAD
INTERFAZ DE CONTROL/ OP DEL SISTEMA
54

FIG. 2

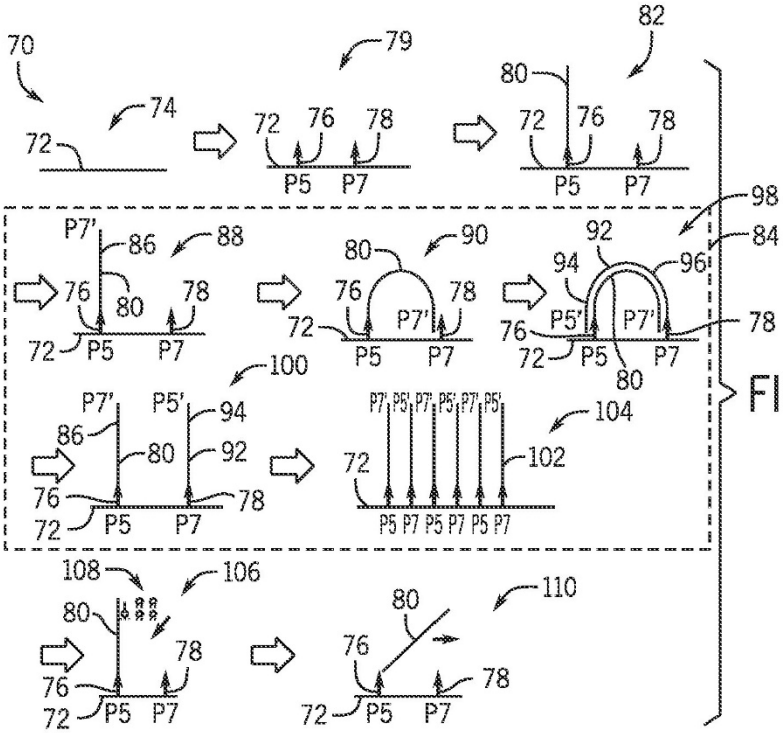
70
74
72
79
72 76 78
P5 P7
82
80
72 76 78
P5 P7
98
84
P7'
86
88
76 80 78
72 P5 P7
80
90
76 78
72 P7'
P5 P7
92
94
96
76 P5' P7' 78
72 P5 80 P7
100
P7' P5' 94
86 80 92
76 78
72 P5 P7
104
P7' P5' P7' P5' P7' P5'
72 102
P5 P7 P5 P7 P5 P7
108
106
80
76 78
72 P5 P7
80
110
76 78
72 P5 P7

FIG. 3

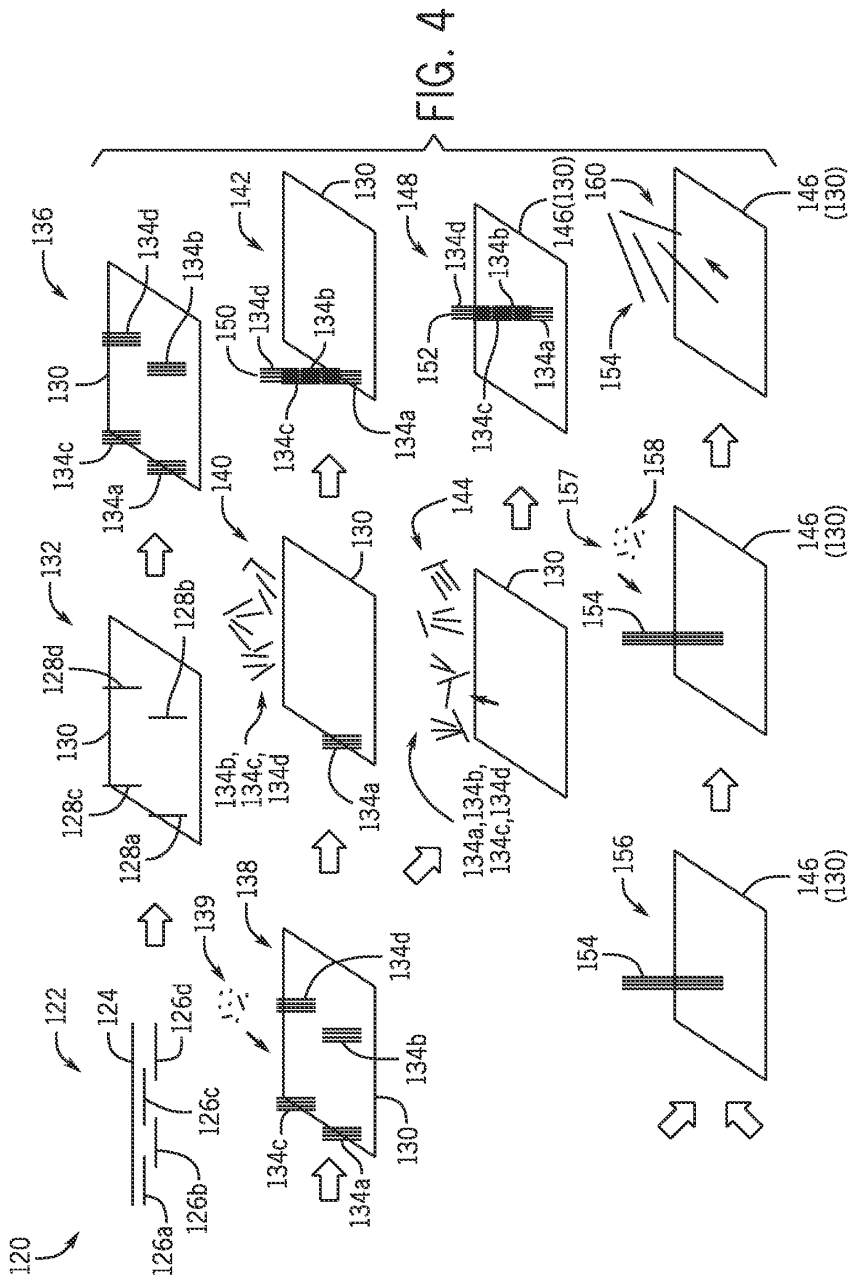
120
122
124
126a
126b
126c
126d
128a
128b
128c
128d
130
132
134a
134b
134c
134d
134a,134b,134c,134d
134b,134c,134d
136
138
139
140
142
144
146(130)
148
150
152
154
156
157
158
160

FIG. 4

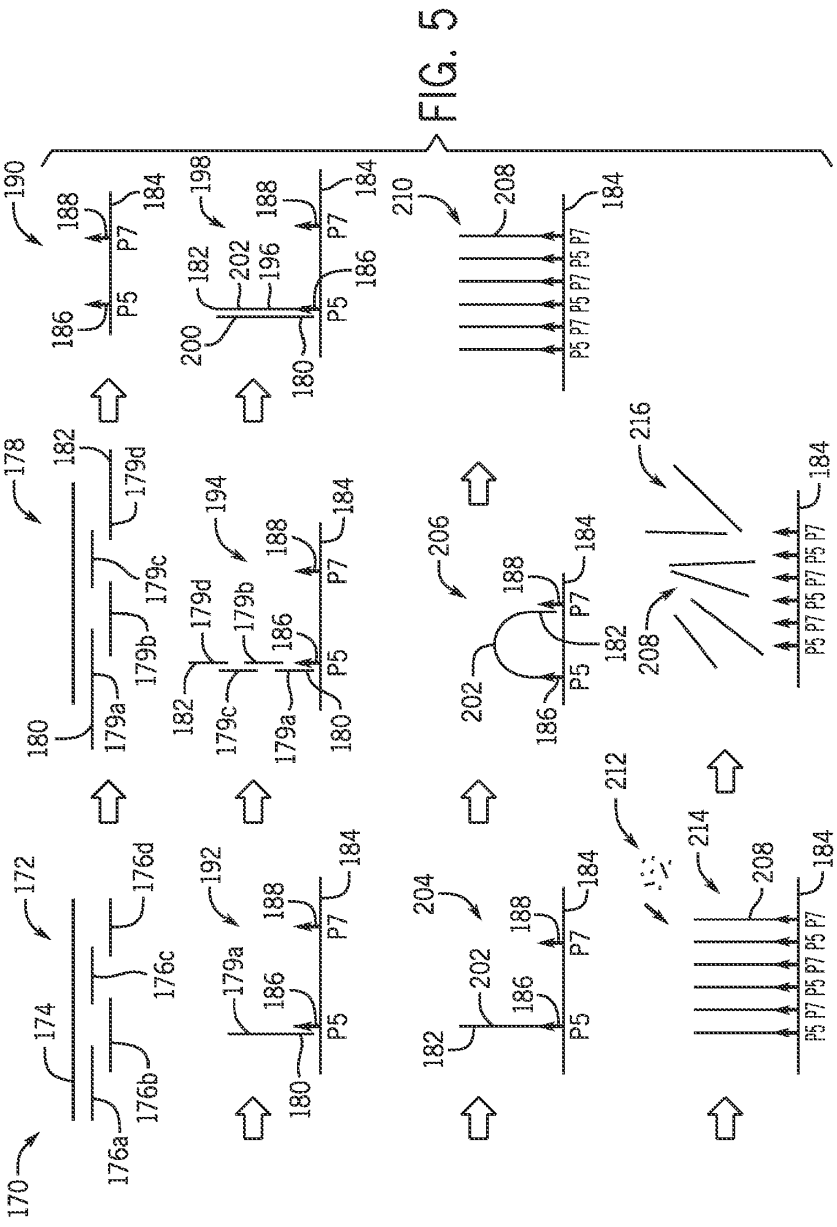
170
172
174
176a
176b
176c
176d
178
180
182
179a
179b
179c
179d
184
186
188
190
192
194
196
198
200
202
204
206
208
210
212
214
216
P5
P7
P5 P7 P5 P7

FIG. 5

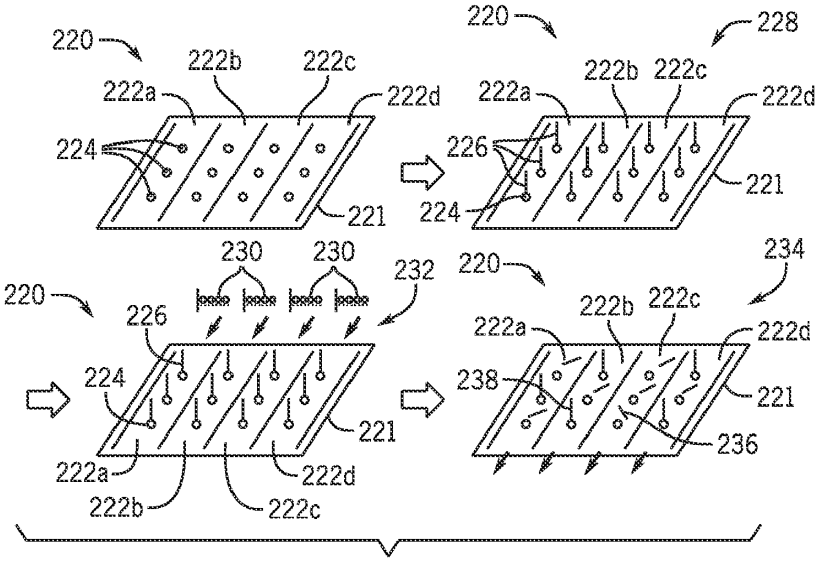
220
222b
222c
222a
222d
224
221
220
228
222b
222c
222a
222d
226
221
224
230
230
220
232
226
224
221
222a
222d
222b
222c
220
234
222b
222c
222a
222d
238
221
236

FIG. 6

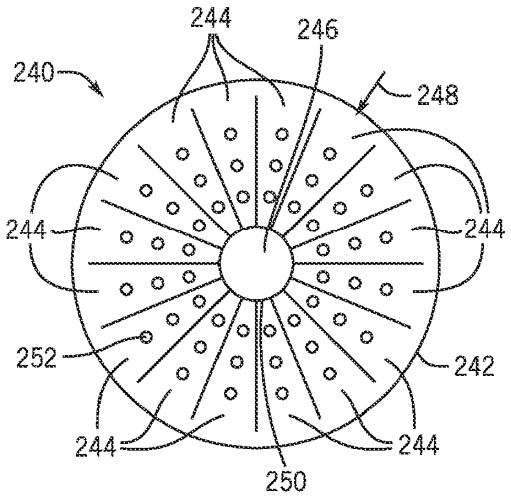
244
246
240
248
244
244
252
242
244
244
250

FIG. 7

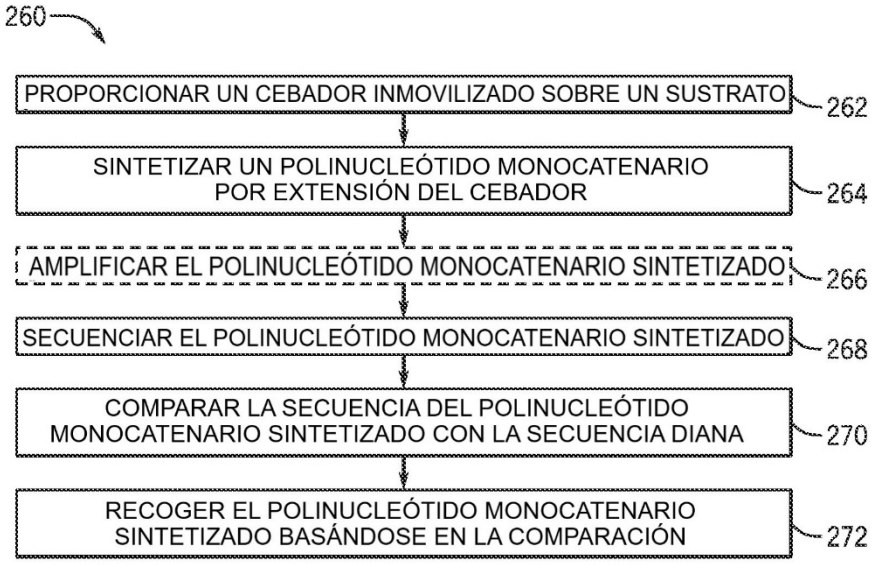
260
PROPORCIONAR UN CEBADOR INMOVILIZADO SOBRE UN SUSTRATO
262
SINTETIZAR UN POLINUCLEÓTIDO MONOCATENARIO POR EXTENSIÓN DEL CEBADOR
264
AMPLIFICAR EL POLINUCLEÓTIDO MONOCATENARIO SINTETIZADO
266
SECUENCIAR EL POLINUCLEÓTIDO MONOCATENARIO SINTETIZADO
268
COMPARAR LA SECUENCIA DEL POLINUCLEÓTIDO MONOCATENARIO SINTETIZADO CON LA SECUENCIA DIANA
270
RECOGER EL POLINUCLEÓTIDO MONOCATENARIO SINTETIZADO BASÁNDOSE EN LA COMPARACIÓN
272

FIG. 8

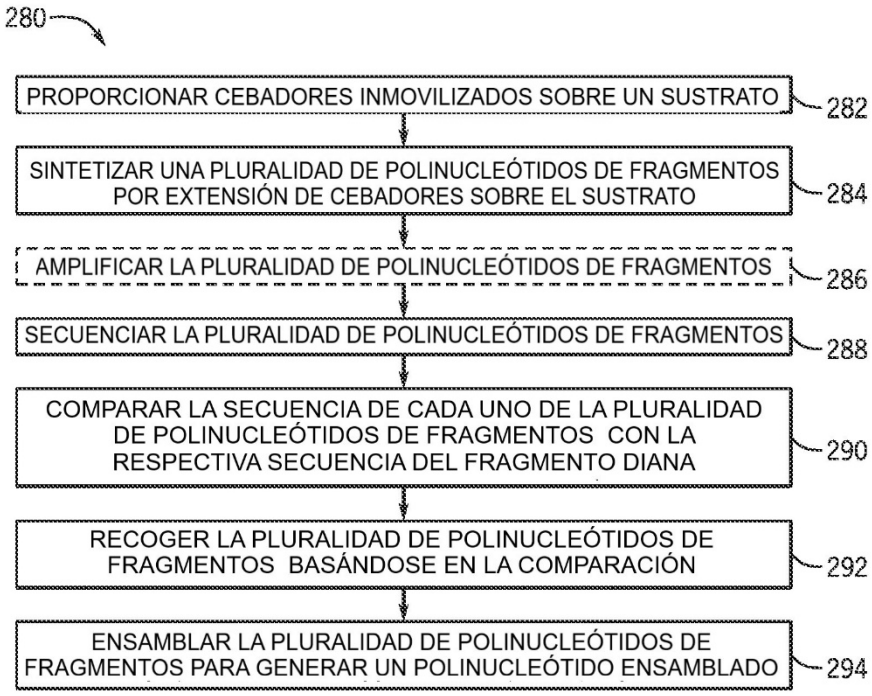
280
PROPORCIONAR CEBADORES INMOVILIZADOS SOBRE UN SUSTRATO
282
SINTETIZAR UNA PLURALIDAD DE POLINUCLEÓTIDOS DE FRAGMENTOS POR EXTENSIÓN DE CEBADORES SOBRE EL SUSTRATO
284
AMPLIFICAR LA PLURALIDAD DE POLINUCLEÓTIDOS DE FRAGMENTOS
286
SECUENCIAR LA PLURALIDAD DE POLINUCLEÓTIDOS DE FRAGMENTOS
288
COMPARAR LA SECUENCIA DE CADA UNO DE LA PLURALIDAD DE POLINUCLEÓTIDOS DE FRAGMENTOS CON LA RESPECTIVA SECUENCIA DEL FRAGMENTO DIANA
290
RECOGER LA PLURALIDAD DE POLINUCLEÓTIDOS DE FRAGMENTOS BASÁNDOSE EN LA COMPARACIÓN
292
ENSAMBLAR LA PLURALIDAD DE POLINUCLEÓTIDOS DE FRAGMENTOS PARA GENERAR UN POLINUCLEÓTIDO ENSAMBLADO
294

FIG. 9

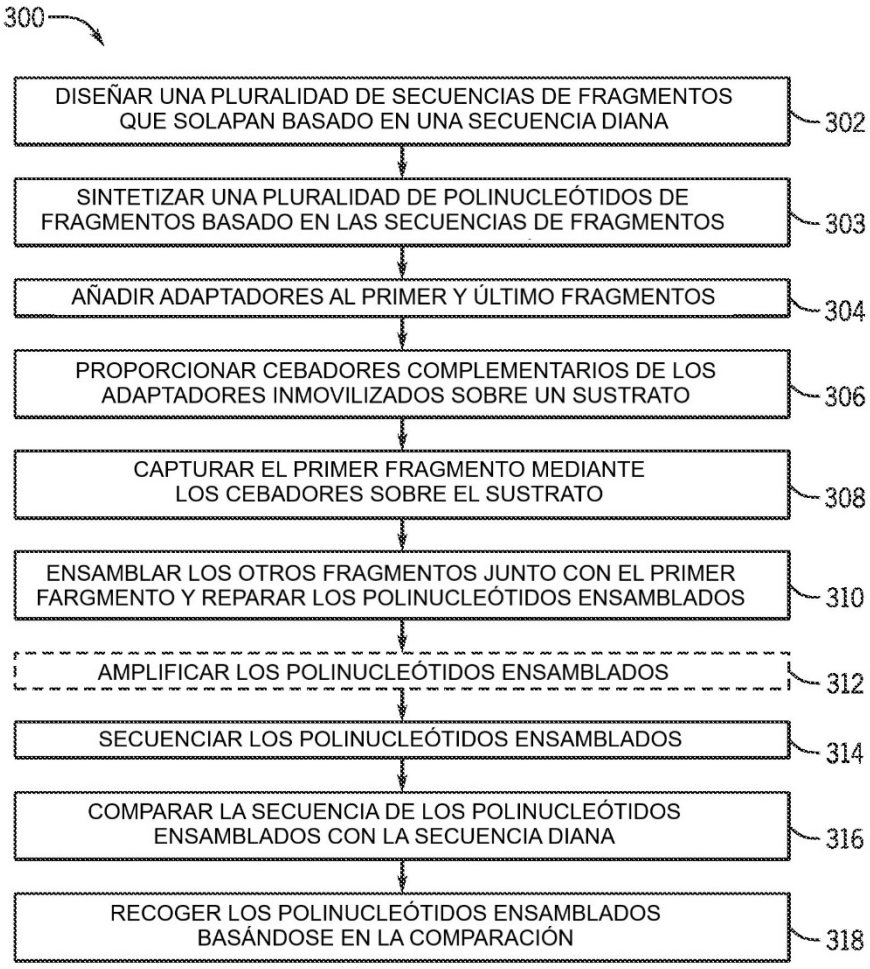
300
DISEÑAR UNA PLURALIDAD DE SECUENCIAS DE FRAGMENTOS QUE SOLAPAN BASADO EN UNA SECUENCIA DIANA
302
SINTETIZAR UNA PLURALIDAD DE POLINUCLEÓTIDOS DE FRAGMENTOS BASADO EN LAS SECUENCIAS DE FRAGMENTOS
303
AÑADIR ADAPTADORES AL PRIMER Y ÚLTIMO FRAGMENTOS
304
PROPORCIONAR CEBADORES COMPLEMENTARIOS DE LOS ADAPTADORES INMOVILIZADOS SOBRE UN SUSTRATO
306
CAPTURAR EL PRIMER FRAGMENTO MEDIANTE LOS CEBADORES SOBRE EL SUSTRATO
308
ENSAMBLAR LOS OTROS FRAGMENTOS JUNTO CON EL PRIMER FARGMENTO Y REPARAR LOS POLINUCLEÓTIDOS ENSAMBLADOS
310
AMPLIFICAR LOS POLINUCLEÓTIDOS ENSAMBLADOS
312
SECUENCIAR LOS POLINUCLEÓTIDOS ENSAMBLADOS
314
COMPARAR LA SECUENCIA DE LOS POLINUCLEÓTIDOS ENSAMBLADOS CON LA SECUENCIA DIANA
316
RECOGER LOS POLINUCLEÓTIDOS ENSAMBLADOS BASÁNDOSE EN LA COMPARACIÓN
318

FIG. 10