

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 977 793**

51 Int. Cl.:

C12Q 1/6806 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **08.11.2018** **PCT/US2018/059908**

87 Fecha y número de publicación internacional: **16.05.2019** **WO19094651**

96 Fecha de presentación y número de la solicitud europea: **08.11.2018** **E 18877065 (5)**

97 Fecha y número de publicación de la concesión europea: **10.04.2024** **EP 3707273**

54 Título: **Reactivos y adaptadores para la secuenciación de ácidos nucleicos y métodos para fabricar dichos reactivos y adaptadores**

30 Prioridad:

08.11.2017 US 201762583462 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
30.08.2024

73 Titular/es:

TWINSTRAND BIOSCIENCES, INC. (100.0%)
3131 Elliot Ave., Suite 750
Seattle, WA 98121, US

72 Inventor/es:

SALK, JESSE J. y
KENNEDY, SCOTT R.

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 977 793 T3

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

DESCRIPCIÓN

Reactivos y adaptadores para la secuenciación de ácidos nucleicos y métodos para fabricar dichos reactivos y adaptadores

Antecedentes

5 La Secuenciación Dúplex es un método para producir lecturas de secuencias de ácidos nucleicos con corrección de errores a partir de moléculas de ácido nucleico de doble hebra. En determinados aspectos de la tecnología, la Secuenciación Dúplex se puede utilizar para secuenciar independientemente ambas hebras de moléculas de ácido nucleico individuales de tal manera que se puede reconocer que las lecturas de la secuencia derivada se han originado a partir de la misma molécula original de ácido nucleico de doble hebra durante secuenciación paralela masiva, pero también se diferencian entre sí como entidades distinguibles después de la secuenciación. Las lecturas de secuencia resultantes de cada hebra luego se comparan con el fin de obtener una secuencia con corrección de errores de la molécula de ácido nucleico de doble hebra original, conocida como una Secuencia de Consenso Dúplex. El proceso de Secuenciación Dúplex permite confirmar si una o ambas hebras de una molécula de ácido nucleico de doble hebra original están representadas en los datos de secuenciación generados y utilizados para formar una Secuencia de Consenso Dúplex.

El alto grado de corrección de errores proporcionado por la tecnología de comparación de hebras de Secuenciación Dúplex reduce los errores de secuenciación de moléculas de ácido nucleico de doble hebra en múltiples órdenes de magnitud en comparación con los métodos de secuenciación estándar de próxima generación. La reducción en los errores de secuenciación mejora la precisión de las lecturas de secuenciación en casi todos los tipos de moléculas de ácido nucleico de doble hebra, sin embargo la utilidad de esto es mayor cuando la población molecular que se secuencia es heterogénea (es decir un subconjunto menor de moléculas porta una variante de secuencia que otros no). La tasa de error de la secuenciación estándar de próxima generación es del orden aproximado de 1/100-1/1000 y cuando menos de 1/100-1/1000 de las moléculas portan una variante de secuencia, su presencia queda oscurecida por la tasa de error de fondo del proceso de secuenciación. La Secuenciación Dúplex, por otro lado puede detectar con precisión variantes de frecuencia extremadamente bajas debido al alto grado de corrección de errores obtenidos.

Además de las aplicaciones de detección de variantes de baja frecuencia, la Secuenciación Dúplex también es adecuada para el genotipado preciso de regiones del genoma difíciles de secuenciar (homopolímeros, microsatélites, G-tetraplex etc.) en donde la tasa de error de la secuenciación estándar es especialmente alta. En el ADN muy dañado (oxidación, desaminación, etc.), que se produce a través de procesos de fijación (es decir FFPE en patología clínica) o ADN antiguo o en aplicaciones forenses en donde el material se ha expuesto a productos químicos o entornos agresivos, la Secuenciación Dúplex es particularmente útil para reducir el alto nivel de error resultante que confiere el daño. Se pueden encontrar ejemplos adicionales no limitantes de la utilidad de la Secuenciación Dúplex en Salk et al, Nature Reviews Genetics 2018, PMID 29576615.

Los métodos que incorporan Secuenciación Dúplex pueden incluir la ligación de uno o más adaptadores de secuenciación a una molécula de ácido nucleico de doble hebra diana, que comprende una primera secuencia de ácidos nucleicos diana de hebra y una segunda secuencia de ácidos nucleicos diana de hebra, para producir un complejo de ácido nucleico diana de doble hebra. El uso de adaptadores de secuenciación de alta pureza para Secuenciación Dúplex, o cualquier tecnología de secuenciación de próxima generación, es importante para obtener datos reproducibles de alta calidad y maximizar el rendimiento de secuencia de una muestra (es decir, el porcentaje relativo de moléculas ingresadas que se convierten en lecturas de secuencia independientes). La Solicitud de Patente China No. 106 367 485 A describe un conjunto de detectores de etiqueta dobles de múltiples posiciones para detectar mutaciones genéticas. Es particularmente importante con la Secuenciación Dúplex debido a la necesidad de recuperar con éxito ambas hebras de las moléculas dúplex originales. Los métodos básicos actuales de síntesis de adaptadores pueden conducir a la acumulación de intermedios y subproductos incompletos entre el producto adaptador deseado creando ineficiencias en la preparación de la biblioteca de ácidos nucleicos y pérdida de información de secuenciación. Esto es particularmente cierto en el caso de determinadas formas de adaptadores de secuenciación Dúplex que se fabrican de tal manera que implican una extensión de polimerasa y otras etapas enzimáticas o químicas que pueden no ser 100 % eficientes. De acuerdo con lo anterior, subsiste la necesidad no satisfecha de métodos y reactivos potenciados para sintetizar adaptadores de secuenciación purificados, y especialmente para uso con Secuenciación Dúplex.

Resumen

En el presente documento se describen una variedad de métodos y reactivos para producir adaptadores de secuenciación para uso en aplicaciones de secuenciación de ácidos nucleicos, tales como Secuenciación Dúplex. Estas realizaciones específicas están dirigidas a mejorar la pureza del adaptador, eficiencia de fabricación, reproducibilidad, coste y flexibilidad del diseño.

La invención se define por las reivindicaciones que proporcionan un método para preparar un adaptador dúplex, que comprende: aparear una hebra de alargamiento y una hebra de plantilla en una región complementaria para formar un primer producto adaptador dúplex intermedio, en el que la hebra de plantilla comprende: una secuencia

identificadora, un primer marcador de captura adherido a la hebra de plantilla a través de un primer sitio de corte, y un segundo marcador de captura unido a la hebra de plantilla a través de un segundo sitio de corte; extender la hebra de alargamiento para duplicar al menos parcialmente la secuencia identificadora para formar un segundo producto adaptador dúplex intermedio que comprende un primer sitio de corte entre la secuencia identificadora y el primer marcador de captura; cortar el segundo producto adaptador dúplex intermedio para formar un tercer producto adaptador dúplex intermedio y un primer subproducto escindido que comprende el primer marcador de captura; eliminar productos no deseados que comprenden el primer marcador de captura; y cortar en el segundo sitio de corte para formar el adaptador dúplex y un segundo subproducto escindido que comprende el segundo marcador de captura, y eliminar productos no deseados.

En algunas realizaciones, el marcador de captura está en un extremo 5' de la hebra de plantilla de tal manera que las hebras de plantilla no apareadas, el primer y segundo productos adaptadores dúplex intermedios y los subproductos comprenden el marcador de captura. En algunas realizaciones, los productos no deseados incluyen hebras de plantilla no apareadas, primer y segundo productos adaptadores dúplex intermedios y subproductos. En algunas realizaciones, los métodos proporcionados comprenden además proporcionar una superficie que comprende una fracción de extracción configurada para unirse al marcador de captura. En algunas realizaciones, las etapas de apareamiento, extensión y/o corte ocurren en una solución líquida, y en el que la solución líquida se expone a la superficie después de la etapa de corte. En algunas realizaciones, la hebra de plantilla se une a la superficie antes de la etapa de apareamiento. En algunas realizaciones, la hebra de plantilla se une a la superficie mediante el marcador de captura durante las etapas de apareamiento y extensión. En algunas realizaciones, los métodos proporcionados comprenden además recolectar un producto de adaptador dúplex purificado. En algunas realizaciones, el marcador de captura es un primer marcador de captura en un extremo 5' de la hebra de plantilla, y en el que la hebra de plantilla comprende un segundo marcador de captura en un extremo 3', y el método incluye además las etapas de proporcionar una segunda superficie que comprende una segunda fracción de extracción, y cortar una región 3' de la hebra de plantilla para liberar el adaptador dúplex de la segunda superficie. En algunas realizaciones, el primer y segundo marcadores de captura son diferentes. En algunas realizaciones, el primer y segundo marcadores de captura son los mismos. En algunas realizaciones, la hebra de plantilla y la hebra de alargamiento están ligadas mediante un dominio ligador. En algunas realizaciones, el dominio ligador comprende nucleótidos. En algunas realizaciones, el dominio ligador forma un bucle que comprende nucleótidos de hebra sencilla. En algunas realizaciones, el dominio ligador contiene una o más moléculas de nucleótidos o no nucleótidos modificadas. En algunas realizaciones, el método comprende cortar el dominio ligador en cualquier etapa después de la etapa de apareamiento para proporcionar dos brazos de hebra sencilla en el extremo 5' del producto adaptador, en el que el dominio ligador contiene una o más moléculas de nucleótidos o no nucleótidos modificadas. En algunas realizaciones, extender la hebra de alargamiento comprende extender la hebra de alargamiento desde un extremo 3' de la hebra de alargamiento mediante reacción enzimática. En algunas realizaciones, el primer y segundo sitio de corte comprenden al menos uno de una molécula de nucleótido o no nucleótido modificada y un sitio de reconocimiento de endonucleasa de restricción. En algunas realizaciones, la hebra de plantilla comprende una estructura de bucle en horquilla que tiene el marcador de captura.

La presente divulgación proporciona métodos para preparar un adaptador dúplex que incluye las etapas de aparear una hebra de alargamiento y una hebra de plantilla en una región complementaria, en los que la hebra de plantilla comprende una secuencia identificadora, un primer marcador de captura, un primer sitio de corte, un segundo marcador de captura, y un segundo sitio de corte para formar un primer producto adaptador dúplex intermedio, en el que el primer marcador de captura se adhiere a la hebra de plantilla a través del primer sitio de corte, extender la hebra de alargamiento para duplicar al menos parcialmente la secuencia identificadora para formar un segundo producto adaptador dúplex intermedio, cortar el segundo producto adaptador dúplex intermedio en el segundo sitio de corte para formar el adaptador dúplex y un subproducto escindido que comprende el segundo marcador de captura, y eliminar productos no deseados, cortar el primer sitio de corte para liberar el adaptador dúplex, y retirar productos adicionales no deseados.

En algunas realizaciones, los métodos proporcionados comprenden además proporcionar al menos una fracción de extracción configurada para unirse al primer marcador de captura, y capturar productos no deseados que incluyen el primer marcador de captura. En algunas realizaciones, los métodos proporcionados comprenden además proporcionar al menos una fracción de extracción configurada para unirse al segundo marcador de captura, y capturar productos no deseados que incluyen el segundo marcador de captura. En algunas realizaciones, la fracción de extracción está unida a una superficie. En algunas realizaciones, los productos no deseados comprenden uno o más de un exceso de hebra de plantilla, exceso de hebra de alargamiento, complejos preadaptadores no extendidos o extendidos de manera incompleta, y subproductos de fragmentos de escisión.

La presente divulgación proporciona métodos para preparar un conjunto de adaptadores de secuenciación dúplex que tienen una secuencia identificadora de doble hebra que incluye las etapas de proporcionar soportes sólidos para la síntesis de oligonucleótidos que comprenden una pluralidad de hebras de plantilla unidas a los mismos, en los que las hebras de plantilla comprenden una secuencia de nucleótidos que se extiende desde el soporte sólido de síntesis de oligonucleótidos a un extremo terminal 5', y en el que una porción de la secuencia de nucleótidos incluye una secuencia identificadora que distingue cada hebra de plantilla de las otras hebras de plantilla, aparear una hebra de alargamiento con cada hebra de plantilla en una región complementaria para formar una pluralidad de adaptadores de secuenciación preliminar, extender cada hebra de alargamiento en una dirección de 5' a 3' de tal manera que la secuencia identificadora esté presente en cada hebra de cada adaptador de secuenciación preliminar, y liberar los adaptadores

de secuenciación preliminar de los soportes sólidos de síntesis de oligonucleótidos para proporcionar un conjunto de dúplex adaptadores de secuenciación que tienen una secuencia identificadora de doble hebra.

En algunas realizaciones, cada adaptador de secuenciación preliminar comprende dos brazos de hebra sencilla próximos al soporte sólido de síntesis de oligonucleótidos. En algunas realizaciones, los métodos proporcionados comprenden además cortar enzimáticamente un extremo 3' de los adaptadores de secuenciación dúplex para formar un extremo 3' ligable. En algunas realizaciones, los métodos proporcionados comprenden además sintetizar la pluralidad de hebras de plantilla en los soportes sólidos de síntesis de oligonucleótidos. En algunas realizaciones, los soportes sólidos de síntesis de oligonucleótidos son perlas de vidrio de poro controlado (CPG) o perlas de poliestireno macroporoso (MPPS).

La presente divulgación proporciona métodos para fabricar un adaptador dúplex que incluye las etapas de proporcionar una hebra de plantilla que tiene una secuencia identificadora, una primera estructura de bucle en horquilla en una región 5' y una segunda estructura de bucle en horquilla en una región 3', en los que la primera estructura de bucle en horquilla comprende un primer bucle de nucleótido de hebra sencilla que tiene un primer sitio de corte y una porción de tallo de doble hebra en 5', la segunda estructura de bucle en horquilla comprende un segundo bucle de nucleótido de hebra sencilla que tiene un marcador de captura y una porción de tallo de doble hebra en 3' que tiene un segundo sitio de corte, y la hebra de plantilla comprende además una secuencia identificadora en una región media entre la porción de tallo de doble hebra 5' y la porción de tallo de doble hebra 3', extender enzimáticamente la hebra de plantilla desde un extremo terminal 3' hasta encontrarse con el extremo terminal 5' de tal manera que la secuencia identificadora sea de doble hebra, y cortar el primer y segundo sitios de corte para proporcionar un adaptador dúplex que tiene una porción de hebra sencilla en un extremo 5' del adaptador dúplex, y que tiene un dominio de ligación en un extremo de 3' del adaptador dúplex.

La presente divulgación proporciona métodos para fabricar un adaptador dúplex, que incluye las etapas de proporcionar una hebra de plantilla que comprende, en una dirección de 5' a 3', una porción de hebra sencilla que tiene una molécula de nucleótido o no nucleótido modificada y una secuencia identificadora, una porción de tallo de doble hebra que tiene un sitio de corte, y un bucle de nucleótido de hebra sencilla que tiene un marcador de captura, en el que la porción de tallo comprende una región de secuencia complementaria entre la región 3' de la hebra de plantilla y una región media de la hebra de plantilla, formando de esta manera el bucle de nucleótido de hebra sencilla, inmovilizar la hebra de plantilla a través del marcador de captura, extender la porción de hebra sencilla que tiene la secuencia identificadora desde un extremo terminal 3', escindir la hebra de plantilla en el nucleótido modificado, eliminar una región 5' de la hebra de plantilla para generar una región de hebra sencilla 3' de la hebra de plantilla, proporcionar una segunda hebra que tiene una secuencia al menos parcialmente complementaria a la región de hebra sencilla 3' de la hebra de plantilla, aparear la segunda hebra con la región de hebra sencilla 3' de la hebra de plantilla para generar un complejo preadaptador, y cortar el sitio de corte para proporcionar un adaptador dúplex que tiene una secuencia identificadora de doble hebra.

La presente divulgación proporciona métodos para fabricar un adaptador dúplex, que incluye las etapas de proporcionar una hebra de plantilla que comprende, en una dirección de 5' a 3', una porción de hebra sencilla en una región 5', una primera estructura de bucle en horquilla, una secuencia identificadora, y una segunda estructura de bucle de horquilla, en la que la primera estructura de bucle de horquilla comprende (a) una porción de tallo de doble hebra 5' que tiene una región de secuencia complementaria entre la región 5' de la hebra de plantilla y una región media de la hebra de plantilla y formar un primer bucle de nucleótido de hebra sencilla entre la región de secuencia de complementariedad y (b) un bucle de nucleótido de hebra sencilla que tiene una molécula de nucleótido o no nucleótido modificada, la segunda estructura de bucle en horquilla comprende (a) una porción de tallo de doble hebra en 3' que tiene una región de secuencia de complementariedad entre la región 3' de la hebra de plantilla una región media de la hebra de plantilla y que forma un segundo bucle de nucleótido de hebra sencilla entre la región de secuencia de complementariedad, (b) un sitio de corte, y (c) un bucle de nucleótido de hebra sencilla que tiene un marcador de captura, que extiende enzimáticamente la hebra de plantilla desde un extremo terminal 3' sobre la porción de tallo de doble hebra 5' de la primera estructura de bucle en horquilla de tal manera que la secuencia identificadora y la porción de hebra sencilla se hagan de doble hebra, cortar la molécula de nucleótido o no nucleótido modificada para crear una mella de hebra sencilla que permite la liberación de un subproducto de hebra sencilla, liberar el subproducto de hebra sencilla y cortar el sitio de corte para proporcionar un adaptador dúplex que tiene una única porción de hebra en un extremo 5' del adaptador dúplex, y que tiene un dominio de ligación en un extremo 3' del adaptador dúplex.

La presente divulgación proporciona métodos para fabricar un adaptador de secuenciación dúplex purificado que tiene un identificador molecular único (UMI) en cada hebra, que incluye las etapas para proporcionar un adaptador de secuenciación preliminar que comprende una región hibridada de doble hebra, dos brazos de hebra sencilla, un saliente que comprende el UMI físico en un extremo de la región hibridada de doble hebra que está más alejado de los dos brazos de hebra sencilla, y un marcador de captura en el extremo 5' del saliente, que extiende una hebra de la región hibridada de doble hebra utilizando el saliente como una plantilla, produciendo de esta manera un producto de extensión, cortar el producto de extensión en una región de doble hebra 3' con respecto al UMI físico en un sitio de escisión, produciendo de esta manera un adaptador de secuenciación dúplex y un subproducto que comprende los nucleótidos 3' del sitio de escisión y el marcador de captura y separar el adaptador de secuenciación dúplex de los

productos no deseados para proporcionar un adaptador de secuenciación dúplex purificado que tiene un UMI físico en cada hebra.

En algunas realizaciones, una hebra de plantilla comprende una secuencia índice.

5 En algunas realizaciones, extender mediante reacción enzimática comprende actividad de ADN polimerasa. En algunas realizaciones, la ADN polimerasa se selecciona de Pol I, Pol II, Pol III, Pol IV, Pol V, polimerasa Taq, polimerasa alfa, polimerasa beta, polimerasa delta, polimerasa lambda, polimerasa sigma, polimerasa epsilon, polimerasa mu, polimerasa zeta, polimerasa un, y polimerasa theta.

10 En algunas realizaciones, los métodos proporcionados comprenden además ligar el adaptador dúplex a una molécula de ácido nucleico de doble hebra. En algunas realizaciones, la molécula de ácido nucleico de doble hebra es una molécula de ADN de doble hebra o una molécula de ARN de doble hebra. En algunas realizaciones, la molécula de ácido nucleico de doble hebra comprende al menos una molécula de nucleótido o no nucleótido modificada.

15 En algunas realizaciones, la ligación comprende la actividad de la al menos una ligasa. En algunas realizaciones, la al menos una ligasa se selecciona de una ADN ligasa y una ARN ligasa. En algunas realizaciones, la ligación comprende actividad ligasa en un dominio de ligación. En algunas realizaciones, el dominio de ligación es una secuencia de nucleótidos de uno o más nucleótidos degenerados o semidegenerados. En algunas realizaciones, el dominio de ligación es una secuencia de nucleótidos de uno o más nucleótidos no degenerados. En algunas realizaciones, el dominio de ligación contiene uno o más nucleótidos modificados. En algunas realizaciones, el dominio de ligación comprende un saliente en T, un saliente en A, un saliente en CG, un extremo romo, una secuencia de recombinación, un saliente de digestión de restricción, u otra región ligable. En algunas realizaciones, al menos una hebra del dominio de ligación está fosforilada. En algunas realizaciones, el dominio de ligación comprende una secuencia de escisión por endonucleasa de restricción.

20 En algunas realizaciones, la secuencia de escisión de la endonucleasa de restricción se escinde mediante una endonucleasa de restricción para producir un extremo romo, o una región ligable saliente. En algunas realizaciones, el dominio de ligación está en 3' para la secuencia identificadora. En algunas realizaciones, el dominio de ligación está en 5' para la secuencia identificadora.

25 En algunas realizaciones, una secuencia identificadora es o comprende una secuencia de identificador de molécula única (SMI). En algunas realizaciones, una secuencia SMI es una secuencia SMI endógena. En algunas realizaciones, la secuencia SMI endógena está relacionada con el punto de corte. En algunas realizaciones, la secuencia SMI comprende al menos un ácido nucleico degenerado o semidegenerado. En algunas realizaciones, la secuencia SMI no es degenerada. En algunas realizaciones, la secuencia SMI es una secuencia de nucleótidos de uno o más nucleótidos degenerados o semidegenerados. En algunas realizaciones, la secuencia SMI es una secuencia de nucleótidos de uno o más nucleótidos no degenerados. En algunas realizaciones, la secuencia SMI de la hebra de plantilla y la SMI de la hebra de alargamiento son complementarias. En algunas realizaciones, la secuencia SMI de la hebra de plantilla y la SMI de la hebra de alargamiento son al menos parcialmente no complementarias. En algunas realizaciones, la secuencia SMI de la hebra de plantilla y la SMI de la hebra de alargamiento no son complementarias. En algunas realizaciones, la secuencia SMI comprende al menos una molécula de nucleótido o no nucleótido modificada. En algunas realizaciones, la secuencia SMI comprende un dominio de unión a cebador.

30 En algunas realizaciones, una molécula de nucleótido o no nucleótido modificada se selecciona de 2-Aminopurina, 2,6-Diaminopurina (2-Amino-dA), 5-bromo dU, desoxiuridina, dT Invertida, Didesoxi-T Invertida, Didesoxi-C, 5-Metil dC, desoxiinosina, Super T[®], Súper G[®], Ácidos Nucleicos Bloqueados, 5-Nitroindol, Bases de ARN 2'-O-Metilo, Hidroximetilo dC, Iso-dG, Iso-dC, Fluoro C, Fluoro U, Fluoro A, Fluoro G, 2-MetoxiEtoxi A, 2-MetoxiEtoxi MeC, 2-MetoxiEtoxi G, 2-MetoxiEtoxi T, 8-oxo-A, 8-oxoG, 5-hidroximetil-2'-desoxicitidina, 5'-metilisocitosina, tetrahidrofurano, isocitosina, isoguanosina, uracilo, nucleótido metilado, Nucleótido de ARN, nucleótido de ribosa, 8-oxo-G, BrdU, Loto dU, furano, tinte fluorescente, nucleótido de azida, nucleótido abásico, 5-nitroindolnucleótido y nucleótido de digoxenina.

40 En algunas realizaciones, un sitio de corte es o comprende una secuencia de reconocimiento de endonucleasa de restricción.

45 En algunas realizaciones, un marcador de captura es o comprende al menos uno de grupos Acridita, azida, azida (éster NHS), digoxigenina (éster NHS), I-Ligador, modificador Amino C6, modificador Amino C12, modificador Amino C6 dT, modificador Amino no ligado, hexinilo, 5-octadiinilo dU, biotina, biotina (azida), biotina dT, biotina TEG, biotina dual, biotina PC, destiobiotina TEG, modificador de tiol C3, ditiol, modificador de tiol C6 S-S y succinilo.

50 En algunas realizaciones, una fracción de extracción es o comprende al menos uno de amino silano, epoxi silano, isotiocianato, aminofenil silano, aminopropil silano, mercapto silano, aldehído, epóxido, fosfonato, estreptavidina, avidina, un hapteno que reconoce un anticuerpo, una secuencia de ácidos nucleicos particular, partículas magnéticamente atraíbles (Dynabeads) y resinas fotolábiles.

Como se utilizan en esta solicitud, los términos "alrededor de" y "aproximadamente" se utilizan como equivalentes. Cualquier número utilizado en esta solicitud con o sin alrededor/aproximadamente pretende cubrir cualquier fluctuación normal apreciada por un experto con conocimientos básicos en la técnica pertinente.

- 5 Otras características, objetos, y ventajas de la presente invención son evidentes en la descripción detallada que sigue. Se debe entenderse, sin embargo, que la descripción detallada, aunque indica realizaciones de la presente invención, se proporciona únicamente a modo de ilustración, no de limitación.

Breve descripción de los dibujos

- 10 Muchos aspectos de la presente divulgación se pueden entender mejor con referencia a los siguientes dibujos. Los componentes de los dibujos no están necesariamente a escala. En lugar de ello, se pone énfasis en ilustrar varios principios de la presente divulgación. Todos los ejemplos a continuación se deben tomar como realizaciones representativas no limitantes de los conceptos descritos en el presente documento.

- 15 La FIG. 1 representa un método de Apareamiento de Oligonucleótidos (OA) para fabricar Adaptadores Dúplex de acuerdo con la técnica anterior. La Figura 1A ilustra el apareamiento directo de hebras adaptadoras al menos parcialmente complementarias para formar un agrupamiento de adaptadores con secuencias identificadoras de molécula única (SMI) complementarias y sustancialmente únicas. La Figura 1B ilustra el apareamiento directo de hebras adaptadoras complementarias para formar un agrupamiento de adaptadores con secuencias SMI no complementarias únicas.

- 20 La FIG. 2 Representa un método de Extensión Enzimática (EE) para fabricar adaptadores dúplex de acuerdo con la técnica anterior. La Figura 2A ilustra el apareamiento de una hebra de alargamiento con una hebra de plantilla en una región al menos parcialmente complementaria. La Figura 2B ilustra la extensión de la hebra de alargamiento mediante una reacción enzimática, tal como la que implica una polimerasa, que da como resultado la formación de una secuencia de doble hebra que tiene un SMI de doble hebra. La Figura 2C ilustra formas enzimáticas, químicas u otras formas de escisión en una región (por ejemplo, un dominio de ligación) de la secuencia de doble hebra que da como resultado un extremo ligable en el producto adaptador deseado, y un fragmento escindido.

- 25 La FIG. 3 representa intermedios incompletos de la síntesis del adaptador en cada etapa de la síntesis. El Panel A ilustra el apareamiento de la hebra de alargamiento y la hebra de plantilla durante una forma de síntesis del adaptador EE. Las hebras de plantilla residuales y las hebras de alargamiento residuales que no se recogieron durante esta etapa (Intermedios 1 y 2, respectivamente) pueden permanecer en la mezcla de reacción. Después del apareamiento, las hebras de alargamiento y las hebras de plantilla forman complejos preadaptadores que se deben extender enzimáticamente para crear una secuencia adaptadora de doble hebra. Posteriormente, los complejos preadaptadores, residuales no extendidos, o incompletamente extendidos (Intermedio 3) también pueden quedar en la mezcla de reacción después de la etapa de extensión. Otra fuente de intermedios incompletos durante la síntesis del adaptador ocurre cuando un sitio de corte de la región 3' de doble hebra no se corta o se corta de manera incompleta, lo que resulta en complejos adaptadores residuales no cortados (Intermedio 4), como se ilustra en el Panel B. El Panel C ilustra una etapa de procesamiento final de la síntesis del adaptador, que también produce un subproducto de reacción indeseable, un fragmento escindido, junto con un producto adaptador deseado.

- 40 La FIG. 4 representa un ejemplo de intermedios incompletos de síntesis de adaptadores que conducen a resultados aberrantes de Secuenciación Dúplex. El Panel A ilustra un fragmento de ADN de biblioteca con salientes 3' A en cada extremo (centro) flanqueado por un adaptador completamente procesado (izquierda) y un adaptador no cortado (derecha). El Panel B ilustra la ligación de ambos adaptadores al fragmento de ADN de la biblioteca, en el que la ligación del adaptador no cortado (derecha) deja una muesca abierta en un 1ª hebra del fragmento de ADN de la biblioteca. El Panel C ilustra la amplificación por PCR asimétrica resultante de la mella abierta en donde solo se amplifica la 2ª hebra.

- 45 La FIG. 5 representa una realización de un método para un esquema de síntesis de adaptador en fase de solución utilizando un marcador de captura. El Panel A ilustra la hibridación de una hebra de plantilla y una hebra de alargamiento en una región al menos parcialmente complementaria para formar un complejo preadaptador, en el que la hebra de plantilla comprende un marcador de captura 5'. Los intermedios incompletos durante esta etapa incluyen el exceso de hebras de plantilla y de hebras de alargamiento (Intermedios 1 y 2, respectivamente). La extensión enzimática de la hebra de alargamiento del complejo preadaptador produce un complejo adaptador no cortado que comprende una secuencia de doble hebra que tiene un SMI en ambas hebras. Los intermedios incompletos posteriores durante esta etapa incluyen complejos preadaptadores no extendidos o incompletamente extendidos (Intermedio 3). Durante la siguiente etapa, se corta un sitio de corte de la región 3' de doble hebra (por ejemplo, un dominio de ligación) mediante tratamiento con, por ejemplo, una enzima endonucleasa de restricción, produciendo de esta manera el producto adaptador deseado y un subproducto del fragmento escindido. El Panel B ilustra un complejo adaptador no cortado que no se pudo cortar o que se cortó de forma incompleta mediante tratamiento con enzima (Intermedio 4). El panel C ilustra el producto adaptador deseado y el subproducto del fragmento de escisión. De forma notable, el subproducto del fragmento de escisión retiene el marcador de captura 5' después de cortar el sitio de corte de la región 3'. El Panel D ilustra el proceso de selección negativa de purificación del adaptador que implica la adición de una superficie funcionalizada que es capaz de unir el marcador de captura a la mezcla de reacción. La superficie

funcionalizada se une solo a los intermedios y subproductos que comprenden el marcador de captura, que incluye los intermedios 1, 3 y 4, y los subproductos. El producto adaptador deseado no marcado que tiene un SMI dúplex se purifica mediante eliminación de la superficie funcionalizada y los intermedios/subproductos unidos.

La FIG. 6 representa una realización de un método para un esquema de síntesis de adaptador en fase de solución utilizando dos marcadores de captura. El Panel A ilustra el apareamiento de una hebra de plantilla y una hebra de alargamiento en una región al menos parcialmente complementaria para formar un complejo preadaptador, en el que la hebra de plantilla comprende un marcador de captura escindible en 3' y un marcador de captura en 5'. Los intermedios incompletos durante esta etapa incluyen el exceso de hebras de plantilla y hebras de alargamiento (Intermedios 1 y 2, respectivamente). La extensión enzimática de la hebra de alargamiento del complejo preadaptador produce un complejo adaptador no cortado que comprende una secuencia de doble hebra que tiene un SMI en cada hebra. Los intermedios incompletos posteriores durante este procesamiento incluyen complejos preadaptadores, no extendidos, o incompletamente extendidos (Intermedio 3). Durante la siguiente etapa, el sitio de corte del dominio de ligación se corta mediante tratamiento con, por ejemplo, una enzima endonucleasa de restricción, produciendo de esta manera el producto adaptador deseado y un subproducto del fragmento escindido. El Panel B ilustra un complejo adaptador no cortado que no se pudo cortar o se cortó de forma incompleta mediante tratamiento con enzima (Intermedio 4). El Panel C ilustra el producto adaptador deseado y el subproducto del fragmento de escisión. Después del tratamiento con la enzima de restricción, el producto adaptador deseado retiene el marcador de captura escindible en 3' y el subproducto del fragmento de escisión retiene el marcador de captura en 5'. El Panel D ilustra los procesos de enriquecimiento/selección negativa y enriquecimiento/selección positiva de purificación de adaptadores que implican la adición gradual de superficies funcionalizadas que son capaces de unirse al marcador de captura 5', o unirse al marcador de captura escindible 3'. Durante el proceso de enriquecimiento/selección negativo, una primera superficie funcionalizada se une a los intermedios y subproductos que llevan el marcador de captura 5', que incluyen los Intermedios 1, 3 y 4, y los subproductos. El complejo adaptador que comprende un marcador de afinidad escindible en 3' se purifica inicialmente mediante la separación de la primera superficie funcionalizada con intermedios/subproductos unidos de la mezcla de reacción restante. Durante el proceso de enriquecimiento/selección positiva, se introduce una segunda superficie funcionalizada en la mezcla de reacción y se une al complejo adaptador mediante el marcador de captura escindible en 3'. Después de una etapa de lavado o similar, la superficie funcionalizada unida al adaptador se trata enzimáticamente para cortar el marcador de captura 3' escindible, liberando de esta manera el producto adaptador deseado en solución. A continuación se purifica el producto adaptador deseado que tiene un SMI dúplex mediante separación de la segunda superficie funcionalizada.

La FIG. 7 representa una realización de un método para un esquema de síntesis de adaptador en fase sólida utilizando un marcador de captura. El Panel A ilustra la adición de una hebra de plantilla que comprende un marcador de captura 5' y una superficie funcionalizada capaz de unir el marcador de captura a una mezcla de reacción. La hebra de plantilla se une a la superficie funcionalizada mediante el marcador de captura y se agrega una hebra de alargamiento para facilitar el apareamiento. El Panel B ilustra el apareamiento de la hebra de plantilla y la hebra de alargamiento en una región al menos parcialmente complementaria que forma un complejo preadaptador. Luego, el complejo preadaptador se extiende enzimáticamente para formar una secuencia de doble hebra que tiene un SMI en cada hebra. El Panel C ilustra el complejo adaptador no cortado extendido enzimáticamente que contiene un SMI de doble hebra. Después de cortar una región 3' de la secuencia de doble hebra (por ejemplo, el dominio de ligación) en un sitio de corte utilizando, por ejemplo, una enzima endonucleasa de restricción, el producto adaptador deseado se separa del fragmento escindido unido a la superficie funcionalizada como se ilustra en el Panel D. Luego, el producto adaptador deseado no marcado se purifica al separar la superficie funcionalizada con intermedios/subproductos unidos del producto adaptador deseado que tiene un SMI dúplex.

La FIG. 8 representa otra realización de un método para un esquema de síntesis de adaptadores en fase sólida. El Panel A ilustra una hebra de plantilla preunida a un soporte sólido de síntesis de oligonucleótidos (por ejemplo, perlas de vidrio de poro controlado) en un extremo terminal 3' y que se extiende desde el soporte sólido hasta un extremo terminal 5' en una mezcla de reacción. Se agrega una hebra de alargamiento a la mezcla de reacción donde se aparea con la hebra de plantilla en una región al menos parcialmente complementaria para formar un complejo preadaptador. El Panel B ilustra la extensión enzimática de la hebra de alargamiento seguida de escisión en un sitio de corte para formar un extremo ligable del complejo adaptador. Después de la escisión, el producto adaptador deseado que tiene un SMI dúplex se libera del soporte sólido de síntesis de oligonucleótidos y se recupera.

La FIG. 9 representa una realización de un método para un esquema de síntesis de adaptador híbrido en fase de solución/fase sólida utilizando un marcador de captura. El Panel A muestra la extensión enzimática de la hebra de alargamiento del complejo preadaptador en solución. Después de la extensión enzimática, se introduce una superficie funcionalizada para interactuar con un marcador de captura 5' en la hebra de plantilla del complejo adaptador no cortado (Panel B). El Panel C ilustra un evento de escisión en un sitio de corte de la región 3' de doble hebra (por ejemplo, un dominio de ligación), liberando de esta manera el producto adaptador purificado que tiene un SMI dúplex en solución.

La FIG. 10 representa otra realización de un método para un esquema de síntesis de adaptador que utiliza marcadores de captura simples o dobles y un ligador escindible. La Figura 10A ilustra un método de enriquecimiento/selección negativo para la síntesis de adaptadores utilizando una hebra de plantilla que comprende un marcador de captura 5' para eliminar subproductos indeseables y/o productos de síntesis de adaptadores intermedios indeseables, y un

ligador escindible que se puede cortar para generar una estructura de adaptador dúplex en forma de Y. La Figura 10B ilustra un método de síntesis de adaptador que comprende esquemas de enriquecimiento/selección tanto negativos como positivos utilizando una hebra de plantilla que comprende un marcador de captura 5' para eliminar subproductos indeseables y/o productos de síntesis de adaptador intermedios indeseables, y un segundo marcador de captura proporcionado en una secuencia ligadora de un estructura de bucle de horquilla configurada para seleccionar los productos adaptadores deseados en la que el ligador se ha escindido con éxito para formar una estructura adaptadora dúplex en forma de Y.

La FIG. 11 representa la incorporación de una secuencia índice durante la síntesis del adaptador. La Figura 11A ilustra una realización de un esquema de síntesis de adaptadores en fase sólida que tiene secuencias índice en las regiones no complementarias de las hebras de alargamiento y plantilla. Las hebras de alargamiento y plantilla pueden comprender las secuencias índice antes de la etapa de hibridación para generar un complejo adaptador preindexado adecuado para el marcado específico de una muestra de una biblioteca de ácidos nucleicos. La Figura 11B ilustra una realización de un esquema de síntesis de adaptador en fase sólida que tiene una secuencia índice (denominada "XXXX") presente en la hebra de plantilla. Después de las etapas de apareamiento y extensión de la síntesis del adaptador, la secuencia índice está presente en la región de doble hebra de la secuencia adaptadora tanto en la hebra de alargamiento como en la hebra de plantilla.

La FIG. 12 representa una realización de un método de un esquema de síntesis de adaptador inverso. Una hebra de plantilla que comprende una secuencia SMI y una estructura en forma de horquilla próxima a un extremo 3' se aparea con una hebra de alargamiento en una región al menos parcialmente complementaria. La estructura de bucle en horquilla incluye una porción de tallo autocomplementaria y una porción de ligador de nucleótido de hebra sencilla. Una polimerasa extiende el extremo 3' terminal de la hebra de plantilla para encontrarse con el extremo 5' de la hebra de alargamiento apareada para formar una región de doble hebra del complejo adaptador que tiene la secuencia SMI en ambas hebras. Se incorpora una base modificada o no estándar en la hebra de plantilla en la porción de tallo de tal manera que un evento de escisión dirigido en la base modificada o no estándar libera el producto adaptador deseado mientras se genera un extremo ligable del adaptador dúplex liberado. Otras etapas incluyen la reparación de mellas (ligación) y la separación de productos no deseados del producto adaptador deseado que tiene un SMI dúplex.

La FIG. 13 representa otras realizaciones de métodos de un esquema de síntesis de adaptador inverso. La Figura 13A ilustra un método de síntesis de adaptador inverso que utiliza una hebra de plantilla única que tiene regiones autocomplementarias en una región 5' y en una región 3' que son adecuadas para formar estructuras de bucle en horquilla. Se utiliza una polimerasa para extender la región complementaria entre el terminal 3' y el terminal 5' de la hebra de plantilla. Las Figuras 13A y 13B ilustran variaciones de realizaciones que utilizan una incorporación de base modificada o no estándar para liberar el adaptador de una superficie funcionalizada mientras se forma un extremo ligable, y/o para escindir la característica de horquilla 5' para formar un complejo adaptador en forma de Y. La Figura 13C ilustra una realización de síntesis de adaptador inverso en la que se genera una característica asimétrica 5' después de la extensión enzimática para formar un SMI de doble hebra en el complejo adaptador. La Figura 13D ilustra una variación de una realización de síntesis de adaptador inverso utilizando incorporación de base modificada o no estándar para formar un complejo adaptador en forma de Y sin utilizar una etapa de ligación.

Definiciones

Para que la presente divulgación se comprenda más fácilmente, a continuación se definen primero determinados términos. A lo largo de la especificación se establecen definiciones adicionales para los siguientes términos y otros términos.

En esta solicitud, a menos que se desprenda lo contrario del contexto, (i) se puede entender que el término "un" significa "al menos uno"; (ii) se puede entender que el término "o" significa "y/o"; (iii) se puede entender que los términos "que comprende" y "que incluye" abarcan componentes o etapas detalladas, ya sea que se presenten por sí mismas o junto con uno o más componentes o etapas adicionales; (iv) se puede entender que los términos "alrededor de" y "aproximadamente" permiten una variación estándar como lo entenderían los expertos con conocimientos básicos en la técnica; y (v) cuando se proporcionan rangos, se incluyen los criterios de valoración. Como se utiliza en esta solicitud, el término "comprender" y variaciones del término, tales como "que comprende" y "comprende", no pretenden excluir otros aditivos, componentes, números enteros o etapas.

Acerca de: El término "acerca de", cuando se utiliza en el presente documento en referencia a un valor, se refiere a un valor que es similar, en contexto, al valor al que se hace referencia. En general, los expertos en la técnica, familiarizados con el contexto, apreciarán el grado relevante de variación abarcado por "alrededor de" en ese contexto. Por ejemplo, en algunas realizaciones, el término "alrededor de" puede abarcar un rango de valores que están dentro del 25 %, 20 %, 19 %, 18 %, 17 %, 16 %, 15 %, 14 %, 13 %, 12 %, 11 %, 10 %, 9 %, 8 %, 7 %, 6 %, 5 %, 4 %, 3 %, 2 %, 1 %, o menos del valor referido.

Ácido nucleico: Como se utiliza en el presente documento, en su sentido más amplio, se refiere a cualquier compuesto y/o sustancia que esté o se pueda incorporar a una cadena de oligonucleótidos. En algunas realizaciones, un ácido nucleico es un compuesto y/o sustancia que se incorpora o se puede incorporar en una cadena de oligonucleótidos mediante un ligamiento fosfodiéster. Como quedará claro por el contexto, en algunas realizaciones, "ácido nucleico"

se refiere a un residuo de ácido nucleico individual (por ejemplo, un nucleótido y/o nucleósido); en algunas realizaciones, "ácido nucleico" se refiere a una cadena de oligonucleótidos que comprende residuos de ácidos nucleicos individuales. En algunas realizaciones, un "ácido nucleico" es o comprende ARN; en algunas realizaciones, un "ácido nucleico" es o comprende ADN. En algunas realizaciones, un ácido nucleico es, comprende, o consiste en uno o más residuos de ácido nucleico naturales. En algunas realizaciones, un ácido nucleico es, comprende, o consiste en uno o más análogos de ácido nucleico. En algunas realizaciones, un análogo de ácido nucleico difiere de un ácido nucleico en que no utiliza una estructura principal de fosfodiéster. Por ejemplo, en algunas realizaciones, un ácido nucleico es, comprende, o consiste en uno o más "ácidos nucleicos de péptido", que son conocidos en la técnica y tienen enlaces peptídicos en lugar de enlaces fosfodiéster en la estructura principal, se consideran dentro del alcance de la presente tecnología. Alternativamente, o adicionalmente, en algunas realizaciones, un ácido nucleico tiene uno o más ligamientos fosforotioato y/o 5'-N-fosforamidita en lugar de enlaces fosfodiéster. En algunas realizaciones, un ácido nucleico es, comprende, o consiste en uno o más nucleósidos naturales (por ejemplo, adenosina, timidina, guanosina, citidina, uridina, desoxiadenosina, desoxitimidina, desoxiguanosina, y desoxicitidina). En algunas realizaciones, un ácido nucleico es, comprende, o consiste en uno o más análogos de nucleósidos (por ejemplo, 2-aminoadenosina, 2-tiotimidina, inosina, pirrolo-pirimidina, 3-metil adenosina, 5-metilcitidina, C-5 propinil-citidina, C-5 propinil-uridina, 2-aminoadenosina, C5 bromuridina, CS-fluoruridina, C5-yodouridina, C5-propinil-uridina, C5-propinil-citidina, C5-metilcitidina, 2-aminoadenosina, 7-desazaadenosina, 7-desazaguanosina, 8-oxoadenosina, 8-oxoguanosina, 0(6)-metilguanina, 2-tiocitidina, bases metiladas, bases intercaladas y combinaciones de las mismas). En algunas realizaciones, un ácido nucleico comprende uno o más azúcares modificados (por ejemplo, 2'-fluororibosa, ribosa, 2'-desoxirribosa, arabinosa, hexosa o ácidos Nucleicos Bloqueados) en comparación con los de los ácidos nucleicos de origen natural habituales. En algunas realizaciones, un ácido nucleico tiene una secuencia de nucleótidos que codifica un producto génico funcional tal como un ARN o una proteína. En algunas realizaciones, un ácido nucleico incluye uno o más intrones. En algunas realizaciones, un ácido nucleico puede ser un producto de ARN no codificante de proteínas, tal como un ARNm, un ARN ribosómico o un ARN guía CRISPER/Cas9. En algunas realizaciones, un ácido nucleico tiene un propósito regulador en un genoma. En algunas realizaciones, un ácido nucleico no surge de un genoma. En algunas realizaciones, un ácido nucleico incluye secuencias intergénicas. En algunas realizaciones, un ácido nucleico deriva de un elemento extracromosómico o un genoma no nuclear (mitocondrial, cloroplasto, etc.). En algunas realizaciones, los ácidos nucleicos se preparan mediante uno o más aislamiento de una fuente natural, síntesis enzimática mediante polimerización basada en una plantilla complementaria (in vivo o in vitro), reproducción en una célula o sistema recombinante, y síntesis química. En algunas realizaciones, un ácido nucleico es al menos 2, 3, 4, 5, 6, 7, 8, 9, 10, 15, 20, 25, 30, 35, 40, 45, 50, 55, 60, 65, 70, 75, 80, 85, 90, 95, 100, 110, 120, 130, 140, 150, 160, 170, 180, 190, 200, 225, 250, 275, 300, 325, 350, 375, 400, 425, 450, 475, 500, 600, 700, 800, 900, 1000, 1500, 2000, 2500, 3000, 3500, 4000, 4500, 5000 o más residuos de longitud. En algunas realizaciones, un ácido nucleico es parcial o totalmente de hebra sencilla; en algunas realizaciones, un ácido nucleico es parcial o totalmente de doble hebra. En algunas realizaciones, un ácido nucleico tiene una secuencia de nucleótidos que comprende al menos un elemento que codifica, o es el complemento de una secuencia que codifica, un polipéptido. En algunas realizaciones, un ácido nucleico tiene actividad enzimática. En algunas realizaciones, el ácido nucleico cumple una función mecánica, por ejemplo en un complejo de ribonucleoproteína o un ARN de transferencia. En algunas realizaciones, un ácido nucleico funciona como un aptámero. En algunas realizaciones se puede utilizar un ácido nucleico para el almacenamiento de datos. En algunas realizaciones, un ácido nucleico se puede sintetizar químicamente in vitro.

Elemento que Define la Hebra (SDE): Como se utiliza en el presente documento, el término "Elemento que Define la Hebra" o "SDE", se refiere a cualquier material que permita la identificación de una hebra específica de un material de ácido nucleico de doble hebra y por tanto la diferenciación de la otra hebra/complementaria (por ejemplo, cualquier material que hace que los productos de amplificación de cada uno de los dos ácidos nucleicos de hebra sencilla resultantes de un ácido nucleico de doble hebra diana sean sustancialmente distinguibles entre sí después de la secuenciación u otra interrogación de ácidos nucleicos). En algunas realizaciones, un SDE puede ser o comprender uno o más segmentos de secuencia sustancialmente no complementaria dentro de una secuencia adaptadora. En realizaciones particulares, un segmento de secuencia sustancialmente no complementaria dentro de una secuencia adaptadora se puede proporcionar mediante una molécula adaptadora que comprende una forma de Y o una forma de "bucle". En otras realizaciones, un segmento de secuencia sustancialmente no complementaria dentro de una secuencia adaptadora puede formar una "burbuja" no apareada en medio de secuencias complementarias adyacentes dentro de una secuencia adaptadora. En otras realizaciones, un SDE puede abarcar una modificación de ácido nucleico. En algunas realizaciones, un SDE puede comprender la separación física de hebras emparejadas en compartimentos de reacción físicamente separados. En algunas realizaciones, un SDE puede comprender una modificación química. En algunas realizaciones, un SDE puede comprender un ácido nucleico modificado. En algunas realizaciones, un SDE se puede relacionar con una variación de secuencia en una molécula de ácido nucleico causada por daño aleatorio o semialeatorio, modificación química, modificación enzimática u otra modificación de la molécula de ácido nucleico. En algunas realizaciones, la modificación puede ser la desaminación de metilcitosina. En algunas realizaciones, la modificación puede implicar sitios de mellas de ácido nucleico. Varias realizaciones de SDE se divulgan además en la Publicación de Patente Internacional No. WO2017/100441, .

Identificador de Molécula Única (SMI): Como se utiliza en el presente documento, el término "identificador de molécula única" o "SMI", (al que se puede hacer referencia como una "etiqueta", un "código de barras", un "código de barras molecular", un "identificador molecular", una "secuencia identificadora", un "Identificador Molecular Único", o "UMI", entre otros nombres) se refiere a cualquier material (por ejemplo, una secuencia de nucleótidos, una característica de

una molécula de ácido nucleico) que sea capaz de distinguir una molécula individual en una gran población heterogénea de moléculas. En algunas realizaciones, un SMI puede ser o comprender un SMI aplicado exógenamente. En algunas realizaciones, un SMI aplicado exógenamente puede ser o comprender una secuencia degenerada o semidegenerada. En algunas realizaciones, los SMI sustancialmente degenerados se pueden conocer como Identificadores Moleculares Únicos Aleatorios (R-UMI). En algunas realizaciones, un SMI puede comprender un código (por ejemplo, una secuencia de ácidos nucleicos) de un agrupamiento de códigos conocidos. En algunas realizaciones, los códigos SMI predefinidos se conocen como Identificadores Moleculares Únicos Definidos (D-UMI). En algunas realizaciones, un SMI puede ser o comprender un SMI endógeno. En algunas realizaciones, un SMI endógeno puede ser o comprender información relacionada con puntos de corte específicos de una secuencia diana, o características relacionadas con los extremos terminales de moléculas individuales que comprenden una secuencia diana. En algunas realizaciones, un SMI se puede relacionar con una variación de secuencia en una molécula de ácido nucleico causada por daño aleatorio o semialeatorio, modificación química, modificación enzimática u otra modificación a la molécula de ácido nucleico. En algunas realizaciones, la modificación puede ser la desaminación de metilicitosina. En algunas realizaciones, la modificación puede implicar sitios de mellas de ácido nucleico. En algunas realizaciones, un SMI puede comprender elementos tanto exógenos como endógenos. En algunas realizaciones, un SMI puede comprender elementos SMI físicamente adyacentes. En algunas realizaciones, los elementos SMI pueden ser espacialmente distintos en una molécula. En algunas realizaciones, un SMI puede ser un ácido no nucleico. En algunas realizaciones, un SMI puede comprender dos o más tipos diferentes de información de SMI. Varias realizaciones de SMI se divulgan además en la Publicación de Patente Internacional No. WO2017/100441, .

Hebra de alargamiento: Como se utiliza en el presente documento, el término “hebra de alargamiento” o “molécula de alargamiento” se refiere a una molécula de ácido polinucleico utilizada para la síntesis de adaptadores. En algunas realizaciones, la hebra de alargamiento es un oligonucleótido sintetizado de secuencia definida. En otras realizaciones, la hebra de alargamiento es un oligonucleótido sintetizado que tiene uno o más nucleótidos indefinidos (por ejemplo, aleatorios o semialeatorios). En algunas realizaciones, la hebra de alargamiento está compuesta por cuatro o más nucleótidos. En algunas realizaciones, la hebra de alargamiento comprende un SDE. Por ejemplo, la hebra de alargamiento puede incluir una o más porciones que no son complementarias con respecto a una hebra de plantilla. En algunas realizaciones, la hebra de alargamiento comprende un SMI. En algunas realizaciones, la hebra de alargamiento comprende una secuencia índice (por ejemplo, código de barras índice, etiqueta índice) u otra etiqueta de identificación. En algunas realizaciones, la hebra de alargamiento comprende un sitio de unión al cebador. En algunas realizaciones, la hebra de alargamiento comprende un sitio de unión del cebador dentro del SDE. En algunas realizaciones, la hebra de plantilla comprende una secuencia cebadora de lectura. En algunas realizaciones, la hebra de alargamiento comprende una o más regiones complementarias a una hebra de plantilla.

Hebra de plantilla: Como se utiliza en el presente documento, el término “hebra de plantilla” o “molécula de plantilla” se refiere a una molécula de ácido polinucleico utilizada para la síntesis de adaptadores. En algunas realizaciones, la hebra de plantilla está compuesta por aproximadamente 10 o más, pero preferiblemente 20 o más nucleótidos. En algunas realizaciones la hebra de plantilla es un oligonucleótido sintetizado de secuencia definida. En otras realizaciones la hebra de plantilla es un oligonucleótido sintetizado que tiene uno o más nucleótidos indefinidos (por ejemplo, aleatorios o semialeatorios). En algunas realizaciones, la hebra de plantilla tiene regiones de autocomplementariedad. En algunas realizaciones, la hebra de plantilla comprende un SDE. Por ejemplo, la hebra de plantilla puede incluir una o más porciones que no son complementarias con respecto a una hebra de alargamiento. En algunas realizaciones, la hebra de plantilla comprende un SMI. En algunas realizaciones, la hebra de plantilla comprende una secuencia índice (por ejemplo, código de barras índice, etiqueta índice) u otra etiqueta de identificación. En algunas realizaciones, la hebra de plantilla comprende un sitio de unión al cebador. En algunas realizaciones, la hebra de plantilla comprende un sitio de unión al cebador dentro del SDE. En algunas realizaciones, la hebra de plantilla comprende una secuencia cebadora de lectura. En algunas realizaciones, la hebra de plantilla comprende una o más regiones complementarias a una hebra de alargamiento. En algunas realizaciones, la hebra de plantilla comprende uno o más nucleótidos modificados o no estándar. En algunas realizaciones, dichos nucleótidos modificados o no estándar son escindibles enzimáticamente. En algunas realizaciones, una hebra de plantilla comprende un dominio de ligación. En algunas realizaciones, el dominio de ligación comprende una porción escindible o un sitio de reconocimiento. En algunas realizaciones, una hebra de plantilla comprende un marcador de captura configurada para unirse a una fracción asociada de unión.

Marcador de captura: Como se utiliza en el presente documento, el término “marcador de captura” (que también se puede denominar como una “etiqueta de captura”, “fracción de captura”, “marcador de afinidad”, “etiqueta de afinidad”, “etiqueta de epítipo”, “etiqueta”, “fracción o grupo químico “presa”, entre otros nombres) se refiere a una fracción que se puede integrar en o sobre una molécula diana o sustrato, con fines de purificación. En algunas realizaciones, el marcador de captura se selecciona de un grupo que comprende una molécula pequeña, un ácido nucleico, un péptido o cualquier fracción que se pueda unir de forma única. En algunas realizaciones, el marcador de captura está fijado al 5' de una molécula de ácido nucleico. En algunas realizaciones, el marcador de captura está fijado al extremo 3' de una molécula de ácido nucleico. En algunas realizaciones, el marcador de captura está conjugado con un nucleótido dentro de la secuencia interna de una molécula de ácido nucleico que no está en ninguno de los extremos. En algunas realizaciones, el marcador de captura es una secuencia de nucleótidos dentro de la molécula de ácido nucleico. En algunas realizaciones, el marcador de captura se selecciona de un grupo de biotina, biotina desoxitimidina dT, biotina NHS, biotina TEG, destiobiotina NHS, digoxigenina NHS, DNP TEG, tioles, entre otros. En algunas realizaciones, los

marcadores de captura incluyen, sin limitación, biotina, avidina, estreptavidina, un hapteno reconocido por un anticuerpo, una secuencia de ácidos nucleicos particular y partículas magnéticamente atraíbles. En algunas realizaciones, la modificación química (por ejemplo, Acridite™- modificado, adenilado, modificado con azida, modificado con alquino, Linker™ modificado, etc.) de moléculas de ácido nucleico pueden servir como un marcador de captura.

Fracción de extracción: Como se utiliza en el presente documento, el término “fracción de extracción” (que también se puede denominar como un “compañero de unión”, “compañero de afinidad”, una fracción “cebo” o grupo químico entre otros nombres) se refiere a una fracción aislable o cualquier tipo de Molécula que permite la separación por afinidad de los ácidos nucleicos que llevan el marcador de captura de los ácidos nucleicos que carecen del marcador de captura. En algunas realizaciones, la fracción de extracción se selecciona de un grupo que comprende una molécula pequeña, un ácido nucleico, un péptido, un anticuerpo o cualquier fracción que se pueda ligar de forma única. La fracción de extracción puede estar ligada o ser ligable a una fase sólida u otra superficie para formar una superficie funcionalizada. En algunas realizaciones, la fracción de extracción es una secuencia de nucleótidos ligada a una superficie (por ejemplo, una superficie sólida, perla, partícula magnética, etc.). En algunas realizaciones, la fracción de extracción se selecciona de un grupo de avidina, estreptavidina, un anticuerpo, una etiqueta de polihistadina, una etiqueta FLAG o cualquier modificación química de una superficie para la química de adhesión. Los ejemplos no limitantes de estos últimos incluyen grupos azida y alquino que pueden formar enlaces 1,2,3-triazol mediante métodos “Click”, o tiol, una azida y alquino terminal, las superficies modificadas con tiol pueden reaccionar covalentemente con oligonucleótidos modificados con Acridita y superficies modificadas con aldehídos y cetonas que pueden reaccionar para fijar oligonucleótidos marcados con I-Linker™.

Superficie funcionalizada: Como se utiliza en el presente documento, el término “superficie funcionalizada” se refiere a una superficie sólida, una perla, u otra estructura fija que sea capaz de unirse o inmovilizar un marcador de captura. En algunas realizaciones, la superficie funcionalizada comprende una fracción de extracción capaz de unirse a un marcador de captura. En algunas realizaciones, una fracción de extracción está ligada directamente a una superficie. En algunas realizaciones, la modificación química de la superficie funciona como una fracción de extracción. En algunas realizaciones, una superficie funcionalizada puede comprender vidrio de poro controlado (CPG), vidrio poroso magnético (MPG), entre otras superficies de vidrio o no vidrio. La funcionalización química puede implicar modificación de cetonas, modificación de aldehídos, modificación de tioles, modificación de azidas y modificaciones de alquinos, entre otras. En algunas realizaciones, la superficie funcionalizada y un oligonucleótido utilizado para la síntesis de adaptadores se ligan utilizando uno o más de un grupo de químicas de inmovilización que forman enlaces amida, enlaces alquilamina, enlaces tiourea, enlaces diazo, enlaces hidracina, entre otras químicas de superficie. En algunas realizaciones, la superficie funcionalizada y un oligonucleótido utilizado para la síntesis de adaptadores se ligan utilizando uno o más de un grupo de reactivos que incluyen EDAC, NHS, peryodato de sodio, glutaraldehído, disulfuros de piridilo, ácido nitroso, biotina, entre otros reactivos de ligamiento.

Sitio de corte: También denominado “sitio de escisión” y “sitio de mella”, es el enlace, o par de enlaces entre nucleótidos en una molécula de ácido nucleico. En el caso de moléculas de ácido nucleico de doble hebra, tales como el ADN de doble hebra, el sitio de corte puede implicar enlaces (comúnmente enlaces fosfodiéster) que son inmediatamente adyacentes entre sí en una molécula de doble hebra, de tal manera que después del corte se forma un extremo “romo”. El sitio de corte también puede implicar dos enlaces de nucleótidos que están en cada hebra sencilla del par que no son inmediatamente opuestos entre sí de tal manera que cuando se escinde queda un “extremo pegajoso”, por lo que regiones de nucleótidos de hebra sencilla permanecen en los extremos terminales de las moléculas. Los sitios de corte se pueden definir mediante una secuencia de nucleótidos particular que sea capaz de ser reconocida por una enzima, tal como una enzima de restricción, u otra endonucleasa con capacidad de reconocimiento de secuencia, tal como CRISPER/Cas9. El sitio de corte puede estar dentro de la secuencia de reconocimiento de dichas enzimas (es decir enzimas de restricción de tipo 1) o adyacente a ellas por algún intervalo definido de nucleótidos (es decir enzimas de restricción de tipo 2). Los sitios de corte también se pueden definir por la posición de nucleótidos modificados que son capaces de ser reconocidos por determinadas nucleasas. Por ejemplo, los sitios abásicos se pueden reconocer y escindidos por la endonucleasa VII así como por la enzima FPG. La enzima UDG puede reconocer y convertir la base de uracilo en sitios abásicos. Los nucleótidos que contienen ribosa en una secuencia de ADN se pueden reconocer y escindidos por RNAsaH2 cuando se hibridan con secuencias de ADN complementarias.

Sustancialmente: Como se utiliza en el presente documento, el término “sustancialmente” se refiere a la condición cualitativa de exhibir una extensión o grado total o casi total de una característica o propiedad de interés. Un experto con conocimientos básicos en las técnicas biológicas comprenderá que los fenómenos biológicos y químicos rara vez, o nunca, llegan a su fin y/o proceden a su plenitud o logran o evitan un resultado absoluto. Por lo tanto, el término “sustancialmente” se utiliza en el presente documento para captar la posible falta de integridad inherente a muchos fenómenos biológicos y químicos.

Descripción detallada

La presente tecnología está dirigida a reactivos y adaptadores para aplicaciones de secuenciación de ácidos nucleicos, tal como la Secuenciación Dúplex. Otros aspectos de la tecnología están dirigidos a métodos para fabricar y utilizar dichos reactivos y adaptadores. En particular, algunas realizaciones de la tecnología están dirigidas a fabricar un adaptador dúplex purificado que tiene una secuencia identificadora única (por ejemplo, un SMI) en cada hebra del

adaptador. Por ejemplo, diversas realizaciones de la presente tecnología incluyen sintetizar y purificar un conjunto de adaptadores de secuenciación dúplex para preparar una biblioteca de secuencias adecuada para realizar métodos de Secuenciación Dúplex. Diversos aspectos de la presente tecnología son adecuados para preparar adaptadores de doble hebra para uso en Secuenciación Dúplex así como otras aplicaciones de secuenciación.

Los detalles específicos de varias realizaciones de la tecnología se describen a continuación con referencia a las FIG. 1A-13D. Las realizaciones pueden incluir, por ejemplo, métodos para preparar un adaptador dúplex y reactivos asociados para uso en dichos métodos. Algunas realizaciones de la tecnología están dirigidas a preparar un adaptador dúplex purificado para mejorar aspectos de la Secuenciación Dúplex. Aunque muchas de las realizaciones se describen en el presente documento con respecto a la síntesis y generación de adaptadores y otros reactivos para uso con métodos de Secuenciación Dúplex, otras aplicaciones y usos de los adaptadores y reactivos además de los descritos en el presente documento están dentro del alcance de la presente tecnología. Adicionalmente, varias otras realizaciones de la tecnología pueden tener configuraciones o procedimientos diferentes a los descritos en el presente documento. Por lo tanto, un experto con conocimientos básicos en la técnica comprenderá que la tecnología puede tener otras realizaciones con elementos adicionales y que la tecnología puede tener otras realizaciones sin varias de las características mostradas y descritas a continuación con referencia a las Figuras 1A-13D. Varios aspectos de la presente tecnología se describen en detalle en las siguientes secciones. El uso de secciones no pretende limitar la tecnología actual.

La Secuenciación Dúplex es un método para producir secuencias de ADN con corrección de errores a partir de moléculas de ácido nucleico de doble hebra y que se describió originalmente en la Publicación de Patente Internacional No. WO 2013/142389 y en la Patente de EE.UU. No. 9,752,188. En determinadas realizaciones, los métodos que incorporan Secuenciación Dúplex pueden incluir la ligación de uno o más adaptadores de secuenciación a una molécula de ácido nucleico de doble hebra diana, que comprende una primera secuencia de ácidos nucleicos diana de hebra y una segunda secuencia nucleica diana de hebra, para producir un complejo de ácido nucleico diana de doble hebra. En diversas realizaciones, el complejo de ácido nucleico diana resultante en la Secuenciación Dúplex puede incluir al menos una secuencia de Identificador Molecular Único (SMI), que puede implicar una secuencia degenerada o semidegenerada aplicada exógenamente (que se puede denominar como una "etiqueta" código de barras", un "Identificador Molecular Único" o "UMI", entre otros nombres), información endógena relacionada con los puntos de corte específicos de la molécula de ácido nucleico de doble hebra diana, o una combinación de los mismos. El SMI puede hacer que la molécula de ácido nucleico diana sea sustancialmente distinguible de la pluralidad de otras moléculas en una población que se está secuenciando. La característica sustancialmente distinguible del elemento SMI se puede transportar independientemente por cada una de las hebras simples que forman la molécula de ácido nucleico de doble hebra de tal manera que se pueda reconocer que los productos de amplificación derivados de cada hebra provienen de la misma molécula de ácido nucleico de doble hebra sustancialmente única original después de la secuenciación. En otras realizaciones, el SMI puede incluir información adicional y/o se puede utilizar en otros métodos para los cuales dicha funcionalidad de distinción de moléculas es útil, tal como los descritos en las publicaciones mencionadas anteriormente.

En algunas realizaciones, cada complejo de secuencia de ácidos nucleicos diana de doble hebra en la Secuenciación Dúplex puede incluir además un elemento que hace que los productos de amplificación de los dos ácidos nucleicos de hebra sencilla que forman el ácido nucleico de doble hebra diana sean sustancialmente distinguibles entre sí después de la secuenciación. En algunas realizaciones, este elemento se puede denominar Elemento que Define la Hebra (SDE). En una realización, el SDE puede comprender sitios de cebador asimétricos comprendidos dentro de los adaptadores de secuenciación, o, en otras disposiciones, se pueden introducir asimetrías de secuencia en las moléculas adaptadoras que no están dentro de las secuencias de cebador, de tal manera que al menos una posición en las secuencias de nucleótidos del primer complejo de secuencia de ácidos nucleicos diana de hebra y el segundo complejo de secuencia de ácidos nucleicos diana son diferentes entre sí después de la amplificación y secuenciación. En otras realizaciones, el SMI puede comprender otra asimetría bioquímica entre las dos hebras que difiere de las secuencias de nucleótidos canónicas A, T, C, G o U, pero se convierte en al menos una secuencia de nucleótidos canónica diferente en las dos moléculas amplificadas y secuenciadas. En aún otra realización, el SDE puede ser un medio para separar físicamente las dos hebras antes de la amplificación, de tal manera que los productos de amplificación derivados de la primera secuencia de ácidos nucleicos diana de hebra y la segunda secuencia de ácidos nucleicos diana de hebra se mantengan en un aislamiento físico sustancial de uno y otros con el fin de mantener una distinción entre los dos. Se pueden utilizar otras disposiciones o metodologías para proporcionar una función SDE que permita distinguir la primera y la segunda hebras, tales como las descritas en las publicaciones mencionadas anteriormente, u otros métodos que sirvan al propósito funcional descrito.

Se han descrito previamente varios métodos para sintetizar adaptadores de secuenciación dúplex y se describen además en la Patente de EE.UU. No. 9,752,188 y en la Publicación de Patente Internacional No. WO2017/100441,.

Un método para fabricar SMI en Adaptadores Dúplex incluye la hibridación de oligonucleótidos que comprenden secuencias conocidas (incluidos oligonucleótidos ligados químicamente o dos porciones de un único bucle de oligonucleótido continuo). Se muestran ejemplos no limitantes de un método de Apareamiento de Oligonucleótidos (OA) para fabricar adaptadores dúplex que tienen un SMI en cada hebra en la Figura 1A y la Figura 1B. En este enfoque, se pueden sintetizar oligonucleótidos de secuencia conocida predeterminada y las correspondientes secuencias del complemento. En una realización mostrada en la Figura 1A, un oligonucleótido y su oligonucleótido

complementario se pueden aparear para formar un complejo adaptador que tiene un SMI dúplex de secuencia predeterminada. Luego se pueden combinar dos o más adaptadores con diferentes secuencias SMI para formar un agrupamiento de adaptadores que comprende dos o más SMI. Se pueden mezclar múltiples adaptadores con diferentes secuencias SMI para generar un conjunto de adaptadores que puede ser adecuado para algunas aplicaciones de Secuenciación Dúplex. Como se ilustra, el apareamiento directo de hebras adaptadoras complementarias para formar un agrupamiento de adaptadores con secuencias SMI complementarias únicas proporciona un conjunto de adaptadores adecuados para la preparación de bibliotecas de ácidos nucleicos.

Sin embargo, las secuencias SMI creadas en el método OA de síntesis de adaptadores no necesitan ser complementarias. Por ejemplo, la Figura 1B ilustra el apareamiento directo de hebras adaptadoras complementarias para formar un agrupamiento de adaptadores con secuencias SMI no complementarias únicas. Por ejemplo, los oligonucleótidos de secuencia conocida predeterminada se pueden aparear de tal manera que un SMI de hebra sencilla en una hebra se pueda relacionar informáticamente con el de la hebra opuesta de tal manera que los productos de amplificación derivados de las dos hebras se puedan reconocer como derivados a partir de los mismos dúplex de doble hebra de partida. Luego se pueden combinar dos o más versiones diferentes de estos adaptadores para formar un agrupamiento de dos o más tipos de adaptadores que comprenden dos o más SMI. Un ejemplo no limitante de este método de Apareamiento de Oligonucleótidos (OA) para fabricar SMI en Adaptadores Dúplex se muestra en la Figura 1B. De acuerdo con lo anterior, dichos adaptadores comprenderían secuencias SMI en regiones no complementarias (por ejemplo, en brazos en forma de Y, dentro de una burbuja, etc.). En el ejemplo ilustrado en la Figura 1B, diferentes SMI están representados por diferentes series de cuatro letras (es decir "PPPP" y "QQQQ"), que se registran en el momento de la mezcla como pertenecientes a una misma, pero que al menos no son parcialmente complementarias.

En ambos ejemplos, conocer sustancialmente la secuencia de las hebras individuales que se están apareando proporciona información importante para uso posterior. Por ejemplo, al menos en parte, la etiqueta SMI se puede utilizar para relacionar lecturas de secuencia que se derivan de una hebra con las de la otra hebra del dúplex de ADN original. De acuerdo con lo anterior, los adaptadores proporcionan la capacidad de distinguir sustancialmente aquellas hebras de una molécula de doble hebra fundadora particular de aquellas derivadas de otras moléculas de doble hebra originales en la población que se secuencia.

En algunos casos, el método OA para crear SMI en Adaptadores Dúplex proporciona simplicidad de síntesis. Por ejemplo, se puede generar un agrupamiento de 96 adaptadores diferentes al aparearse individualmente 96 hebras "superiores" con 96 hebras "inferiores" correspondientes en 96 pocillos individuales, seguido al mezclar de los adaptadores apareados individualmente. En un ejemplo no limitante, aunque se requiere la síntesis de 196 oligonucleótidos individuales, la preparación del agrupamiento de adaptadores de 96 etiquetas SMI requiere poco esfuerzo más allá de simplemente aparear oligonucleótidos y mezclar. Si está presente en cantidades de adaptador aproximadamente equimolares, y con un adaptador que contiene Dúplex SMI ligado a ambos extremos de cada fragmento de biblioteca, con el uso únicamente de información de estas etiquetas SMI, hay $96 \times 96 = 9216$ formas diferentes para marcar un fragmento determinado. Aunque en algunas aplicaciones donde se secuencia una gran cantidad de fragmentos de biblioteca, 9216 combinaciones de etiquetas son inadecuadas para distinguir sustancialmente cada fragmento de biblioteca entre sí (es decir si se secuenciaran 100,000 fragmentos, se esperaría que aproximadamente 10 fragmentos independientes se etiquetaran de manera idéntica), cuando se combina con información SMI codificada en puntos de corte aleatorios, semialeatorios o de otro modo heterogéneos de moléculas de biblioteca, la diversidad acumulativa de los elementos SMI exógenos más endógenos es suficiente para que cada fragmento de biblioteca esté sustancialmente marcado de forma única de tal manera que pueda ser sustancialmente diferenciadas de otras moléculas en el mismo agrupamiento de bibliotecas.

Para aplicaciones en las que los fragmentos de ácido nucleico que se secuencian no tienen, o tienen un grado menor, de diversidad de puntos de corte, se puede confiar menos o ninguna en la información SMI relacionada con los puntos de corte de la biblioteca endógena para contribuir a la información SMI agregada. Si no existe diversidad de puntos de corte, la probabilidad de un "choque de etiquetas" (es decir cuando dos moléculas independientes llevan la misma información SMI) se vuelve alta cuando se utiliza un pequeño agrupamiento de adaptadores SMI producidos por el método OA para una serie de secuenciación de tamaño típico. Por ejemplo, cuando los fragmentos de la biblioteca de secuenciación se generan mediante digestión enzimática específica de la secuencia, esencialmente toda la información de Dúplex SMI debe estar contenida dentro de las propias etiquetas del adaptador. Si la profundidad de la secuenciación excede el número de variantes de etiquetas, necesariamente se producirán choques de etiquetas y puede disminuir la eficacia de la secuenciación dúplex.

En algunos casos, un enfoque de extensión enzimática (EE) para fabricar Adaptadores Dúplex proporciona una cantidad mucho mayor de secuencias SMI generadas. Por ejemplo, y en una realización, un oligonucleótido se puede hibridar con otro oligonucleótido que contiene una secuencia de nucleótidos degenerada o semidegenerada en una región de no complementariedad. Los oligonucleótidos hibridados luego se pueden ligar químicamente o pueden ser dos porciones de un oligonucleótido continuo que cuando se hibrida forma un "bucle" o una forma de "U". Luego se puede utilizar una enzima capaz de polimerizar nucleótidos para copiar una región de hebra sencilla degenerada o semidegenerada de tal manera que se sintetice un complemento. De este modo se produce una secuencia de doble hebra complementaria degenerada o semidegenerada que puede servir como al menos un elemento SMI durante la Secuenciación Dúplex. El sitio de ligación en la molécula adaptadora puede modificarse a partir de este producto de

extensión mediante manipulación enzimática o química, por ejemplo mediante digestión de restricción, actividad transferasa terminal de una polimerasa u otra enzima o cualquier otro método conocido en la técnica.

Aunque se requieren más etapas en el enfoque EE, se necesitan tan solo dos lotes de oligonucleótidos para generar una cantidad mucho mayor de secuencias SMI. En un ejemplo particular, para Adaptadores EE Dúplex con una secuencia completamente degenerada de 10 pares de bases, teóricamente existirán $4^{10} = \sim$ sesenta millones de etiquetas SMI. Cuando estos adaptadores se ligan a ambos extremos de los fragmentos de la biblioteca, hay $4^{20} = \sim 3.6 \times 10^{15}$ combinaciones SMI diferentes. Incluso cuando se secuencian miles de millones de fragmentos de bibliotecas donde no existe diversidad de puntos de corte, la probabilidad de choques de etiquetas es extremadamente baja. Un ejemplo no limitante de un método EE para producir adaptadores de Secuenciación Dúplex se ilustra en la Figura 2. Con referencia a la Figura 2A, la hebra superior, o la hebra de alargamiento, se aparea con una hebra inferior, o hebra de plantilla, que contiene una secuencia SMI degenerada o semidegenerada, formando de esta manera un complejo preadaptador. Luego, el extremo 3' de la hebra de alargamiento se extiende para convertir tanto la secuencia SMI como un dominio de ligación en de doble hebra, creando un complejo adaptador no cortado (Figura 2B). El sitio de corte del dominio de ligación del adaptador se escinde luego con una enzima endonucleasa de restricción para dejar un saliente 3' "T" que es compatible para la ligación con un saliente 3' "A" en un fragmento de biblioteca preparado (Figura 2C). En determinadas realizaciones, el dominio de ligación resultante es un par de bases único de timina (T) que sobresale en el extremo 3' de la hebra de extensión extendida, pero en otras realizaciones, puede ser un extremo romo, o un tipo diferente o extremo 3' o 5'. saliente "pegajoso". En este ejemplo particular, "CUT" implica el uso de una endonucleasa específica de secuencia, tal como una enzima de restricción, para escindir de una manera que crea inherentemente el extremo ligable. En otras realizaciones, después de la escisión, un procesamiento enzimático o químico adicional, tal como con una transferasa terminal, puede crear el extremo ligable.

Además del producto adaptador final deseado, en la reacción pueden quedar otros productos incompletos. Por ejemplo, el exceso de oligonucleótidos de hebra sencilla (por ejemplo, hebra de alargamiento y hebra de plantilla), complejo preadaptador no extendido o extendido incompletamente, complejo adaptador no cortado y fragmentos intermedios y subproductos escindidos son impurezas que pueden disminuir la eficiencia de la ligación del adaptador. La Figura 3 es un esquema que ilustra la formación de dichos intermedios y subproductos durante el proceso de síntesis del adaptador. Con referencia a la Figura 3A, La hebra de plantilla de hebra sencilla residual y la hebra de alargamiento que no están hibridadas o en complejo pueden formar productos "Intermedios 1 y 2", respectivamente. Más referencia a la Figura 3A, el complejo preadaptador, residual no extendido, o incompletamente extendido puede formar productos "Intermedios 3" que permanecen en solución. El procesamiento adicional de productos de preadaptador de doble hebra incluye cortar un extremo 3' del complejo para generar un extremo ligable (por ejemplo, dentro de un dominio de ligación). Con referencia a la Figura 3B, los complejos preadaptadores de doble hebra no cortados pueden formar productos "Intermedios 4" que permanecen en solución. Incluso cuando se corta con éxito, se genera un fragmento escindido ("Subproducto") junto con el adaptador dúplex deseado.

Los productos intermedios y los subproductos pueden servir simplemente para contaminar los agrupamientos de adaptadores más grandes que se generan, reduciendo potencialmente la eficiencia del adaptador y el rendimiento general.

Más allá de simplemente disminuir la disponibilidad de productos adaptadores correctos y completamente formados, algunos intermedios y subproductos incompletos pueden ser activamente perjudiciales para varias etapas del proceso de Secuenciación Dúplex. Por ejemplo, y como se ilustra en la Figura 4, un complejo adaptador intermedio no cortado con una "A" que sobresale en 3' de la actividad transferasa terminal de la polimerasa se puede forzar a ligarse a un fragmento de biblioteca con cola A bajo determinadas condiciones. Debido a que el extremo 5' de la hebra inferior no cortado normalmente no está fosforilado, no se puede ligar mientras que la hebra superior sí. La ligación de dicho adaptador a un extremo de un fragmento de biblioteca creará un producto en el que solo se han formado tres de los cuatro enlaces fosfodiéster previstos (Figura 4B). La Figura 4C ilustra que la amplificación por PCR de este producto solo generará copias secuenciables de una hebra y la Secuenciación Dúplex ya no es posible en este escenario.

Se pueden realizar una variedad de optimizaciones para mejorar la integridad de la reacción de síntesis, pero no se pueden evitar todos los subproductos incompletos. La purificación de productos adaptadores EE completos puede mejorar la eficiencia de la Secuenciación Dúplex, pero puede ser difícil fabricarlos basándose únicamente en las diferentes propiedades de tamaño molecular de los productos completos frente a los incompletos. Sin embargo, y de acuerdo con diversos aspectos de la presente tecnología, determinadas modificaciones a los oligonucleótidos utilizados para la síntesis pueden alterar significativamente las propiedades químicas de los productos adaptadores completos versus incompletos y facilitar la purificación.

A. Diversas realizaciones de métodos y reactivos para la síntesis de adaptadores en fase de solución

La presente divulgación, entre otras cosas, proporciona métodos y reactivos para la síntesis de adaptadores en fase de solución. En algunas realizaciones que incluyen dichos métodos, se pueden utilizar uno o más marcadores de captura para el enriquecimiento/selección de producto(s) adaptador(es) deseado(s) a partir de productos de síntesis crudos, por ejemplo, mediante enriquecimiento/selección positiva de productos deseados o enriquecimiento/selección negativa para excluir o reducir la abundancia de productos/subproductos/subproductos no deseados.

Por ejemplo, en algunas realizaciones que incluyen enriquecimiento positivo, un oligonucleótido adaptador puede tener un marcador de captura que es o comprende una fracción química fijada (por ejemplo, biotina) que se puede utilizar para purificar el(los) producto(s) adaptador(es) deseado(s) mediante captura en una o más etapas de purificación posteriores, por ejemplo, mediante una fracción de extracción (por ejemplo, estreptavidina) unida a una superficie funcionalizada. En algunas realizaciones que incluyen el enriquecimiento negativo, se puede utilizar un marcador de captura que es o comprende una fracción química fijada (por ejemplo, biotina) para purificar o separar subproductos no deseados de una reacción (por ejemplo, subproductos escindidos unidos a un marcador de captura después del corte del producto preadaptador, intermediarios preadaptadores, hebras no apareadas, etc.) mediante captura en una o más etapas de purificación posteriores, por ejemplo, mediante una fracción de extracción (por ejemplo, estreptavidina) unida a una superficie funcionalizada.

Ejemplo de agente de captura único

La Figura 5 ilustra una realización que incluye enriquecimiento/selección negativa y el uso de un único agente de captura y adaptadores en forma de Y. Específicamente, en la Figura 5, panel A, una hebra de alargamiento y una hebra de plantilla se aparean en una región al menos parcialmente complementaria en una mezcla de reacción para crear un complejo preadaptador. En algunas realizaciones, la hebra de plantilla y la hebra de alargamiento son oligonucleótidos sintetizados. En una realización particular, la hebra de alargamiento puede incluir una secuencia de oligonucleótidos predefinida o conocida. En determinadas realizaciones, la hebra de plantilla puede incluir porciones de secuencia predefinida e incluir una región de secuencia de nucleótidos aleatoria o semialeatoria (por ejemplo, una secuencia degenerada o semidegenerada). De acuerdo con lo anterior, dichas porciones de secuencia pueden ser una secuencia identificadora o SMI. La hebra de plantilla también puede incorporar un marcador de captura, tal como en un extremo 5' en este ejemplo. En un ejemplo particular, se puede sintetizar una hebra de plantilla (por ejemplo, en fragmentos de vidrio de poro controlado (CPG) o similares) en una dirección de 3' a 5' tal como mediante el método de fosforamidita, y se puede ligar una fracción química (por ejemplo, ligada covalentemente, ligada no covalentemente, ligada iónicamente u otra química de enlace) al terminal 5' después de la síntesis del oligonucleótido, o como parte de la síntesis del oligonucleótido, tal como mediante la incorporación de una molécula de fosforamidita no canónica en el terminal 5', cerca del terminal 5' o en una posición interna en el oligonucleótido. En la realización ilustrada en la Figura 5, la hebra de plantilla incluye un SMI en una región media o central de la secuencia. También se muestra, que la hebra de plantilla incluye un marcador de captura en el extremo 5' de la hebra (por ejemplo, biotina). En algunas realizaciones, se incluye un sitio de corte, tal como una secuencia de plantilla para generar un sitio de reconocimiento de endonucleasa de restricción de doble hebra, entre el SMI y un marcador de captura. Dicho sitio de reconocimiento de endonucleasa de restricción de doble hebra se puede cortar enzimáticamente con la endonucleasa de restricción correspondiente para separar el marcador de captura de un complejo adaptador como se describe con más detalle a continuación. Además, cortar en el sitio de corte puede proporcionar al complejo adaptador un extremo ligable en 3' (por ejemplo, un dominio de ligación). De manera similar, un par de sitios de reconocimiento de nucleasa de hebra sencilla (es decir una "nickasa") puede cumplir un propósito funcional similar para escindir la secuencia de estructura principal de las hebras de ADN de doble hebra.

Las enzimas utilizadas para dichos eventos de escisión pueden incluir endonucleasas de restricción, tales como endonucleasas de restricción de tipo I o endonucleasas de restricción de tipo II u otras endonucleasas de restricción, que incluyen nickasas o cualquier endonucleasa modificada que se ha modificada químicamente o mediante ingeniería genética de la estructura del organismo derivado. El sitio de escisión de la endonucleasa se puede reconocer por la naturaleza de la secuencia del polinucleótido (es decir el motivo de reconocimiento) o la endonucleasa podría ser una que reconozca y escinda nucleótidos o componentes modificados o no canónicos de la estructura principal de azúcar-fosfato, directa o indirectamente. La endonucleasa puede implicar sistemas que son específicos de sitios de reconocimiento largos tales como TALEN, meganucleasas/endonucleasas dirigidas, MegaTALEN, nucleasas con dedos de zinc o similares. La endonucleasa puede ser un sistema fácilmente personalizable que implica una secuencia de ARN guía como CRISPER/Cas9 o CPF1 o similar. El sitio de reconocimiento puede incluir tanto una región complementaria de un ARN guía como un motivo PAM. Un experto en la técnica reconocerá que un gran número de enzimas o sistemas enzimáticos similares para escindir directa o indirectamente la estructura principal se podrían aplicar con un propósito similar. También será evidente que combinaciones de enzimas, ribozimas u otros catalizadores de escisión molecular específicos se podrían combinar o sustituir completamente por otros métodos químicos de escisión, tales como ligadores fotoescindibles o similares.

Opcionalmente, la hebra de alargamiento incluye una región de no complementariedad con respecto a la hebra de plantilla que puede funcionar como un SDE durante el análisis de Secuenciación Dúplex. Opcionalmente, la hebra de plantilla contiene una región de no complementariedad con respecto a la hebra de alargamiento que puede funcionar como un SDE durante el análisis de Secuenciación Dúplex. La(s) región(es) de no complementariedad pueden incluir, en algunas realizaciones, un sitio de unión del cebador, una secuencia del cebador leído, un SMI, una secuencia índice u otra secuencia de etiqueta, moléculas modificadas o no nucleótidos, y/o uno o más sitios de corte.

El apareamiento de las hebras de alargamiento y de plantilla se puede llevar a cabo en presencia de un exceso de hebra de plantilla y/o de un exceso de hebra de alargamiento, o utilizando cantidades equimolares tanto de hebra de plantilla como de hebra de alargamiento. En algunas realizaciones, se aprecia que determinada cantidad de hebra de alargamiento residual y/o hebra de plantilla puede permanecer sin hibridar en la mezcla de reacción. La hebra de plantilla residual se denomina Intermedio 1 y la hebra de alargamiento residual se denomina Intermedio 2 en la Figura

5, panel A. En la presente realización, utilizando la síntesis de adaptador en fase de solución, una relación molar entre el exceso de hebra de plantilla y la hebra de alargamiento puede garantizar que la mezcla de reacción comprenda toda o la mayor parte de la hebra de plantilla, y que la hebra de plantilla libre/sin hibridar (Intermedio 1) se puede eliminar, por ejemplo, mediante enriquecimiento/selección negativa como se describe en el presente documento.

Tras el apareamiento de la hebra de alargamiento y la hebra de plantilla para crear un complejo preadaptador (Figura 5, panel A), la hebra de alargamiento se extiende enzimáticamente desde el extremo 3' para producir una porción de doble hebra alargada del complejo preadaptador (Figura 5, panel B). La porción de doble hebra incluye un SMI dúplex (por ejemplo, una secuencia SMI en ambas hebras) y un sitio de corte de doble hebra (por ejemplo, un sitio de reconocimiento de endonucleasa de restricción), creando de esta manera un complejo adaptador de doble hebra no cortado (ver Figura 5, panel B). En algunas realizaciones, se aprecia que algunos de los complejos preadaptadores pueden no extenderse durante esta etapa, o extenderse de manera incompleta, y permanecer en la mezcla de reacción. Esto(s) complejo(s) preadaptador(es) residual(es) se denomina(n) Intermedio 3 en la Figura 5, panel A, y estas entidades no deseadas se pueden eliminar, por ejemplo, mediante enriquecimiento/selección negativa como se describe en el presente documento.

A continuación, como se muestra en la Figura 5, panel C, el complejo adaptador de doble hebra no cortado se trata con una enzima apropiada para interactuar con el complejo preadaptador en el sitio del corte. Por ejemplo, se puede proporcionar a la mezcla de reacción una enzima endonucleasa de restricción que corta un sitio de corte de reconocimiento de endonucleasa de restricción (por ejemplo, dentro de un dominio de ligación 3' con respecto al SMI) para producir un extremo ligable 3' en el producto adaptador deseado. En la Figura 5, panel C, el extremo ligable se muestra como un saliente en T, sin embargo, será evidente para un experto en la técnica que el extremo ligable puede tener cualquiera de una variedad de formas, por ejemplo, un extremo romo, un saliente en A-3', un extremo "pegajoso" que comprende un saliente 3' de un nucleótido, un saliente 3' de dos nucleótidos, un saliente 3' de tres nucleótidos, un saliente 3' de 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleótidos, un saliente 3' de un nucleótido, un saliente 5' de dos nucleótidos, un saliente 5' de tres nucleótidos, un saliente 5' de 4, 5, 6, 7, 8, 9, 10, 11, 12, 13, 14, 15, 16, 17, 18, 19, 20 o más nucleótidos, entre otros. La base 5' del sitio de ligación puede estar fosforilada y la base 3' puede tener un grupo hidroxilo, o puede estar, sola o en combinación, desfosforilada o deshidratada o modificada químicamente adicionalmente para facilitar una ligación potenciada o una hebra para evitar la ligación de una hebra, opcionalmente, hasta un momento posterior. El corte del dominio de ligación también produce un fragmento escindido, denominado subproducto en la Figura 5, panel C. La temperatura de fusión de este subproducto puede ser baja y el fragmento escindido puede fundirse bajo determinadas condiciones de temperatura y sal, o puede permanecer apareado bajo determinadas condiciones de temperatura y sal. De esta manera, el producto adaptador deseado se separa de (por ejemplo, se libera de) el marcador de captura. Dicho de otra manera, como el fragmento escindido contiene el marcador de captura fijado, el producto adaptador deseado ya no incluye un marcador de captura después de la etapa de corte.

En algunas realizaciones, se aprecia que algunos de los complejos preadaptadores no se pueden cortar durante esta etapa y permanecer en la mezcla de reacción. Esto(s) complejo(s) preadaptador(es) residual(es) no cortados se denominan Intermedio 4 en la Figura 5, panel B, y estas entidades no deseadas se pueden eliminar, por ejemplo, mediante enriquecimiento/selección negativa como se describe en el presente documento.

Postcorte, como se muestra en la Figura 5, panel D, la mezcla de reacción se pone en contacto con una superficie funcionalizada con una o más fracciones de extracción unidas a la misma. Las fracciones de extracción proporcionadas son capaces de unirse al marcador de captura (por ejemplo, una perla de estreptavidina donde el marcador de captura es biotina) para la inmovilización y separación de moléculas que llevan el marcador de captura. En particular, la fracción de extracción puede ser cualquier miembro de un par de unión, tal como biotina/estreptavidina o hapteno/anticuerpo o secuencias de ácidos nucleicos complementarias (par ADN/ADN, par ADN/ARN, par ARN/ARN, par LNA/ADN, etc.). En la realización ilustrada, un marcador de captura que está adherido a productos intermedios y subproductos del proceso de fabricación del adaptador es capturado por su par de unión (por ejemplo, la fracción de extracción) que está adherido a una fracción aislable (por ejemplo, tal como una partícula atraíble magnéticamente o una partícula grande que se puede sedimentar mediante centrifugación). De acuerdo con lo anterior, el marcador de captura puede ser cualquier tipo de molécula/fracción que permita la separación por afinidad de los ácidos nucleicos que llevan el marcador de captura de los ácidos nucleicos que carecen del marcador de captura. Un ejemplo de un marcador de captura es la biotina que permite la separación por afinidad mediante la unión a estreptavidina ligada o ligable a una fase sólida o un oligonucleótido, que a su vez permite la separación por afinidad a través de la unión a un oligonucleótido complementario ligado o ligable a una fase sólida.

Como se ilustra en la Figura 5, panel D, los intermedios no deseados (por ejemplo, Intermedio 1, Intermedio 3, Intermedio 4) y subproductos que contienen el marcador de captura se unen a la superficie funcionalizada y el producto adaptador deseado se deja libre en solución. De acuerdo con diversas realizaciones, los intermedios y subproductos no deseados pueden ser o comprender cualquiera o todos de: fragmentos escindidos, exceso de hebras de plantilla, complejos preadaptadores no extendidos o incompletamente extendidos, y/o complejos adaptadores no cortados o solo parcialmente cortados y también pueden incluir oligonucleótidos constituyentes que se sintetizaron químicamente de forma incompleta o incorrecta antes de la ligación, tales como productos de longitud incompleta, productos con nucleótidos dañados, productos con inserciones o supresiones relativas a la secuencia de síntesis deseada, productos que se han ligado o entrecruzado erróneamente con otros productos u otras fracciones químicas tales como enzimas

u otros tipos de moléculas o superficies, productos con aductos de nucleótidos no deseados, aquellos con daños tales como oxidación, desaminación, fosforilación, desfosforilación, sumoilación, glicosilación, desglicosilación, putrescinilación, carboxilación, halogenación, formilación, daño por hidrólisis, daño por nucleasa, fotodaño tal como los dímeros de pirimidina, o con cualquier otra forma de síntesis o daño incompleto o erróneo posterior, o cualesquier subproductos, productos intermedios, productos secundarios tales como los enumerados anteriormente, que contienen uno o más de dichos oligonucleótidos con propiedades incompletas o síntesis errónea o daño a los mismos posteriormente. De manera similar, los productos no deseados que contienen errores (tales como emparejamientos erróneos inserciones o supresiones) introducidos por la extensión de la polimerasa, nucleótidos naturales o dañados mal incorporados o como consecuencia de una actividad transferasa terminal o exonucleasa no intencionada, pueden todos ellos implicar productos no deseados y contribuir a complejos preadaptadores no extendidos o incompletamente extendidos y/o complejos adaptadores no cortados o sólo parcialmente cortados.

En algunas realizaciones, los métodos proporcionados permiten la eliminación de todos o sustancialmente todos los intermedios y/o subproductos no deseados o reducen sustancialmente su abundancia. La recolección de(los) producto(s) adaptador(es) deseado(s) se puede realizar de cualquier manera apropiada para la aplicación. A modo de ejemplo específico, en algunas realizaciones, la recolección de de(los) producto(s) adaptador(es) deseado(s) se puede lograr mediante una o más de la eliminación de la superficie funcionalizada mediante filtración por tamaño, métodos magnéticos, métodos de carga eléctrica, métodos de densidad de centrifugación o cualesquier otros métodos o, recolección de fracciones de elución si se utilizan métodos de purificación basados en columnas o similares, o mediante cualquier otra práctica de purificación comúnmente entendida.

Ejemplo de agente de captura múltiple

En algunas realizaciones, un método puede incluir el uso de dos o más marcadores de captura (por ejemplo, 2, 3, 4, 5, 6, 7, 8, 9, 10 o más). En algunas realizaciones, al menos dos marcadores de captura utilizados en un método son diferentes entre sí (por ejemplo, una molécula pequeña y un péptido). En algunas realizaciones, la inclusión de dos o más marcadores de captura diferentes permite el uso tanto de enriquecimiento/selección positiva como de enriquecimiento/selección negativa. La inclusión de dos o más marcadores de captura puede resultar útil, entre otros, en los casos en los que existe un deseo de controlar los casos en los que un marcador de captura no se incorpora involuntariamente a la oligosíntesis original de la hebra de plantilla y/o cuando existe la preocupación de que moléculas no marcadas (por ejemplo, hebra de alargamiento) puedan pasar al producto adaptador recolectado final.

En los casos en los que sea apropiada una purificación adicional de un producto adaptador deseado, se puede combinar un esquema de enriquecimiento negativo con un esquema de enriquecimiento positivo al utilizar un método de etiqueta de doble afinidad, por ejemplo, como se ejemplifica en la Figura 6. En particular para los métodos que incluyen dos o más agentes de captura, una hebra de plantilla también puede incluir un marcador de captura en un extremo 3' de la hebra junto con el marcador de captura en el extremo 5' como se muestra en la Figura 6. El marcador de captura en el extremo 3' de la hebra de plantilla se puede ligar a la hebra de plantilla mediante una fracción escindible opcional. Una fracción escindible puede ser un grupo uracilo, como se muestra en la Figura 6, o cualquier otro grupo escindible enzimática, química o fotoeléctricamente.

Como se muestra en la Figura 6, panel A, y similar a algunas realizaciones que incluyen un único marcador de captura (ver Figura 5), una hebra de alargamiento y una hebra de plantilla se aparean en una región al menos parcialmente complementaria en una mezcla de reacción para crear un complejo preadaptador. La hebra de alargamiento y/o la hebra de plantilla pueden tener regiones de no complementariedad que, como se discutió anteriormente, pueden incluir sitio(s) de unión de cebador, una secuencia de cebador leída, un SMI, una secuencia índice u otra secuencia de etiqueta, moléculas modificadas o no nucleótidos, y/o uno o más sitios de corte. En algunas realizaciones, la región de no complementariedad puede funcionar como un SDE.

En la realización ilustrada en la Figura 6, la hebra de plantilla incluye un SMI en una región media o central de la secuencia. También se muestra, la hebra de plantilla incluye un primer marcador de captura en el extremo 5' de la hebra (por ejemplo, biotina) y un segundo marcador de captura, que puede ser igual o diferente del primer marcador de captura, en un extremo 3' de la hebra. En algunas realizaciones, primero se incluye un sitio de corte, tal como una secuencia de plantilla para generar un sitio de reconocimiento de endonucleasa de restricción de doble hebra, entre el SMI y el primer marcador de captura. Adicionalmente, un segundo sitio de corte, tal como una molécula de nucleótido o no nucleótido modificada, se coloca en 5' con respecto al segundo marcador de captura mediante un ligamiento escindible de tal manera que el segundo marcador de captura sea un marcador de captura escindible. En algunas realizaciones, el ligamiento escindible es un ligamiento fotoescindible. En algunas realizaciones, el ligamiento escindible comprende un nucleótido de uracilo. En otra realización, los segundos marcadores de captura comprenden biotina ligada a hebras de ácido nucleico que incluyen uno o más residuos modificados. En una realización, un nucleótido modificado puede comprender un nucleótido de ribosa. En aún otra realización, el segundo sitio de corte puede ser un residuo de destiobiotina-TEG fotoescindible o de uracilo escindible con una combinación de uracilo ADN glicosilasa (UDG) y una enzima con actividad ADN liasa de sitio abásico tal como endonucleasa VIII o formamidopirimidina [fapy]-ADN glicosilasa (FPG) o combinaciones premezcladas comerciales (por ejemplo enzima USER™. En determinadas realizaciones, la una o más moléculas de nucleótidos o no nucleótidos modificadas pueden ser un sitio abásico, un uracilo, tetrahidrofurano, 8-oxo-7,8-dihidro-2'-desoxiadenosina (8-oxo-A), 8-oxo-7,8-dihidro-

2'-desoxiguanosina (8-oxo-G), desoxiinosina, 5'-nitroindol, 5-Hidroximetil-2'-desoxicitidina, iso-citosina, 5'-metil-isocitosina o iso-guanosina.

Al igual que con las realizaciones que incluyen un único marcador de captura (ver Figura 5), el apareamiento de una hebra de alargamiento y una hebra de plantilla se puede llevar a cabo en presencia de un exceso de hebra de plantilla y/o un exceso de hebra de alargamiento, o utilizando cantidades equimolares tanto de la hebra de plantilla como de la hebra de alargamiento. En la presente realización, utilizando la síntesis de adaptador en fase de solución con etapas de enriquecimiento/selección negativa y de enriquecimiento/selección positiva, el exceso de hebra de plantilla (Intermedio 1) y el exceso de hebra de alargamiento (Intermedio 2) se pueden eliminar durante el procesamiento. La hebra de plantilla libre/sin hibridar (Intermedio 1) se puede eliminar, por ejemplo, mediante enriquecimiento/selección negativa y la hebra de alargamiento libre/sin hibridar (Intermedio 2) se puede eliminar, por ejemplo, mediante enriquecimiento/selección positiva como se describe en el presente documento.

La Figura 6, el panel B muestra que el complejo preadaptador se puede extender enzimáticamente, y la Figura 6, panel C muestra que el complejo adaptador no cortado se trata enzimáticamente (por ejemplo, con una enzima endonucleasa de restricción) para cortar en el sitio de corte. Como se muestra en la Figura 6, panel C, la etapa de corte produce ambos complejos adaptadores cortados que tienen un segundo marcador de captura en 3' con respecto a la región de doble hebra que comprende el SMI y, opcionalmente, ligado al extremo 3' de la hebra de plantilla mediante un ligador escindible, y fragmentos escindidos que incluyen el primer marcador de captura. Como se indica en la Figura 6 paneles de A-C, los Intermedios 1, 3 y 4, y los subproductos de los fragmentos de escisión también se forman durante este proceso sustancialmente como se describe para algunas realizaciones de un único agente de captura como se describe en la Figura 5.

A continuación se puede realizar una etapa de enriquecimiento/selección negativa, por ejemplo, mediante la introducción de la mezcla de reacción en una superficie funcionalizada (S_1) que es capaz de unir el primer marcador de captura ubicada en 5' (y, en la realización ilustrada, el segundo marcador de captura no ubicado en 3') (Figura 6, panel D). La interacción de unión entre el primer marcador de captura y la superficie funcionalizada (S_1) a través de una fracción de extracción/compañero de unión asociado inmovilizará los intermedios y subproductos no deseados que incluyen el primer marcador de captura en posición 5', mientras que el producto adaptador deseado con el segundo marcador de captura en posición 3' permanecerá libre en la solución (Figura 6, panel D). La eliminación de la superficie funcionalizada (S_1) proporciona una primera etapa de purificación del adaptador.

En una etapa de enriquecimiento/selección positiva, una superficie funcionalizada (S_2) capaz de unirse al segundo marcador de captura en posición 3' se introduce en la mezcla de reacción. La interacción de unión entre el segundo marcador de captura y la superficie funcionalizada (S_2) a través de una fracción de extracción/compañero de unión asociado inmovilizará los productos adaptadores deseados. Una vez que se inmovilizan los productos adaptadores deseados, la mezcla de reacción se puede lavar y/o los productos adaptadores inmovilizados se pueden separar de la mezcla de reacción restante. Ventajosamente, la hebra de alargamiento libre/no hibridada (Intermedio 2), que no lleva ningún marcador de captura, se puede eliminar/separar eficazmente del producto adaptador deseado. En realizaciones adicionales, la superficie funcionalizada (S_2) se puede lavar para eliminar intermedios residuales, subproductos u otros contaminantes. En las etapas adicionales, la mezcla de reacción que comprende la superficie funcionalizada (S_2) con productos adaptadores unidos se puede tratar con un reactivo (por ejemplo, una enzima) que escinde y separa el marcador de captura en posición 3' (escindible) del producto adaptador deseado, liberando de esta manera el producto adaptador deseado en solución. Por ejemplo, una mezcla de reacción que comprende la superficie funcionalizada (S_2) se puede tratar con la enzima USER (Reactivo de Escisión Específico de Uracilo) (una combinación comercial de UDG y Endonucleasa VIII), o cualquier otro tratamiento que pueda escindir en un grupo ligador incorporado (por ejemplo, un grupo uracilo como se muestra en la Figura 6 o RNaseH2 que escinde la estructura principal de fosfato de azúcar en las bases de ribosa). El producto adaptador deseado luego se puede recoger como eluato o separar al eliminar la superficie funcionalizada (secundaria) (S_2).

Si bien las realizaciones que comprenden un primer marcador de captura y un segundo marcador de captura (escindible) que es diferente del primer marcador de captura se describen con referencia a la Figura 6, los expertos en la técnica entenderán que el primer marcador de captura y el segundo marcador de captura pueden ser el mismo marcador de captura. En tales casos, se puede introducir una superficie funcionalizada que lleva una fracción de extracción que es el compañero de unión para el primer y segundo marcadores de captura después de la etapa de corte. El primer y segundo marcadores de captura se inmovilizarán mediante la interacción entre los marcadores de captura y la(s) superficie(s) funcionalizada(s) (no mostradas). Después de la inmovilización, se puede dirigir el primer sitio de corte (por ejemplo, mediante enzima de restricción) de tal manera que las moléculas preadaptadoras de doble hebra se puedan cortar y liberar de la inmovilización del primer marcador de captura. En este estadio, los intermedios y subproductos no deseados permanecen unidos a la superficie funcionalizada. Adicionalmente, en este estadio, los productos adaptadores deseados permanecen unidos a la superficie funcionalizada mediante el segundo marcador de captura. Las moléculas libres o no unidas que no llevan las primeras o segundas marcadores de captura se pueden eliminar o separar de otro modo de aquellas que llevan los marcadores de captura. En una siguiente etapa, se puede dirigir la escisión de un ligador escindible próximo a o 5' al segundo marcador de captura para iniciar la liberación del producto adaptador deseado en la solución. La eliminación y/o separación de la superficie funcionalizada del producto deseado no unido en solución proporciona un producto adaptador purificado tanto

enriquecido/seleccionado negativamente (eliminación de intermedios y subproductos no deseados) como enriquecido/seleccionado positivamente.

Al igual que con algunas realizaciones que incluyen un único agente de captura, algunas realizaciones que incluyen dos o más agentes de captura permiten la eliminación de intermedios y subproductos no deseados que pueden ser o comprender cualquiera o todos de: fragmentos escindidos, exceso de hebras de plantilla, exceso de hebras de alargamiento, complejos de preadaptador no extendidos o incompletamente extendidos, y/o complejos de adaptador no cortado o sólo parcialmente cortados u otros enumerados anteriormente. En algunas realizaciones, los métodos proporcionados permiten la eliminación de todos o sustancialmente todos los intermedios y/o subproductos no deseados para proporcionar un producto de adaptador dúplex purificado.

10 Aspectos adicionales

Tipos de adaptador

Si bien la mayoría de los ejemplos en la presente divulgación representan adaptadores en forma de Y o de bucle, se puede utilizar cualquier estructura de adaptador conocida de acuerdo con diversas realizaciones, tales como la descrita en el documento WO2017/100441, .

15 Por ejemplo, se contemplan además diversas formas de adaptadores que comprenden burbujas (por ejemplo, regiones internas de no complementariedad).

Separación

Como se describe en el presente documento, diversos métodos incluyen al menos una etapa de separación. Se contempla específicamente que se pueda incluir cualquiera de una variedad de etapas de separación en diversas realizaciones. Por ejemplo, en algunas realizaciones, la separación puede ser o comprender separación física, separación por tamaño, separación magnética, separación por solubilidad, separación de carga, separación por hidrofobicidad, separación por polaridad, separación por movilidad electroforética, separación por densidad, separación por elución química, separación de perlas SBIR, etc. Por ejemplo, un grupo físico puede tener una propiedad magnética, una propiedad de carga o una propiedad de insolubilidad. En realizaciones, cuando el grupo físico tiene una propiedad magnética y se aplica un campo magnético, las secuencias de ácidos nucleicos adaptadoras asociadas que incluyen el grupo físico se separan de las secuencias de ácidos nucleicos adaptadoras que no incluyen el grupo físico. En otra realización, cuando el grupo físico tiene una propiedad de carga y se aplica un campo eléctrico, las secuencias de ácidos nucleicos adaptadoras asociadas que incluyen el grupo físico se separan de la secuencia de ácidos nucleicos adaptadora que no incluye el grupo físico. En las realizaciones, cuando el grupo físico tiene una propiedad de insolubilidad y las secuencias de ácidos nucleicos adaptadoras están contenidas en una solución para la cual el grupo físico es insoluble, las secuencias de ácidos nucleicos adaptadoras que comprenden el grupo físico se precipitan lejos de la secuencia de ácidos nucleicos adaptador que no incluyen el grupo físico que permanece en solución.

En diversas realizaciones se puede incluir cualquiera de una variedad de métodos de separación física. A modo de ejemplo específico, un conjunto no limitante de métodos incluye: filtración selectiva por tamaño, centrifugación por densidad, separación por HPLC, separación por filtración en gel, selección de cada perla con sus diminutas pinzas una por una, separación por FPLC, centrifugación en gradiente de densidad y cromatografía en gel, entre otros.

En diversas realizaciones se puede incluir cualquiera de una variedad de métodos de separación magnética. Normalmente, los métodos de separación magnética abarcarán la inclusión o adición de uno o más grupos físicos que tienen una propiedad magnética tal que, cuando se aplica un campo magnético, las moléculas que incluyen dicho(s) grupo(s) físico(s) se separan de las que no. A modo de ejemplo específico, los grupos físicos que exhiben una propiedad magnética incluyen, pero no se limitan a, materiales ferromagnéticos tales como hierro, níquel, cobalto, disprosio, gadolinio y aleaciones de los mismos. Las perlas paramagnéticas comúnmente utilizadas para la separación química y bioquímica incorporan dichos materiales dentro de una superficie que reduce la interacción química de los materiales con los productos químicos que se manipulan, tales como el poliestireno, que se puede funcionalizar para las propiedades de afinidad discutidas anteriormente.

Extensión

Con respecto a la extensión, por ejemplo, de una hebra de alargamiento, se puede utilizar cualquier método conocido. En algunas realizaciones en las que se utilizan una o más enzimas, dicha(s) enzima(s) se pueden seleccionar de ADN polimerasas con diversos grados de actividad de desplazamiento de hebra que extienden secuencias de ácidos nucleicos desde un grupo 3'OH en una dirección de 5' a 3'. Ejemplos de ADN polimerasas incluyen T4, T7, phi29, ADN polimerasa Bst, polimerasa Taq, polimerasa Pol I, y muchas otras. En general, se pueden utilizar polimerasas que carecen de actividad de desplazamiento de hebras en las reacciones de llenado de espacios que se describen más adelante en el presente documento. Los ejemplos no limitantes de polimerasas que no desplazan hebras disponibles comercialmente incluyen ADN polimerasas T4 y T7.

Etiquetas de captura

Como se describe en el presente documento, en algunas realizaciones, un marcador de captura puede estar presente en cualquiera de una variedad de configuraciones a lo largo de una hebra de plantilla. Por ejemplo, en algunas realizaciones, un marcador de captura puede estar presente en los extremos 5', los extremos 3', o ambos extremos de una hebra de plantilla. En otras realizaciones, se puede incorporar o fijar un marcador de captura a una hebra de

5 plantilla en una región 5' de la secuencia identificadora. En algunas realizaciones, un marcador de captura puede estar presente en algún lugar en el medio de una hebra de plantilla (es decir, no en el extremo 5' o 3' de la hebra de plantilla). En realizaciones que incluyen dos o más marcadores de captura, cada marcador de captura puede estar presente en una ubicación diferente a lo largo de la hebra de plantilla.

En algunas realizaciones, un marcador de captura se selecciona de un grupo de biotina, biotina desoxitimidina dT, biotina NHS, biotina TEG, Biotina-6-Aminoalil-2'-desoxiuridina-5'-Trifosfato, Biotina-16-Aminoalil-2'-desoxicitidina-5'-Trifosfato, Biotina-16-Aminoalil-2'-desoxicitidina-5'-Trifosfato, N4-Biotina-OBEA-2'-desoxicitidina-5'-Trifosfato, Biotina-16-Aminoaliluridina-5'-Trifosfato, Biotina-16-7-Deaza-7-Aminoalil-2'-desoxiguanosina-5'-Trifosfato, 5'-Biotina-G-Monofosfato, 5'-Biotina-A-Monofosfato, 5'-Biotina-dG-Monofosfato, 5'-Biotina-dA-Monofosfato, destiobiotina NHS, destiobiotina-6-Aminoalil-2'-desoxicitidina-5'-trifosfato, digoxigenina NHS, DNP TEG, tioles, Colicina E2, Im2, glutatión, glutatión-s-transferasa (GST), níquel, polihistidina, etiqueta FLAG, etiqueta myc, entre otros. En algunas realizaciones, los marcadores de captura incluyen, sin limitación, biotina, avidina, estreptavidina, un hapteno reconocido por un anticuerpo, una secuencia de ácidos nucleicos particular y/o una partícula atraíble magnéticamente. En algunas realizaciones, una o más modificaciones químicas de moléculas de ácido nucleico (por ejemplo, Acridite™ modificado entre muchas otras modificaciones, algunas de las cuales se describen en otra parte de la solicitud) puede servir como

10
15
20

Fracciones de extracción

Las fracciones de extracción pueden ser un compañero o par de unión física para el marcador de captura diana y se refiere a una fracción aislable o cualquier tipo de molécula que permita la separación por afinidad de los ácidos nucleicos que llevan el marcador de captura de los ácidos nucleicos que carecen del marcador de captura. Las fracciones de extracción se pueden ligar directamente o indirectamente (por ejemplo, mediante ácido nucleico, mediante anticuerpo, mediante aptámero, etc.) a un sustrato, tal como una superficie sólida. En algunas realizaciones, la fracción de extracción se selecciona de un grupo que comprende una molécula pequeña, un ácido nucleico, un péptido, un anticuerpo o cualquier fracción que se pueda unir de forma única. La fracción de extracción puede estar ligada o ser ligable a una fase sólida u otra superficie para formar una superficie funcionalizada. En algunas realizaciones, la fracción de extracción es una secuencia de nucleótidos ligados a una superficie (por ejemplo, una superficie sólida, una perla, una partícula magnética, etc.). En algunas realizaciones, en las que el marcador de captura es biotina, la fracción de extracción se selecciona de un grupo de avidina o estreptavidina. Un experto en la técnica apreciará que se puede utilizar cualquiera de una variedad de pares de unión por afinidad de acuerdo con diversas realizaciones.

25
30

En determinadas realizaciones, las fracciones de extracción pueden ser propiedades físicas o químicas que interactúan con el marcador de captura dirigido. Por ejemplo, una fracción de extracción puede ser un campo magnético, un campo de carga o una solución líquida en la que un marcador de captura dirigida es insoluble. Dichas propiedades físicas o químicas se pueden aplicar y los ácidos nucleicos adaptadores que llevan el marcador de captura se pueden inmovilizar dentro/contra un recipiente (superficie) o columna. Dependiendo del resultado deseado de enriquecimiento/selección negativo o enriquecimiento/selección positivo, las moléculas no inmovilizadas se pueden retener (enriquecimiento negativo) o las moléculas inmovilizadas se pueden retener (enriquecimiento positivo) para purificación/procesamiento o uso adicional.

35
40

Superficies sólidas

Cuando el compañero de afinidad/fracción de extracción se une a una superficie sólida o sustrato y se une al marcador de captura, las secuencias de ácidos nucleicos adaptadoras que incluyen el marcador de captura se pueden separar de la secuencia de ácidos nucleicos adaptadoras que no incluyen el marcador de afinidad. Una superficie o sustrato sólido puede ser una perla, una partícula aislable, una partícula magnética u otra estructura fija.

45

Como se describe en el presente documento y será apreciado por un experto en la técnica, se puede utilizar cualquiera de una variedad de superficies funcionalizadas de acuerdo con diversas realizaciones. Por ejemplo, en algunas realizaciones, una superficie funcionalizada puede ser o comprender una perla (por ejemplo, una perla de vidrio de poro controlado, una perla de poliestireno macroporoso, etc.). Sin embargo, un experto en la técnica entenderá que muchos otros pares químicos de fracción/superficie podrían utilizarse de manera similar para lograr el mismo propósito. Se entenderá que las superficies funcionalizadas específicas descritas en el presente documento se entienden sólo como ejemplos, y que se puede utilizar cualquier otra estructura o sustrato fijo apropiado capaz de asociarse (por ejemplo, ligarse, unirse, etc.) con una o más fracciones de extracción.

50
55

Corte de ácidos nucleicos

Varios aspectos de la presente tecnología, que incluye la síntesis, fabricación, procesamiento y purificación de adaptadores dúplex, incorporan escisión enzimática, escisión enzimática de una hebra, escisión enzimática de ambas

hebras, incorporación de un ácido nucleico modificado seguido de un tratamiento enzimático que conduce a la escisión o una o ambas hebras, incorporación de un ligador fotoescindible, incorporación de un uracilo, incorporación de una base ribosa, incorporación de un aducto de 8-oxo-guanina, uso de una endonucleasa de restricción, uso de enzimas de corte dirigidas al sitio, y similares. En otras realizaciones, se pueden utilizar las endonucleasas, tales como una endonucleasa de ribonucleoproteína (por ejemplo, una enzima Cas, tal como Cas9 o CPF1), u otra endonucleasa programable (por ejemplo, una endonucleasa de búsqueda, una nucleasa con dedos de zinc, una TALEN, una meganucleasa (por ejemplo, nucleasa megaTAL), una argonauta nucleasa, etc.), y cualquier combinación de las mismas.

Como se describe en el presente documento, diversas realizaciones incluyen el uso de una o más endonucleasas que reconocen secuencias o modificaciones de nucleótidos únicas u otras entidades que reconocen bases u otras modificaciones químicas de la estructura principal para cortar y/o escindir un ácido nucleico de doble hebra (por ejemplo, ADN o ARN) en una ubicación específica en una o varias hebras. Los ejemplos incluyen Uracilo (reconocido y que se puede escindir con una combinación de Uracilo ADN glicosilasa y una liasa de sitio abásico tal como la Endonucleasa VIII o FPG, y nucleótidos de ribosa, que se pueden reconocer y escindir por la ARNasaH2 cuando se emparejan con una base de ADN. El ácido nucleico puede ser ADN, ARN o una combinación de los mismos y, opcionalmente, incluir un ácido nucleico peptídico (PNA) o un ácido nucleico bloqueado (LNA) u otro ácido nucleico modificado. En algunas realizaciones, el corte se puede realizar mediante el uso de una o más endonucleasas de restricción. En algunas realizaciones, la escisión se puede realizar utilizando un ligador escindible, por ejemplo, uracilo destiobotina-TEG, escisión de ribosa, u otros métodos. En algunas realizaciones, el ligado escindible puede ser un ligador fotoescindible o un ligador escindible químicamente que no requiere enzimas, o parcialmente.

En algunas realizaciones, por ejemplo, cuando se utiliza una endonucleasa de restricción para cortar un ácido nucleico de doble hebra, se pueden incluir una o más secuencias plantilla de reconocimiento de endonucleasa de restricción en una hebra de plantilla. Cuando se extiende la hebra de alargamiento, la secuencia de reconocimiento de la endonucleasa de restricción de doble hebra resultante es reconocible por su correspondiente endonucleasa de restricción. En algunas realizaciones, se agrega una secuencia de plantilla de reconocimiento de endonucleasa de restricción antes del inicio de una etapa de apareamiento. En algunas realizaciones, se agrega una secuencia de plantilla de reconocimiento de endonucleasa de restricción después de una o más etapas de extensión, corte o eliminación como se describe en el presente documento. En algunas realizaciones, una secuencia de reconocimiento de endonucleasa de restricción se puede generar o generarse en 3' de un SMI. En algunas realizaciones, una secuencia de reconocimiento de endonucleasa de restricción se puede generar o generarse en 5' de un SMI. En algunas realizaciones, una secuencia de reconocimiento de endonucleasa de restricción puede estar en 3' de un SDE. En algunas realizaciones, una secuencia de reconocimiento de endonucleasa de restricción puede estar en 5' de un SDE. En algunas realizaciones, una secuencia de reconocimiento de endonucleasa de restricción puede estar en 3' de un marcador de captura. En algunas realizaciones, una secuencia de reconocimiento de endonucleasa de restricción puede estar en 5' de un marcador de captura. En algunas realizaciones, una secuencia de reconocimiento de endonucleasa de restricción puede estar dentro o proximal a un dominio de ligación. En algunas realizaciones, se puede proporcionar una secuencia de reconocimiento de endonucleasa de restricción de tal manera que el tratamiento posterior con una endonucleasa de restricción correspondiente genere un extremo ligable 3' del adaptador dúplex.

En algunas realizaciones, una hebra de plantilla incluirá 2 o más (por ejemplo, 3, 4, 5, 6, 7, 8, 9, 10 o más) sitios de reconocimiento de endonucleasas de restricción. Un experto en la técnica apreciará que una variedad de endonucleasas de restricción (es decir, enzimas de restricción) que escinden el ADN en o cerca de sitios de reconocimiento (por ejemplo, EcoRI, BamHI, XbaI, HindIII, AluI, Avail, BsaJI, BstNI, DsaV, Fnu4HI, HaeIII, MaeIII, N1aIV, NSI, MspJI, FspEI, NaeI, Bsu36I, NotI, HinfI, Sau3AI, PvuII, SmaI, HgaI, AluI, EcoRV, etc.) pueden estar de acuerdo con diversas realizaciones de la presente tecnología. Los listados de varias endonucleasas de restricción están disponibles tanto en forma impresa como legible por ordenador, y son proporcionados por muchos proveedores comerciales (por ejemplo, New England Biolabs, Ipswich, MA). Se puede encontrar una lista no limitante de endonucleasas de restricción y sitios de reconocimiento asociados en: www.neb.com/tools-and-resources/selection-charts/alphabetized-list-of-recognition-specificities.

En algunas realizaciones, los nucleótidos modificados o no pueden proporcionar una fracción escindible. Por ejemplo, bases de uracilo (se pueden escindir con una combinación de UGD y endonucleasa VIII o FPG como un ejemplo), sitios abásicos (se pueden escindir con la endonucleasa VIII como un ejemplo), 8-oxoguanina (se pueden escindir mediante FPG u OGG1 como ejemplos) y nucleótidos de ribosa (se pueden escindir mediante ARNasaH2 cuando se empareja con ADN en un ejemplo) como se describió anteriormente.

Extremos ligables

En algunas realizaciones, los productos adaptadores se generan con un extremo 3' ligable adecuado para la ligación para dirigirse a secuencias de ácidos nucleicos de doble hebra (por ejemplo, para la preparación de bibliotecas de secuenciación). En algunas realizaciones, una hebra de plantilla comprende un dominio de ligación 3' que comprende una secuencia de plantilla para generar un sitio de corte de doble hebra. Dicho sitio de corte puede ser una secuencia de reconocimiento de endonucleasa de restricción o, en otra realización, pueden ser nucleótidos no estándar escindibles. Después de la extensión de la hebra de alargamiento y el procesamiento adicional en el(los) sitio(s) de corte, los dominios de ligamiento resultantes presentes en cada uno de los productos adaptadores de doble hebra

pueden ser capaces de ligarse a una hebra correspondiente de una secuencia de ácidos nucleicos diana de doble hebra. En algunas realizaciones, uno de los dominios de ligación incluye un saliente en T, un saliente en A, un saliente en CG, un extremo romo u otra secuencia de ácidos nucleicos ligable. En algunas realizaciones, un dominio de ligación 3' de doble hebra comprende un extremo romo. En determinadas realizaciones, al menos una de las secuencias del dominio de ligación incluye un ácido nucleico modificado o no estándar. En algunas realizaciones, un nucleótido modificado puede ser un sitio abásico, un uracilo, tetrahidrofurano, 8-oxo-7,8-dihidro-2'-desoxiadenosina (8-oxo-A), 8-oxo-7,8-dihidro-2'-desoxiguanosina (8-oxo-G), desoxiinosina, 5'-nitroindol, 5-hidroximetil-2'-desoxicitidina, isocitosina, 5'-metil-isocitosina o iso-guanosina. En algunas realizaciones, al menos una hebra del dominio de ligación incluye una base desfosforilada. En algunas realizaciones, al menos uno de los dominios de ligación incluye una base deshidroxilada. En algunas realizaciones, al menos una hebra del dominio de ligación se ha modificado químicamente para hacerla no ligable (por ejemplo, hasta que se realice una acción adicional para hacer que el dominio de ligación sea ligable). En algunas realizaciones se obtiene un saliente 3' mediante el uso de una polimerasa con actividad transferasa terminal. En un ejemplo, la Taq polimerasa puede agregar un solo par de bases sobresalientes. En algunas realizaciones esto es una "A".

15 Nucleótidos no estándar

En algunas realizaciones, la plantilla y/o las hebras de alargamiento proporcionadas pueden incluir uno o más nucleótidos no estándar/no canónicos. En algunas realizaciones, un nucleótido no estándar puede ser o comprender un uracilo, un nucleótido metilado, un nucleótido de ARN, un nucleótido de ribosa, una 8-oxoguanina, un nucleótido biotinilado, un nucleótido de destiobiotina, un nucleótido modificado con tiol, un nucleótido modificado con acridita, un iso-dC, un iso dG, un nucleótido 2'-O-metilo, un nucleótido de inosina, un Ácido Nucleico Bloqueado, un ácido peptídico nucleico, un 5 metil dC, una 5-bromodesoxiuridina, una 2,6-diaminopurina, nucleótido de 2-Aminopurina, un nucleótido abásico, un nucleótido de 5-Nitroindol, un nucleótido adenilado, un nucleótido de azida, un nucleótido de digoxigenina, un ligador I, un nucleótido modificado con 5'hexinilo, un 5-Octadiinil dU, espaciador fotoescindible, un espaciador no fotoescindible, un nucleótido modificado compatible con la química de clic, un tinte fluorescente, biotina, furano, BrdU, Fluoro-dU, Ioto-dU y cualquier combinación de los mismos.

B. Diversas realizaciones de métodos y reactivos para la síntesis de adaptadores en fase sólida

La presente divulgación también proporciona métodos y reactivos para la síntesis de adaptadores en fase sólida. En algunas realizaciones, se pueden utilizar uno o más marcadores de captura para el enriquecimiento/selección del(los) producto(s) adaptador(es) deseado(s), por ejemplo, mediante enriquecimiento/selección positiva o enriquecimiento/selección negativa, similar a los métodos utilizados en los métodos de síntesis de adaptadores basados en solución.

En algunas realizaciones, los métodos que incluyen un proceso de síntesis en fase sólida incluyen una etapa inicial de asociar una hebra de plantilla con una superficie funcionalizada. En algunas realizaciones, una hebra de plantilla incluirá uno o más marcadores de captura, y la superficie funcionalizada incluirá una o más fracciones de extracción capaces de unirse o asociarse de otro modo con uno o más marcadores de captura en la hebra de plantilla en presencia o ausencia de una reacción química específica mediante la cual la fracción de marcador de captura y la fracción de superficie funcionalizada se ligan o asocian químicamente.

La realización ejemplificada en la Figura 7 abarca un esquema de enriquecimiento/selección en el que los productos preadaptadores se retienen en una superficie funcionalizada (por ejemplo, o se inmovilizan de otro modo) durante el procesamiento. En cada etapa del proceso, los intermedios y/o subproductos no deseados que no llevan el marcador de captura se eliminan por lavado o eliminación en varias fases del proceso de síntesis. En algunas realizaciones, los métodos de síntesis en fase sólida permiten el uso de un exceso de la hebra de alargamiento, al menos en parte porque la hebra de alargamiento libre/no hibridada se puede eliminar durante el procesamiento mientras que el exceso de la hebra de plantilla y todos los intermedios/subproductos que llevan el marcador de captura permanecen unidos. En algunas realizaciones, un producto de adaptador dúplex se libera de la superficie funcionalizada (por ejemplo, inmovilización) en una etapa de procesamiento final mientras los intermedios del proceso y los subproductos que llevan el marcador de captura permanecen unidos. De acuerdo con lo anterior, se puede incluir un sitio de reconocimiento de corte de doble hebra en el producto adaptador para garantizar que se haya completado la extensión antes de que pueda ocurrir la liberación del adaptador de la superficie funcionalizada.

En el método de síntesis del adaptador en fase sólida descrito en la Figura 7, una hebra de plantilla que comprende un SMI y un marcador de captura en el extremo 5' de la hebra está asociada (por ejemplo, ligada, unida, inmovilizada, conjugada, reacciona químicamente, unida covalentemente a, unida iónicamente, etc.) a través de su extremo 5'. terminar con una superficie funcionalizada, por ejemplo, mediante la interacción entre un marcador de captura en la hebra de plantilla y una fracción de extracción en la superficie funcionalizada (Figura 7, panel A). En algunas realizaciones, dicha asociación puede ocurrir antes o sustancialmente simultáneamente con una etapa de apareamiento. En algunas realizaciones, la hebra de plantilla también puede incluir una secuencia de plantilla para un dominio de ligación (por ejemplo, para formar un sitio de corte de doble hebra), por ejemplo, 3' con respecto al marcador de captura 5'. Opcionalmente, la hebra de plantilla incluye características tales como un sitio de unión del cebador, una secuencia del cebador leído y/o una región de no complementariedad u otra característica para proporcionar un SDE.

Luego, la hebra de alargamiento (opcionalmente, que incluye un sitio de unión del cebador, una secuencia del cebador leída y/o una región de no complementariedad u otra característica para proporcionar un SDE) y la hebra de plantilla se aparean en una región al menos parcialmente complementaria en una mezcla de reacción para crear un complejo preadaptador (Figura 7, panel B). Debido a que el complejo preadaptador está asociado con la superficie funcionalizada, el exceso de hebra de alargamiento luego se puede eliminar mediante sucesivas etapas de lavado, reduciendo la presencia de intermedios residuales sin una pérdida sustancial del complejo preadaptador. La hebra de alargamiento luego se puede extender enzimáticamente desde el extremo 3' para incluir un SMI y un sitio de corte para crear un complejo adaptador de doble hebra no cortado. Un sitio de corte de doble hebra puede proporcionar tanto un mecanismo de purificación de los productos adaptadores deseados (por ejemplo, separación de entidades unidas a marcadores de captura) como una manera de generar un extremo 3' ligable del producto adaptador. Un experto con conocimientos básicos la técnica apreciará que se puede proporcionar un sitio de corte para lograr uno o más de estos aspectos de procesamiento y/o se pueden proporcionar múltiples sitios de corte para lograr diferentes aspectos de procesamiento.

La Figura 7, panel C ilustra el complejo adaptador de doble hebra no cortado unido después de la extensión enzimática, en el que se crean sitios de corte de enzimas de restricción y SMI de doble hebra. En algunas realizaciones, el producto adaptador deseado se libera en solución mediante tratamiento con una endonucleasa de restricción que corta en el sitio de corte de la enzima de restricción mientras que los productos intermedios restantes, si los hay, y los subproductos son retenidos por la superficie funcionalizada a través del marcador de captura (Figura 7, panel D). El producto adaptador deseado se puede aislar mediante cualquiera de una variedad de métodos, por ejemplo, elución o eliminación de la superficie funcionalizada de la mezcla de reacción. En algunas realizaciones, un método alternativo al esquema ilustrado en la Figura 7, puede incluir la introducción de la superficie funcionalizada después de la etapa de extensión enzimática para capturar complejos adaptadores no cortados (ver, por ejemplo, la Figura 9).

En algunas realizaciones, puede ser deseable asociar una hebra de plantilla que comprende un SMI con una superficie funcionalizada a través del extremo 3' de la hebra de plantilla. Sin desear limitarnos a una teoría particular, se contempla que determinadas polimerasas pueden no ser óptimamente adecuadas para alargar adecuadamente un producto adaptador deseado cuando la hebra de plantilla está unida en el extremo 5'. De acuerdo con lo anterior, en algunas realizaciones, la síntesis del adaptador en fase sólida se puede habilitar aún más mediante la inmovilización de la hebra de plantilla en el extremo 3', como se muestra en la Figura 8. En algunas realizaciones, las hebras de plantilla pueden incluir opcionalmente un sitio de corte de endonucleasa de restricción en el extremo 5' para, por ejemplo, generar un extremo ligable 3' del producto adaptador. En determinadas disposiciones, no mostradas, también se puede incluir un marcador de captura en el extremo 5' de la hebra de tal manera que se pueda eliminar la introducción de una superficie funcionalizada adecuada para extraer moléculas de nucleótidos que llevan el marcador de captura 5' (por ejemplo, esquema de selección negativa para productos adaptadores que hayan sido cortados adecuadamente con extremos de 3' ligables).

En una realización, y como se ilustra en la Figura 8, el extremo 3' de la hebra de plantilla se puede ligar covalentemente a una superficie, como resultado de la síntesis de oligonucleótidos. Una vez que se completa el procesamiento, se puede romper el enlace covalente y se puede liberar el producto adaptador deseado. En una realización alternativa, la hebra de plantilla puede comprender un marcador de captura en 3' y una fracción escindible colocada en 5' con respecto al marcador de captura en 3'.

Como se ilustra en la Figura 8, una hebra de plantilla se puede asociar con perlas de vidrio de poro controlado (CPG) en un extremo 3' de la hebra, como siguiendo un protocolo convencional para sintetizar un oligonucleótido (por ejemplo, la síntesis de la hebra de plantilla) (ver Figura 8, panel A). El experto en fabricación de oligonucleótidos y métodos de purificación de ácidos nucleicos apreciará que se pueden utilizar otras superficies y/o perlas (por ejemplo, perlas de poliestireno microporoso (MPPS)) para la síntesis de oligos y/o para proporcionar de otro modo una hebra de plantilla preunida a una superficie. (por ejemplo, mediante enlace químico).

Después de la síntesis de la hebra de plantilla, por ejemplo, se puede introducir una hebra de alargamiento libre/sin hibridar en la mezcla de reacción y la hebra de plantilla se aparea con una hebra de alargamiento en una región al menos parcialmente complementaria para formar un complejo preadaptador de manera similar a las otras realizaciones descritas anteriormente (Figura 8, panel A). También de manera similar a otras realizaciones descritas en el presente documento y anteriormente, se puede utilizar la extensión enzimática para generar un complejo adaptador de doble hebra que contiene una secuencia SMI en ambas hebras y un sitio de corte 5' (Figura 8, panel B). En algunas realizaciones, por ejemplo, realizaciones que incluyen un dominio de ligación, un evento de escisión (por ejemplo, corte con una enzima de restricción) en el extremo 5' puede dar como resultado la formación de un extremo ligable. En algunas realizaciones, el extremo ligable está libre, por lo que se reducen o eliminan problemas potenciales, tales como los asociados con polimerasas grandes, por ejemplo, sacar prematuramente el producto adaptador de la superficie funcionalizada. Un fragmento de escisión del subproducto se puede eliminar por lavado en la siguiente etapa, o alternativamente, se puede seleccionar negativamente mediante purificación por afinidad (si se proporcionó un marcador de captura; no mostrado). Después de la extensión y el corte, el producto adaptador deseado se puede liberar de las perlas CPG (tal como mediante medios químicos o termolíticos) y recuperarse dando como resultado un producto adaptador purificado (ver Figura 8, panel B). Como se discutió anteriormente, y en algunas realizaciones, la hebra de plantilla puede tener un marcador de captura 3' con una fracción escindible asociada y ser capturada por una superficie funcionalizada diferente (no unida previamente).

C. Realizaciones adicionales de métodos y reactivos para la síntesis de adaptadores

Además de los métodos en fase de solución y en fase sólida descritos anteriormente, también es posible combinar los dos en varios esquemas híbridos, por ejemplo, como se representa en la Figura 9. En algunas realizaciones, dicho enfoque híbrido puede incluir cualquier porción de las etapas en métodos en fase de solución y/o en fase sólida, y en cualquier orden apropiado para la aplicación. En algunas realizaciones, un enfoque híbrido puede ser ventajoso al proporcionar determinados beneficios que no se pueden lograr utilizando solo un método de síntesis en fase de solución o en fase sólida. Por ejemplo, en algunas realizaciones, un método híbrido puede proporcionar determinados beneficios de asociar una hebra de plantilla con una superficie funcionalizada mediante el extremo 5' después de la extensión de la hebra de alargamiento (por ejemplo, evitando que las polimerasas más grandes disocien el complejo preadaptador de una superficie funcionalizada antes de su finalización). Por ejemplo, un enfoque híbrido puede incluir realizar las etapas de apareamiento y extensión en solución y luego los complejos preadaptadores de doble hebra se asocian con una superficie funcionalizada a través de un marcador de captura en el extremo 5' de la hebra de plantilla.

En la realización mostrada en la Figura 9, panel A, el apareamiento de una hebra de alargamiento con una hebra de plantilla en una región de complementariedad al menos parcial, así como extensión (por ejemplo, extensión enzimática) de la hebra de alargamiento (por ejemplo, con una ADN polimerasa) para producir una porción de doble hebra alargada del complejo preadaptador que se produce en solución. Luego, como se ilustra en la Figura 9, panel B, el complejo preadaptador de doble hebra se puede asociar con una superficie funcionalizada, por ejemplo, mediante un agente de captura en el extremo 5' de la hebra de plantilla. Como la extensión ya está completa en este punto, se pueden evitar las restricciones espaciales en el marcador de captura o cerca del mismo. Después de la asociación con la superficie funcionalizada, se puede hacer que ocurra un evento de escisión, por ejemplo en un sitio de corte de endonucleasa de restricción, para liberar solo el producto adaptador deseado (Figura 9, panel C), ya sea inmediatamente después de la asociación (por ejemplo, inmovilización) con la superficie funcionalizada, o después de la realización de una o más etapas de procesamiento adicionales deseadas (por ejemplo, etapas de lavado u otras etapas de purificación o enriquecimiento). En algunas realizaciones, se puede diseñar un evento de escisión para que sea específico de doble hebra, por lo que se requiere que la extensión se haya completado sustancialmente para que se produzca la liberación.

En otras realizaciones, una hebra de plantilla puede comprender una región de autocomplementariedad de tal manera que la hebra de plantilla proporciona una porción de alargamiento y no se requiere una hebra de alargamiento separada. Dicho de otra manera, una hebra de alargamiento se puede asociar con una hebra de plantilla mediante un dominio ligador (por ejemplo, un bucle que comprende nucleótidos de hebra sencilla). El dominio ligador (por ejemplo, el bucle en horquilla) puede comprender una o más fracciones escindibles. En dichas realizaciones, se puede proporcionar una hebra de plantilla que comprende una estructura de bucle de horquilla 3' (por ejemplo, como se describe más adelante con respecto a las Figuras 12-13D) y un marcador de captura 5' como complejo preadaptador adecuado para procesamiento adicional. La plantilla y las hebras/porciones de alargamiento del complejo preadaptador se pueden estructurar como se describió anteriormente para los métodos en fase de solución y/o fase sólida (por ejemplo, pueden incluir una o más secuencias predefinidas, una o más regiones de secuencia degenerada o semidegenerada (por ejemplo, un SMI, una secuencia identificadora), uno o más marcadores de captura, uno o más sitios de corte, etc.). En algunas realizaciones, cualquiera de las etapas del proceso, reactivos y diseños descritos anteriormente para métodos en fase de solución y/o fase sólida pueden ser igualmente aplicables en realizaciones que incluyen una fracción escindible como se describe en las Figuras 10A-B. Por ejemplo, pueden ocurrir una o más etapas de extensión, corte, asociación y/o eliminación, ya sea como se describe para la fase de solución y/o la fase sólida.

La Figura 10A representa una realización que incorpora tanto una fracción escindible como un único agente de captura en el complejo preadaptador. Después de la formación del complejo preadaptador, la Figura 10A representa las etapas de extensión y corte sustancialmente similares a las representadas en la Figura 5 y descritas en la sección titulada "Ejemplo de agente de captura único" anterior, que permite el uso de técnicas de enriquecimiento/selección negativas como se describe. Brevemente, ya sea antes o después de la extensión del complejo preadaptador para formar un complejo preadaptador de doble hebra, el marcador de captura puede estar asociado (por ejemplo, ligado) a una superficie funcionalizada. Después de la asociación con la superficie funcionalizada, se puede cortar un complejo adaptador de doble hebra no cortado en un sitio de corte diseñado, por ejemplo, con una enzima apropiada o, si se desea, se pueden realizar uno o más etapas de lavado u otras etapas de purificación en la mezcla de reacción antes de cortar.

Tras el tratamiento del complejo adaptador de doble hebra no cortado con una enzima apropiada para interactuar con el complejo preadaptador en el sitio de corte, el fragmento escindido que incluye el agente de captura (subproducto) y otros intermedios no deseados que comprenden el marcador de captura (por ejemplo, hebra de plantilla libre/no hibridada, complejo preadaptador no extendido o parcialmente extendido, complejo preadaptador de doble hebra no cortado) se pueden eliminar mediante la eliminación de la superficie funcionalizada dejando el producto adaptador deseado en solución. En algunas realizaciones, una enzima endonucleasa de restricción que corta un sitio de corte de reconocimiento de endonucleasa de restricción (por ejemplo, dentro de un dominio de ligación 3' al SMI) puede producir un extremo ligable 3' en el producto adaptador deseado. En la Figura 10A, el extremo ligable se muestra como un saliente en T; sin embargo, de manera similar a las realizaciones descritas anteriormente, será evidente para un experto en la técnica que el extremo ligable puede tener cualquiera de una variedad de formas, por ejemplo, un

extremo romo, un saliente en A, un extremo "pegajoso" que comprende más de un saliente de nucleótido, entre otros. Independientemente de si se puede introducir un dominio ligable en el producto adaptador deseado o cómo, la fracción escindible presente en el dominio ligador se escinde para abrir el bucle en horquilla para formar un producto adaptador deseado (representado en la Figura 10A como un adaptador en forma de Y).

Alternativamente, en algunas realizaciones, los complejos preadaptadores que incluyen una estructura de bucle en horquilla que liga la plantilla y las hebras de alargamiento (por ejemplo, porciones) también pueden incorporar dos (o más) marcadores de captura, por ejemplo, un primer marcador de captura está presente en el extremo 5' de la hebra de plantilla, y un segundo marcador de captura está presente en el dominio ligador de la estructura de bucle en horquilla. En algunas realizaciones, tal como la que se muestra en la Figura 10B, el primer marcador de captura se sitúa en 5' con respecto a un sitio de corte (por ejemplo, un sitio de corte de endonucleasa de restricción) y el segundo marcador de captura se sitúa entre dos fracciones escindibles en la secuencia de bucle/ligador de hebra sencilla de la estructura de bucle en horquilla. En algunas realizaciones, el uso de un enfoque de marcador de captura doble permite el uso de técnicas de enriquecimiento/selección tanto negativas como positivas para producir un producto adaptador deseado. En la Figura 6 se muestran enfoques de selección positivos y negativos de ejemplo que son compatibles con la inclusión de una estructura de bucle en horquilla y se describieron anteriormente en la sección titulada "Ejemplo de agente de captura múltiple". Brevemente, después de la extensión del complejo preadaptador, el primer marcador de captura se puede asociar (por ejemplo, ligarse) a una primera superficie funcionalizada. Después de la asociación con la primera superficie funcionalizada, el complejo adaptador de doble hebra no cortado se puede cortar en un sitio de corte, por ejemplo, con una enzima apropiada para interactuar con el complejo preadaptador en el sitio de corte, o, si se desea, se pueden realizar una o más etapas de lavado u otras etapas de purificación en la mezcla de reacción antes del corte.

Después del corte, los fragmentos escindidos (subproductos) y otros intermedios no deseados que comprenden el marcador de captura (por ejemplo, hebra de plantilla libre/sin hibridar, complejo preadaptador no extendido o parcialmente extendido, complejo preadaptador de doble hebra no cortado) luego se pueden eliminar mediante eliminación de la superficie funcionalizada dejando en solución el producto adaptador deseado. El segundo marcador de captura se puede asociar (por ejemplo, inmovilizar) a una segunda superficie funcionalizada. En una realización, el primer y segundo marcadores de captura son diferentes y las primera y segunda superficies funcionalizadas o fracciones de extracción se pueden introducir secuencialmente (por ejemplo, en cualquier orden) o simultáneamente. En otras realizaciones, los primeros y segundos marcadores de captura son iguales y/o están capturados por la misma superficie funcionalizada (por ejemplo, una superficie funcionalizada que comprende compañeros de afinidad compatibles para el primer y segundo marcadores de captura, presentación de una fracción de extracción que interactúa con tanto el primer como el segundo marcadores de captura, etc.). Después de la asociación con la segunda superficie funcionalizada, se pueden realizar una o más etapas de lavados u otras etapas de purificación según se desee para eliminar contaminantes, intermedios o subproductos como se describió anteriormente. En el momento deseado, las fracciones escindibles se escinden para liberar el producto adaptador deseado (que se muestra como un adaptador en forma de Y en la Figura 10B). En el ejemplo mostrado en la Figura 10B, una segunda fracción de captura está inmovilizada por la superficie (S2). Los nucleótidos no estándar (por ejemplo, uracilo) flanquean el segundo marcador de captura dentro del bucle de la estructura en forma de horquilla. El enriquecimiento/selección positiva se logra cuando el producto adaptador deseado se libera completamente del segundo marcador de captura luego de la escisión en ambos nucleótidos no estándar. En el ejemplo ilustrado, se puede introducir USER (UDG y Endonucleasa VIII) en la mezcla de reacción para escindir ambas bases de uracilo, liberando de esta manera el producto adaptador deseado que tiene una secuencia identificadora (por ejemplo, una SMI) en ambas hebras.

Una vez que se han escindido las fracciones escindibles, se pueden recoger el producto adaptador deseado, por ejemplo, mediante elución u otra técnica de separación. Las ventajas potenciales de determinadas realizaciones que incorporan una estructura de bucle en horquilla como se describe incluyen garantizar que sustancialmente todos los productos sean de doble hebra (por ejemplo, que no pasen subproductos libres/no hibridados al producto adaptador final) y la capacidad de garantizar sustancialmente relación molar 1: 1 de la plantilla y las hebras de alargamiento (por ejemplo, porciones), lo que reduciría o eliminaría la necesidad de lavar o purificar de otro modo el exceso de plantilla o hebras de alargamiento de la mezcla de reacción.

Secuencias de índice

En algunas realizaciones, se puede utilizar el uso de una o más secuencias de indexación (por ejemplo, códigos de barras de indexación). De acuerdo con diversas realizaciones, se puede integrar una secuencia índice en cualquier posición apropiada para la aplicación dentro de un complejo adaptador (por ejemplo, en una porción de hebra sencilla, una porción de doble hebra). A modo de ejemplo, en algunas realizaciones, se puede integrar una secuencia índice en 5' a un SMI en una porción de doble hebra del adaptador. En algunas realizaciones, se puede integrar una secuencia índice en 3' a un SMI en una porción de doble hebra del adaptador. En algunas realizaciones, una secuencia índice se puede integrar dentro de un SMI (por ejemplo, entre dos secuencias degeneradas o semidegeneradas). De acuerdo con lo anterior, en algunas realizaciones, las secuencias índice pueden ser al menos parcialmente complementarias entre las dos hebras del adaptador de doble hebra. En otras realizaciones, las secuencias índice pueden residir en una porción de hebra sencilla (por ejemplo, en un brazo de un adaptador, dentro de una burbuja formada en el adaptador, etc.). En todavía otras realizaciones, una secuencia índice puede residir parcialmente dentro de una porción no complementaria y dentro de una porción complementaria del adaptador dúplex. Además, un adaptador puede incluir

más de una secuencia índice en cada hebra del adaptador, y cualquier secuencia índice puede estar presente junto con otros códigos de barras o secuencias identificadoras presentes en el adaptador dúplex.

En algunas realizaciones, se pueden utilizar secuencias de indexación para etiquetar muestras de tal manera que las muestras se puedan distinguir entre sí después de la secuenciación, por ejemplo, en celdas de flujo. En algunas realizaciones, las secuencias indexadoras se incorporan en una biblioteca de muestras después de una primera ronda de amplificación utilizando cebadores índice (por ejemplo, las secuencias índice se proporcionan en las colas de los cebadores). La Figura 11A y la Figura 11B ilustran realizaciones de métodos para sintetizar complejos adaptadores que tienen secuencias índice de tal manera que no sean necesarias rondas de amplificación indexadas. Se contempla específicamente que el uso de secuencias índice sea compatible con técnicas en fase solución, fase sólida y/o híbridas como se describen en el presente documento.

A modo de ejemplo concreto, la Figura 11A ilustra un esquema de síntesis de adaptadores en fase sólida que incluye secuencias índice "xxxx" y "yyyy" en regiones no complementarias de las hebras de alargamiento y plantilla (a veces denominadas "cola" o "brazo" de un adaptador en forma de Y). Aunque las secuencias índice mostradas en la Figura 11A se representan como cuatro nucleótidos de longitud, se puede utilizar cualquier número de nucleótidos apropiado para la aplicación en una secuencia índice particular (por ejemplo, entre aproximadamente 2 y 40 o más nucleótidos). En algunas realizaciones, las hebras de alargamiento y/o plantilla se pueden sintetizar de tal manera que incluyan las secuencias índice (por ejemplo, códigos de barras de indexación) antes de la etapa de apareamiento para generar un complejo adaptador preindexado adecuado para el marcado específico de muestra de una biblioteca de ácidos nucleicos. El uso de dichos complejos adaptadores pre-indexados es compatible con los métodos descritos en el presente documento que incluyen, pero no se limitan a, técnicas de etiquetado de captura única (que se representan en la Figura 11A), múltiples técnicas de captura de etiquetas e incorporación de horquillas.

Brevemente, la Figura 11A muestra la inclusión de dos secuencias índice "xxxx" y "yyyy" en regiones no complementarias de la plantilla y las hebras de alargamiento, y la formación de un complejo preadaptador. Al utilizar un esquema de síntesis en fase sólida como se representa, la hebra de plantilla se puede inmovilizar en una superficie funcional (S) a través del marcador de captura 5' antes, durante o después de una etapa de hibridación. Una vez apareada, la hebra de alargamiento experimenta extensión como se describe en otra parte del presente documento, por ejemplo, mediante extensión enzimática desde el extremo 3' para producir un complejo preadaptador de doble hebra extendido. Después de la asociación con la superficie funcionalizada y la extensión de la hebra de alargamiento, el complejo adaptador de doble hebra no cortado se puede cortar en un sitio de corte, por ejemplo, con una enzima apropiada para interactuar con el complejo preadaptador en el sitio de corte, o, si se desea, se pueden realizar una o más etapas de lavados u otras etapas de purificación en la mezcla de reacción antes del corte. Tras el tratamiento del complejo adaptador de doble hebra no cortado con una enzima apropiada para interactuar con el sitio de corte, el fragmento escindido (subproducto), que incluye el agente de captura, permanece unido a la superficie funcionalizada. La superficie funcionalizada se puede separar de la solución, eliminando de esta manera productos intermedios no deseados (por ejemplo, hebra de plantilla libre/no hibridada, complejo preadaptador no extendido o parcialmente extendido, complejo preadaptador de doble hebra no cortado) y subproductos que llevan el marcador de captura. Para lograr un enriquecimiento/selección negativa del producto adaptador deseado, la superficie funcionalizada que tiene una o más fracciones de extracción unidas a la misma es capaz de unirse al marcador de captura (por ejemplo, una perla de estreptavidina donde el marcador de captura es biotina) para la inmovilización y separación de moléculas que lleva el marcador de captura. Se puede utilizar cualquier marcador de captura y fracciones de extracción compatibles como se describe en el presente documento, o aquellos conocidos en la técnica.

Como adición o alternativa a la inclusión de una o más secuencias índice en una región de no complementariedad entre una hebra de plantilla y una hebra de alargamiento (como se muestra en la Figura 11A), se pueden integrar una o más secuencias índice en la porción de doble hebra de una secuencia adaptadora (a veces denominada "tallo" de una secuencia adaptadora).

La Figura 11B ilustra una realización de un esquema de síntesis de adaptador en fase sólida que tiene una secuencia índice presente en la hebra de plantilla. Específicamente, después de las etapas de apareamiento y extensión de la síntesis del adaptador, sustancialmente como se describe en otra parte del presente documento, la secuencia índice está presente en la región de doble hebra del complejo adaptador tanto en la hebra de alargamiento como en la plantilla. En algunas realizaciones, una ventaja de incluir una secuencia índice en la porción de doble hebra del complejo adaptador es que la secuencia índice se puede adquirir en la misma secuencia leída que la secuencia SMI/ácidos nucleicos diana leída, en lugar de requerir una lectura separada.

Extensión en dirección opuesta al extremo ligable

Los ejemplos anteriores ilustran métodos que utilizan una polimerasa de 5' a 3' para extender enzimáticamente la hebra de alargamiento utilizando la hebra de plantilla como plantilla. Los siguientes ejemplos ilustran realizaciones de métodos para formar complejos adaptadores que utilizan una polimerasa de 5' a 3' para extenderse enzimáticamente desde un extremo 3' extensible de tal manera que la polimerasa se desplace en la dirección que se aleja del extremo último ligable de la molécula adaptadora. Como se muestra en ejemplos de determinadas realizaciones, esto se puede facilitar mediante una estructura de horquilla mediante la cual una porción de la hebra de plantilla se autopiega sobre sí misma para servir también como hebra de extensión para generar una región de doble hebra del complejo a

partir de la cual se puede producir una extensión a partir de la cual se genera un SMI de doble hebra. En algunas realizaciones, dichos métodos son particularmente susceptibles de incorporar una o más bases modificadas o no estándar en la hebra de plantilla (incluida la porción que se autoaparea para formar la hebra de extensión, por ejemplo, para generar sitios de escisión para eliminar 3' subproductos complejos, liberar productos adaptadores deseados de una superficie funcionalizada con fines de enriquecimiento y para generar un extremo 3' ligable (por ejemplo, pegajoso) del adaptador sin el uso de enzimas de restricción o endonucleasas específicas de secuencia.

La Figura 12 representa una realización de un método de síntesis de adaptadores que utiliza un esquema de extensión de secuencia en dirección inversa (es decir, en el que la hebra de plantilla se extiende enzimáticamente). En esta realización, la hebra de plantilla incluye una secuencia SMI y una estructura en forma de horquilla que está proximal al extremo 3'. En la Figura 12, la estructura de bucle en horquilla incluye una porción de tallo autocomplementaria que incluye una o más bases modificadas o no estándar, y una porción de ligador de nucleótido de hebra sencilla que incorpora un marcador de captura. La hebra de plantilla se autoaparea en una región de complementariedad al menos parcial dentro de la región 3' de la hebra. De acuerdo con lo anterior, después de una etapa de apareamiento, la hebra de plantilla incluye, en una dirección de 5' a 3', una región 5' de hebra sencilla que incluye una secuencia identificadora de hebra sencilla (por ejemplo, una secuencia SMI) y la estructura de bucle en horquilla 3' que comprende un bucle de nucleótido de hebra sencilla que tiene el marcador de captura y una porción de tallo de doble hebra 3'.

Después del apareamiento, se puede introducir una superficie funcionalizada. Alternativamente, la superficie funcionalizada se puede introducir después de las etapas de apareamiento adicionales, una etapa de extensión o después de la etapa de corte (por ejemplo, consistente con enfoques de fase sólida, fase de solución e híbridos). La asociación con la superficie funcionalizada se puede proporcionar mediante fracciones de extracción que son capaces de unirse al marcador de captura (por ejemplo, una perla de estreptavidina donde el marcador de captura es biotina) para la inmovilización y separación de moléculas que llevan el marcador de captura. Se puede utilizar cualquier marcador de captura y fracción de extracción compatible como se describe en el presente documento o se conoce en la técnica. En cualquier momento después de la asociación con la superficie funcionalizada y durante todo el proceso de síntesis, se pueden realizar una o más etapas de lavado u otras etapas de purificación según se desee para eliminar contaminantes, reactivos no deseados, intermedios o subproductos como se describe en otra parte del presente documento.

En una siguiente etapa, se aparea una segunda hebra con la hebra de plantilla en una región al menos parcialmente complementaria situada dentro de la porción de hebra sencilla 5' de la hebra de plantilla. Como se muestra, la secuencia SMI reside entre la región al menos parcialmente complementaria y la estructura de bucle en horquilla 3'. Después de la hibridación de la hebra de plantilla y la segunda hebra, la hebra de plantilla se puede extender enzimáticamente mediante una ADN polimerasa que no desplaza la hebra (por ejemplo, ADN polimerasa T4, ADN polimerasa T7) para extender la secuencia de ADN de doble hebra en el espacio entre el grupo 3' hidroxilo de la hebra de plantilla y el extremo terminal 5' de la segunda hebra. De esta manera, la secuencia identificadora (SMI) se convierte en de doble hebra.

El complejo preadaptador de doble hebra extendido se inmoviliza en la superficie funcionalizada mediante la interacción del marcador de captura 5' con una o más fracciones de extracción unidas a la superficie. Una vez que se han completado las etapas de lavado deseadas u otras etapas de purificación, se puede desencadenar un evento de escisión, por ejemplo, en la región de nucleótidos modificados o no estándar (que se muestra en la Figura 12 como una región que incluye bases de ribosa) para liberar el producto adaptador deseado de la superficie funcionalizada. En el ejemplo ilustrado, se puede introducir ARNasaH (por ejemplo, ARNasaH2) en la mezcla de reacción para cortar las bases de ribosa y liberar el producto adaptador deseado que tiene una secuencia identificadora (por ejemplo, SMI) en ambas hebras y que tiene un extremo ligable (por ejemplo, un saliente en "T"). La recolección del producto adaptador puede ocurrir en cualquier momento después de su liberación, incluso antes o después de la reparación de mellas como se describe a continuación.

En algunas realizaciones, se introducirán una o más mellas en este proceso, por ejemplo, al final de una etapa de extensión. Dichas mellas se pueden reparar a través del uso de una ligasa apropiada para completar la formación del producto adaptador de doble hebra.

Cabe señalar que aunque se muestran dos oligonucleótidos apareados en este proceso, un experto en la técnica apreciará que se podría lograr un resultado similar con tres oligonucleótidos distintos, por lo que la hebra de plantilla, que se muestra, se autoaparea para formar una horquilla que podría estar compuesta por dos oligonucleótidos separados donde la hebra de plantilla y la porción de extensión de la hebra no están ligadas entre sí y simplemente están apareadas. Uno de estas hebras se podría ligar a la superficie y la otra aparearse a ella, seguido del apareamiento de la tercera hebra mostrada que forma una porción del adaptador en forma de Y en este ejemplo no limitante. Este enfoque también se podría aplicar a otros ejemplos discutidos en el presente documento.

Las Figuras 13A-13C representan realizaciones adicionales de métodos que utilizan un esquema de síntesis de adaptador inverso. Cada una de estas realizaciones adicionales incorpora conceptos discutidos anteriormente y se contempla específicamente que esos diversos conceptos puedan aplicarse a las realizaciones siguientes de cualquier manera apropiada para la aplicación. Por ejemplo, la naturaleza, ubicación y/u orientación del(los) marcador(es) de captura, horquillas y nucleótidos modificados o no estándar pueden variar de las realizaciones analizadas a

continuación. Un experto en la técnica, al exponerse a esta divulgación, apreciará muchas permutaciones de los métodos divulgados en el presente documento, cada uno de los cuales se contempla dentro del alcance de la presente divulgación.

La Figura 13A muestra el método de síntesis de adaptador inverso utilizando una única hebra de plantilla que tiene secuencias autocomplementarias tanto en la región 5' como en la región 3'. Como se ilustra, una primera secuencia 5' es complementaria a una segunda secuencia 5', en la que la primera secuencia 5' está separada de la segunda secuencia 5' con una secuencia ligadora no complementaria que puede incluir una primera fracción escindible, como se muestra en la Figura 13A. Después del apareamiento de la primera y segunda secuencias 5', la secuencia ligadora forma una horquilla 5'. Asimismo, una primera secuencia 3' es complementaria a una segunda secuencia 3', en la que la primera y la segunda secuencia 3' están separadas por una secuencia ligadora no complementaria. Después del apareamiento de la primera y segunda secuencias 3', la región ligadora forma una horquilla 3'. La secuencia ligadora no complementaria en la horquilla 3' puede incluir un marcador de captura, y la región complementaria de la horquilla 3' puede incluir además una segunda fracción escindible (por ejemplo, uno o más nucleótidos modificados o no estándar, tales como una región de bases ribosa como se representa en la Figura 13A).

Después del apareamiento, se puede introducir una superficie funcionalizada. Alternativamente, la superficie funcionalizada se puede introducir después de etapas adicionales que incluyen una etapa de extensión o después de la etapa de corte (por ejemplo, consistente con los enfoques de fase sólida, fase de solución e híbridos discutidos anteriormente). La asociación con la superficie funcionalizada se puede proporcionar mediante fracciones de extracción que son capaces de unirse al marcador de captura (por ejemplo, una perla de estreptavidina donde el marcador de captura es biotina) para la inmovilización y separación de moléculas que llevan el marcador de captura. Se puede utilizar cualquier marcador de captura y fracción de extracción compatible como se describe en el presente documento, o se conoce en la técnica. En cualquier momento después de la asociación con la superficie funcionalizada y durante todo el proceso de síntesis, se pueden realizar una o más etapas de lavado u otras etapas de purificación según se desee para eliminar contaminantes, reactivos no deseados, intermedios o subproductos como se describe en otra parte del presente documento.

Se utiliza una polimerasa, tal como una polimerasa que carece de actividad de desplazamiento de hebras (por ejemplo, ADN polimerasa T4 o T7), para extender las porciones complementarias en las regiones 5' y 3', llenando de esta manera el espacio entre ellas. En particular, la polimerasa puede extender la hebra de plantilla desde el grupo hidroxilo 3' en el extremo 3' en una dirección 5' a 3' para encontrarse con el extremo terminal 5' de la hebra de plantilla. Como tal, una secuencia identificadora (por ejemplo, un SMI que contiene nucleótidos degenerados o semidegenerados) se convierte en de doble hebra. La ligasa se utiliza para sellar y reparar muescas entre los extremos 3' y 5' (Figura 13A). Después de la extensión y/o ligación de mella, los complejos preadaptadores circularizados se pueden tratar opcionalmente con exonucleasa para destruir los ácidos nucleicos restantes en la mezcla de reacción. Debido a que los complejos preadaptadores deseados están unidos a la superficie a través del marcador de captura, la exonucleasa se puede lavar/eliminar.

A continuación, se puede desencadenar un primer evento de escisión para alterar la primera fracción escindible en la horquilla 5' y crear dos colas para formar una región en forma de Y. Ya sea antes, después o sustancialmente al mismo tiempo que el primer evento de escisión, se puede desencadenar un segundo evento de escisión en la segunda fracción escindible (en el presente documento una región que comprende bases de ribosa dentro de la región complementaria de la horquilla 3', como se muestra en la Figura 13A) para liberar el producto adaptador deseado. En algunas realizaciones, el segundo evento de escisión produce un extremo 3' ligable. En algunas realizaciones, la primera fracción escindible y la segunda fracción escindible pueden ser sustancialmente iguales, de tal manera que un único evento de escisión puede tanto liberar el producto adaptador del marcador de captura (y cualquier secuencia escindida asociada) así como crear una forma de Y (ver la realización descrita en la Figura 13B a continuación). La separación de la superficie con el subproducto unido y los intermedios no deseados de los productos adaptadores deseados en solución proporciona un adaptador dúplex purificado que tiene una secuencia identificadora (SMI) en ambas hebras.

La Figura 13B ilustra una realización similar en la Figura 13A, con dos diferencias notables. En primer lugar, la realización representada en la Figura 13B incluye un uracilo como primera fracción escindible y como segunda fracción escindible. En segundo lugar, se integra una secuencia índice en la hebra de plantilla 3' del SMI. Como se ilustra, se utiliza un uracilo para forzar un emparejamiento erróneo cerca del extremo 3' del complejo adaptador intermedio de tal manera que la aplicación de una enzima apropiada, tal como USER, se escinda para formar un extremo pegajoso (por ejemplo, saliente en T) y también para formar un Adaptador en forma de Y. Tanto la Figura 13A como la Figura 13B ilustran métodos que utilizan la incorporación de una o más bases modificadas o no estándar para formar un extremo ligable y/o para escindir la característica de horquilla 5' para formar un complejo adaptador en forma de Y.

La Figura 13C ilustra una variación de una realización de síntesis de adaptador inverso en la que se genera una característica asimétrica (por ejemplo, un SDE). En esta realización, una hebra de plantilla que tiene una base modificada incorporada (por ejemplo, uracilo) seguida de una secuencia SMI de hebra sencilla también tiene una porción complementaria en una región 3'. Como se ilustra, una primera secuencia 3' es complementaria a una segunda secuencia 3', en la que la primera secuencia 3' está separada de la segunda secuencia 3' con una secuencia ligadora no complementaria. Después del apareamiento de la primera y segunda secuencias 3', la secuencia ligadora forma

una horquilla 3'. En esta realización, la horquilla 3' también incluye un sitio de corte y un marcador de captura, que se puede asociar con una superficie funcionalizada que comprende una o más fracciones de extracción compatibles como se describe en otra parte del presente documento. Se utiliza una polimerasa de 5' a 3' para extender enzimáticamente la hebra para formar una molécula de doble hebra con una horquilla en 3'. Después de la extensión, se puede utilizar una enzima (por ejemplo, USER) para formar una mella en la base modificada y se puede eliminar la porción 5' de la hebra de plantilla (por ejemplo, mediante la aplicación de temperatura elevada). Después de la eliminación de la porción 5' de la hebra de plantilla, se puede aparear un oligo parcialmente complementario (por ejemplo, una segunda hebra) con una región del complejo de hebra sencilla. El oligo parcialmente complementario puede tener una región no complementaria de tal manera que después del apareamiento y la ligación de la mella, se introduce asimetría en el complejo adaptador. Se puede desencadenar un evento de escisión en el extremo 3' del complejo adaptador (por ejemplo, en el sitio de corte) que puede liberar el producto deseado de una superficie funcionalizada (si se utiliza síntesis en fase sólida) mientras, en algunas realizaciones, se genera un extremo ligable.

La Figura 13D ilustra una variación de una realización de síntesis de adaptador inverso utilizando incorporación de base modificada o no estándar para formar un complejo adaptador en forma de Y sin utilizar una etapa de ligación. Como se ilustra, la Figura 13D muestra un método de síntesis de adaptador inverso utilizando una hebra de plantilla única que tiene una porción complementaria en una región media y en una región 3'. Como se ilustra, después de una región 5' no complementaria, una primera secuencia de nucleótidos interna es complementaria a una segunda secuencia de nucleótidos interna, en la que la primera secuencia de nucleótidos interna está separada de la segunda secuencia de nucleótidos interna con una secuencia ligadora interna no complementaria. Después del apareamiento de la primera y segunda secuencias internas, la secuencia ligadora forma una horquilla interna. La secuencia ligadora interna se proporciona con una base modificada/escindible (por ejemplo, uracilo).

Adicionalmente, una primera secuencia 3' es complementaria a una segunda secuencia 3', en la que la primera y la segunda secuencia 3' están separadas por una secuencia ligadora no complementaria. Después del apareamiento de la primera y segunda secuencias 3', la región ligadora forma una horquilla 3'. En esta realización, la horquilla 3' también incluye un marcador de captura que se puede asociar con una superficie funcionalizada que comprende una o más fracciones de extracción compatibles como se describe en otra parte del presente documento. Se utiliza una polimerasa para extender las porciones complementarias en la región 3'. En esta realización, la polimerasa cruzará un espacio de nucleótidos formado por la estructura de horquilla interna formada en la región media de la hebra de plantilla y continuará formando una secuencia de doble hebra complementaria a la región 5' de la hebra de plantilla. Después de la extensión, la escisión en la base modificada (por ejemplo, uracilo) ubicada en/dentro de la secuencia ligadora interna y la eliminación de la región 5' original de la hebra de plantilla da como resultado una asimetría de secuencia en el extremo 5' del adaptador. Un evento de digestión/escisión en el extremo 3' del complejo adaptador puede liberar el producto deseado de una superficie funcionalizada (si se utiliza síntesis en fase sólida) mientras, en algunas realizaciones, se genera un extremo ligable. Después de la eliminación de la superficie funcionalizada y los productos intermedios y secundarios asociados, se proporciona un adaptador dúplex purificado que tiene una secuencia identificadora (SMI) en ambas hebras.

En algunas realizaciones, los métodos descritos en el presente documento pueden diferir de los métodos descritos anteriormente, entre otros, al permitir el uso de un exceso de hebra de plantilla y evitar el problema de tener una hebra de alargamiento libre en solución que no se puede separar fácilmente del producto adaptador deseado. Adicionalmente, algunas realizaciones de los métodos descritos en el presente documento pueden proporcionar pureza del adaptador, eficiencia de fabricación, reproducibilidad, coste y flexibilidad del diseño del adaptador mejoradas.

Un experto en la técnica apreciará que cualesquier porciones de esta divulgación se puede utilizar en una amplia variedad de combinaciones. Si bien la presente divulgación describe varias de estas combinaciones, las realizaciones ejemplificadas no pretenden ser limitantes.

Equivalentes y alcance

Las descripciones detalladas anteriores de realizaciones de la tecnología no pretenden ser exhaustivas ni limitar la tecnología a la forma precisa divulgada anteriormente. Aunque realizaciones específicas y ejemplos de la tecnología se describen anteriormente con fines ilustrativos, son posibles diversas modificaciones equivalentes, como reconocerán los expertos en la técnica relevante. Por ejemplo, aunque las etapas se presentan en un orden determinado, realizaciones alternativas pueden realizar etapas en un orden diferente. Las diversas realizaciones descritas en el presente documento también se pueden combinar para proporcionar realizaciones adicionales.

A partir de lo anterior, se apreciará que en el presente documento se han descrito realizaciones específicas de la tecnología con fines ilustrativos, pero no se han mostrado ni descrito en detalle estructuras y funciones bien conocidas para evitar oscurecer innecesariamente la descripción de las realizaciones de la tecnología. Cuando el contexto lo permita, los términos singulares o plurales también pueden incluir el término plural o singular, respectivamente. Además, si bien las ventajas asociadas con determinadas realizaciones de la tecnología se han descrito en el contexto de esas realizaciones, otras realizaciones también pueden exhibir dichas ventajas, y no todas las realizaciones necesitan necesariamente exhibir dichas ventajas para estar dentro de la tecnología. De acuerdo con lo anterior, la

divulgación y la tecnología asociada pueden abarcar otras realizaciones no mostradas o descritas expresamente en el presente documento.

5 Los expertos en la técnica reconocerán, o podrán determinar utilizando únicamente experimentación de rutina, muchos equivalentes a las realizaciones específicas de la tecnología divulgada descrita en el presente documento. El alcance de la presente tecnología no pretende limitarse a lo anterior. El alcance de la presente invención está definido por las reivindicaciones adjuntas.

REIVINDICACIONES

1. Un método para preparar un adaptador dúplex, que comprende:
aparear una hebra de alargamiento y una hebra de plantilla en una región complementaria para formar un primer producto adaptador dúplex intermedio, en el que la hebra de plantilla comprende-
5 una secuencia identificadora,
un primer marcador de captura adherido a la hebra de plantilla a través de un primer sitio de corte, y
un segundo marcador de captura adherido a la hebra de plantilla a través de un segundo sitio de corte;
extender la hebra de alargamiento para duplicar al menos parcialmente la secuencia identificadora para formar un
segundo producto adaptador dúplex intermedio que comprende un primer sitio de corte entre la secuencia
10 identificadora y el primer marcador de captura;
cortar el segundo producto adaptador dúplex intermedio para formar un tercer producto adaptador dúplex intermedio
y un primer subproducto escindido que comprende el primer marcador de captura;
eliminar los productos no deseados que comprenden el primer marcador de captura; y
cortar en el segundo sitio de corte para formar el adaptador dúplex y un segundo subproducto escindido que
15 comprende el segundo marcador de captura, y
eliminar los productos no deseados.
2. El método de la reivindicación 1, en el que el primer marcador de captura está en un extremo 5' de la hebra de
plantilla de tal manera que las hebras de plantilla no apareadas, el primer y segundo productos adaptadores dúplex
intermedios y el primer subproducto escindido comprenden el primer marcador de captura.
- 20 3. El método de la reivindicación 1 o la reivindicación 2, que comprende además proporcionar una superficie que
comprende una fracción de extracción configurada para unirse al primer o segundo marcador de captura.
4. El método de la reivindicación 3, en el que la hebra de plantilla se une a la superficie antes de la etapa de
apareamiento.
5. El método de la reivindicación 3, en el que la superficie que comprende la fracción de extracción está configurada
25 para unirse al primer marcador de captura y la hebra de plantilla se une a la superficie a través del primer marcador
de captura durante las etapas de apareamiento y extensión.
6. El método de la reivindicación 3, en el que el primer marcador de captura está unido a la fracción de extracción en
la superficie configurada para unir el primer marcador de captura después de la etapa de apareamiento y antes de la
primera etapa de corte.
- 30 7. El método de la reivindicación 3, en el que la etapa de eliminación comprende separar la superficie de una solución
líquida que comprende el tercer producto adaptador dúplex intermedio, y en el que los productos no deseados se unen
a la superficie a través del primer marcador de captura.
8. El método de la reivindicación 1, en el que la hebra de plantilla y la hebra de alargamiento están ligadas mediante
un dominio ligador.
- 35 9. El método de la reivindicación 8, en el que el primer marcador de captura está en un extremo 5' de la hebra de
plantilla, y en el que el dominio ligador comprende el segundo marcador de captura flanqueado por dos segundos
sitios de corte.
10. El método de una cualquiera de las reivindicaciones 1-9, que comprende además proporcionar al menos una
fracción de extracción configurada para unirse al segundo marcador de captura, y capturar productos no deseados
40 que incluyen el segundo marcador de captura.
11. El método de la reivindicación 10, en el que la fracción de extracción configurada para unir el segundo marcador
de captura se proporciona antes o después de la segunda etapa de corte.
12. El método de una cualquiera de las reivindicaciones 1-11, en el que los productos no deseados comprenden uno
o más de un exceso de hebra de plantilla, exceso de hebra de alargamiento, complejos preadaptadores no extendidos
o extendidos de forma incompleta, segundos y terceros productos adaptadores dúplex intermedios no cortados, y
45 subproductos de fragmentos de escisión.
13. El método de una cualquiera de las reivindicaciones 1-12, en el que el primer marcador de captura es diferente del
segundo marcador de captura.

14. El método de una cualquiera de las reivindicaciones 1-12, en el que después de la segunda etapa de corte, el método comprende además eliminar los productos no deseados que comprenden el segundo marcador de captura.

15. El método de una cualquiera de las reivindicaciones 8-14, en el que el corte en el segundo sitio de corte comprende cortar el dominio ligador para proporcionar dos brazos de hebra sencilla en el extremo 5' del producto adaptador, en el que el dominio ligador contiene una o más moléculas de nucleótidos o no nucleótidos modificadas.

5

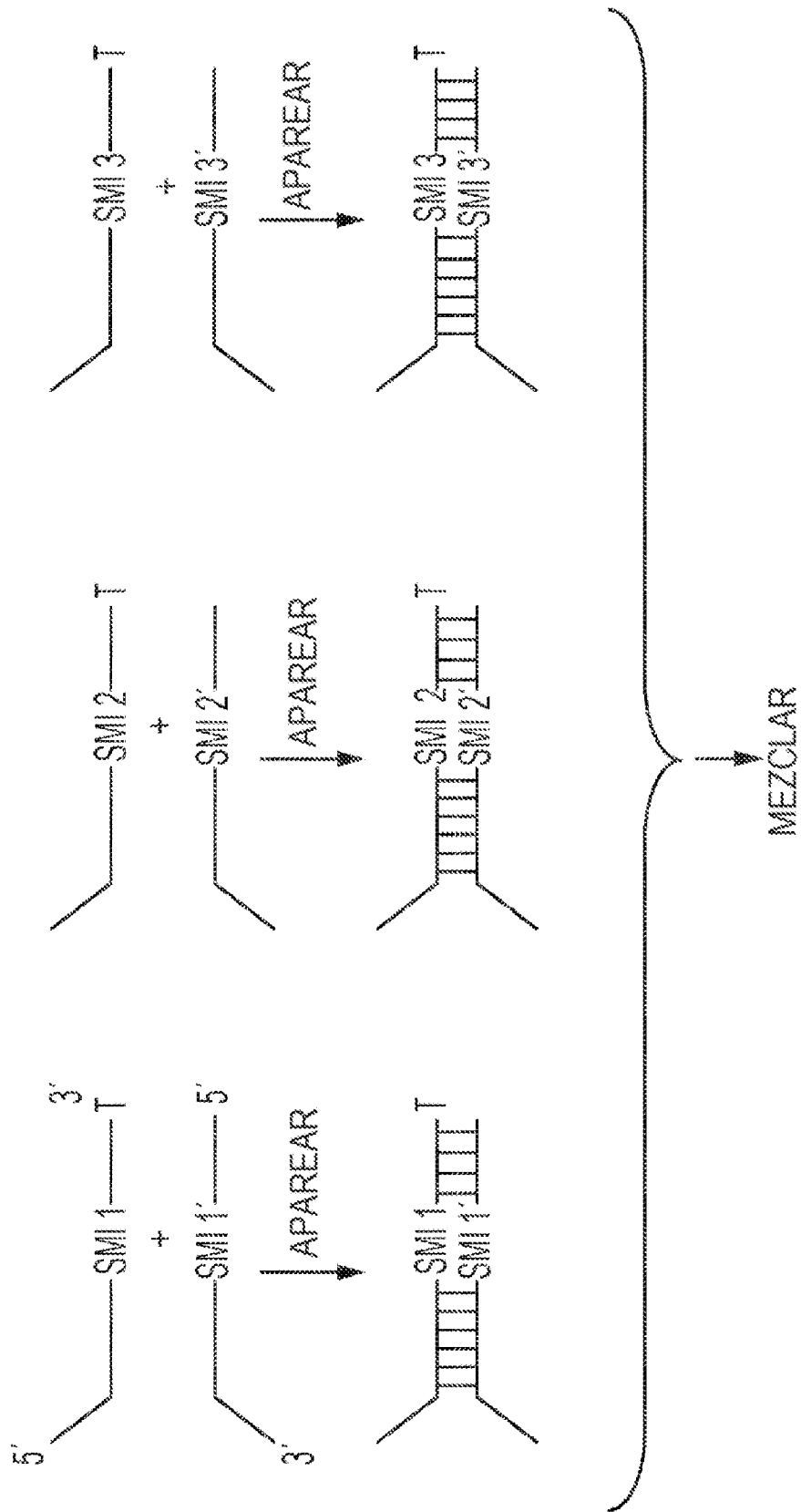
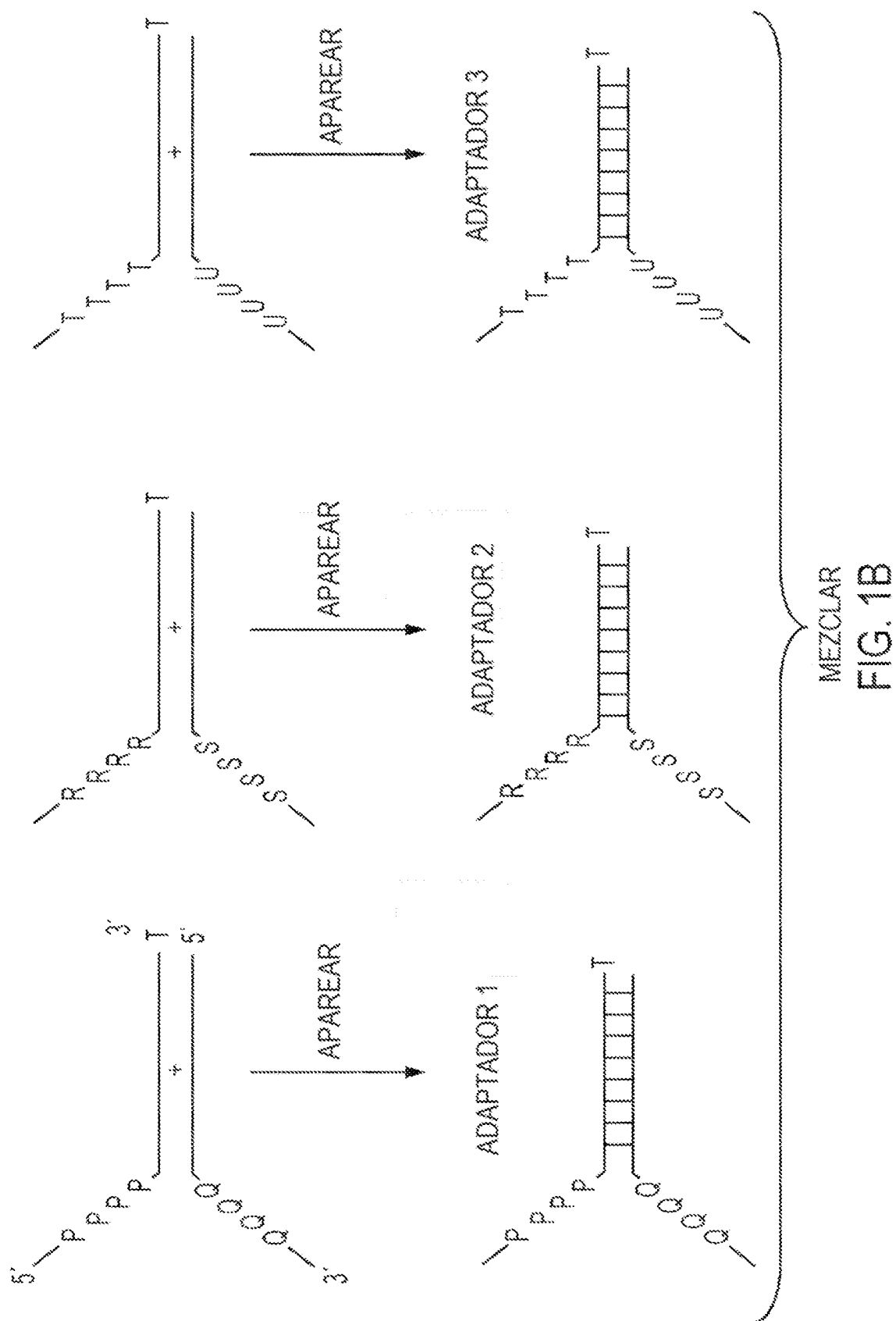


FIG. 1A



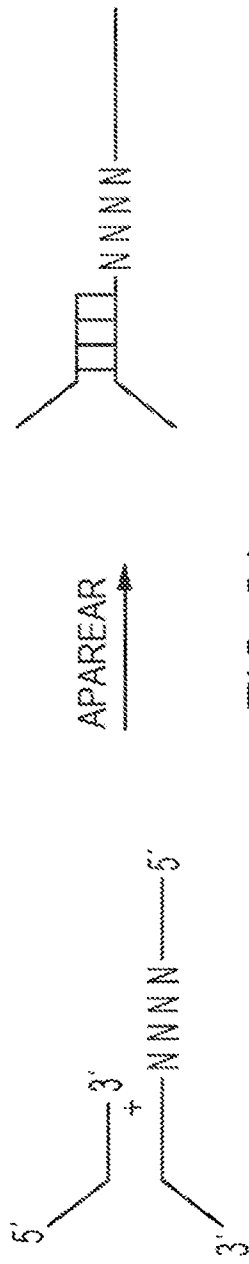


FIG. 2A

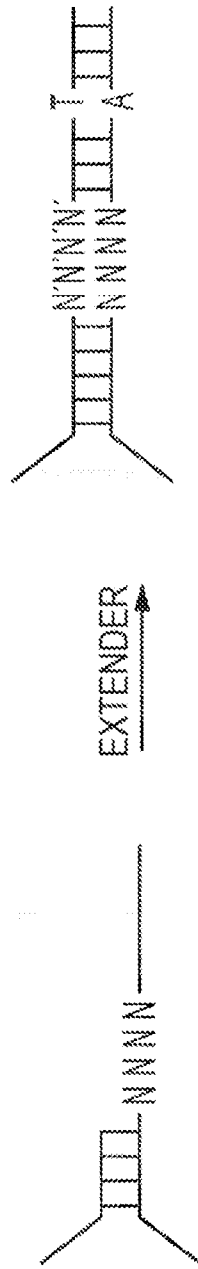


FIG. 2B

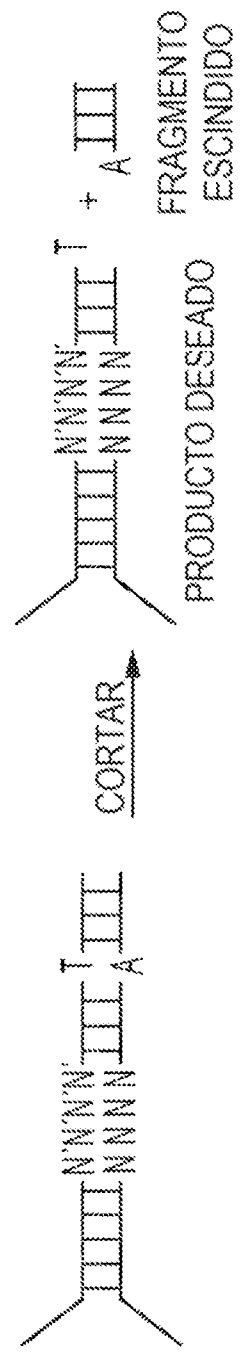


FIG. 2C

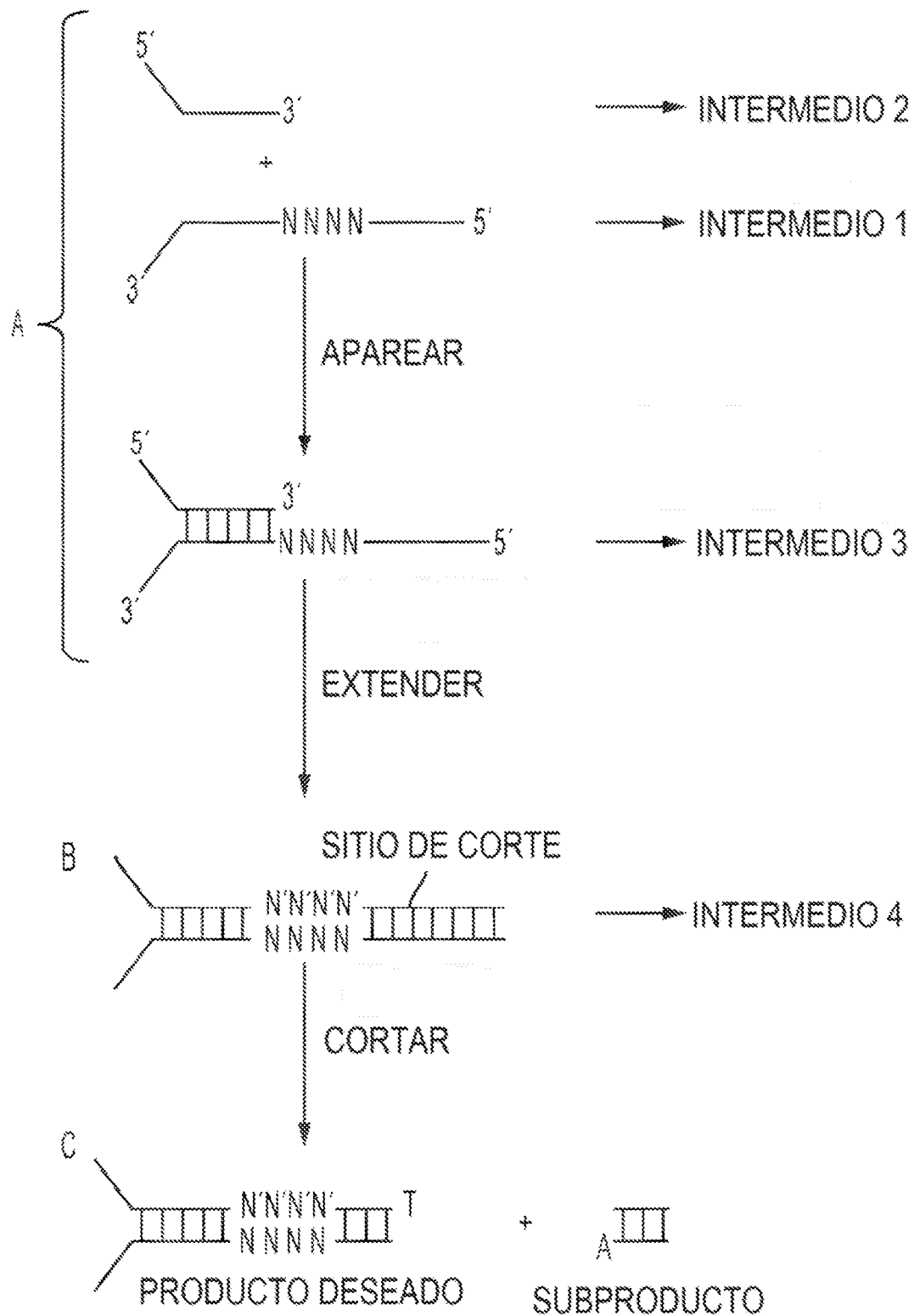


FIG. 3

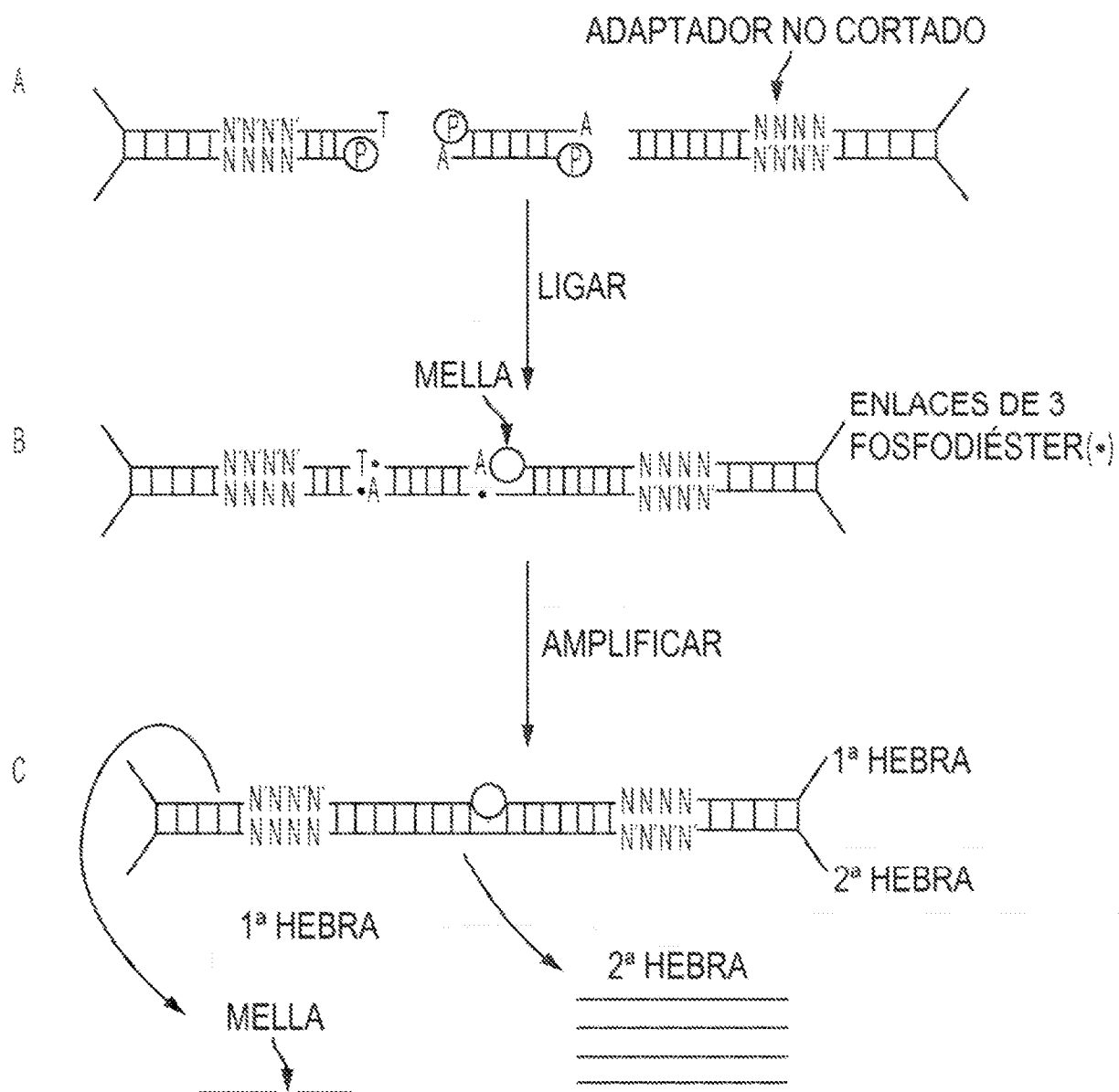


FIG. 4

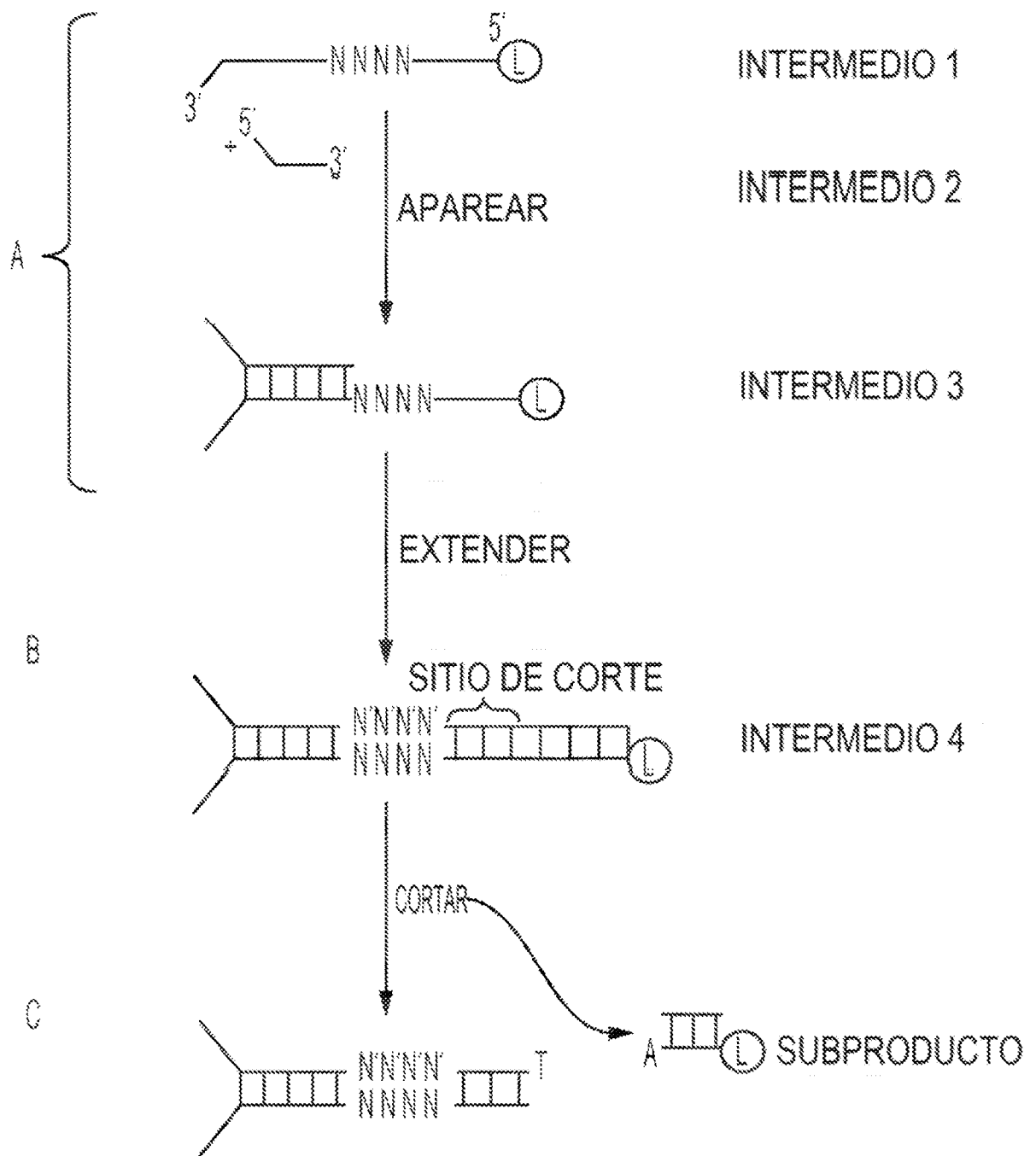


FIG. 5

CONTINUA DESDE PANEL C

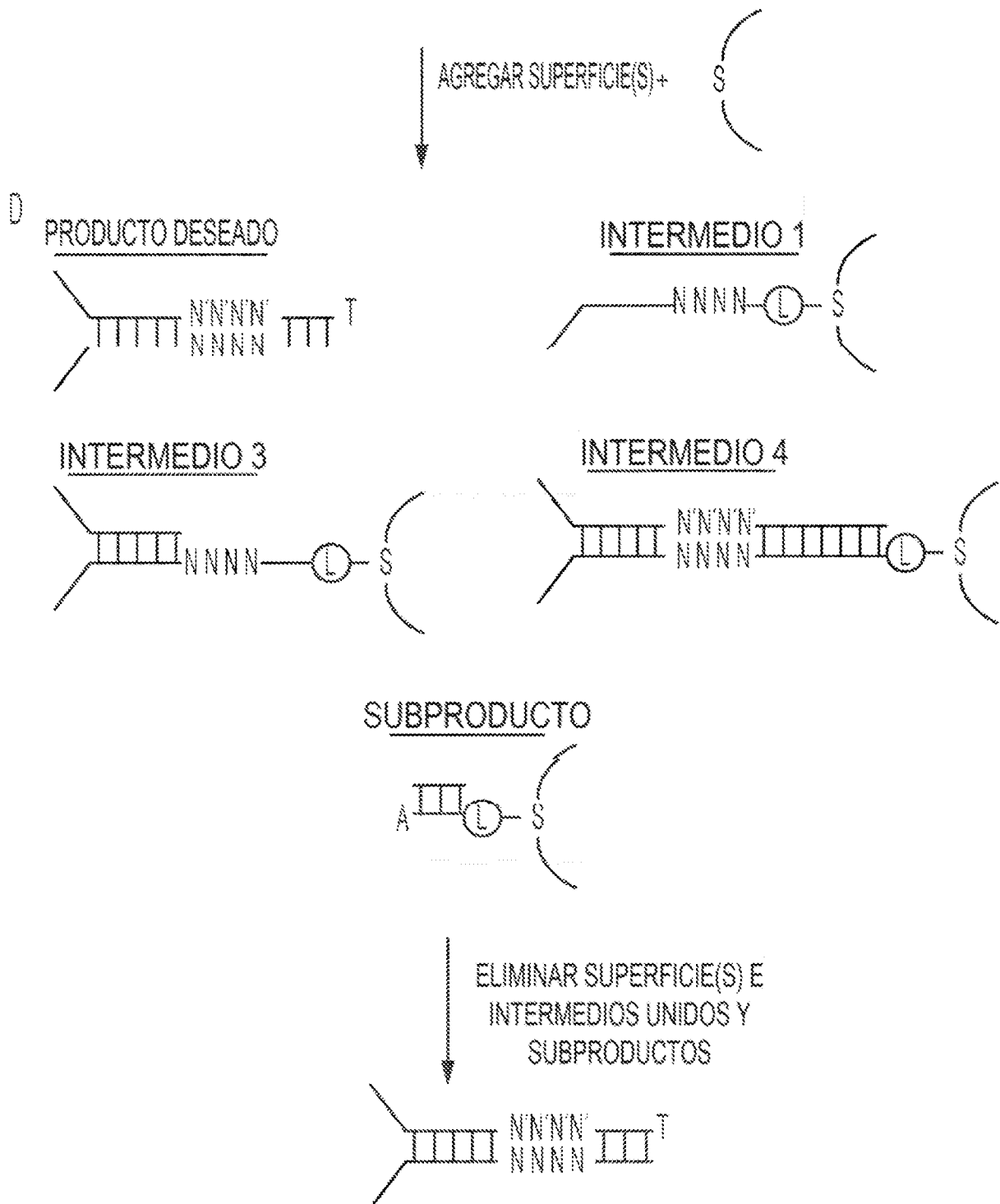


FIG. 5 Continuación

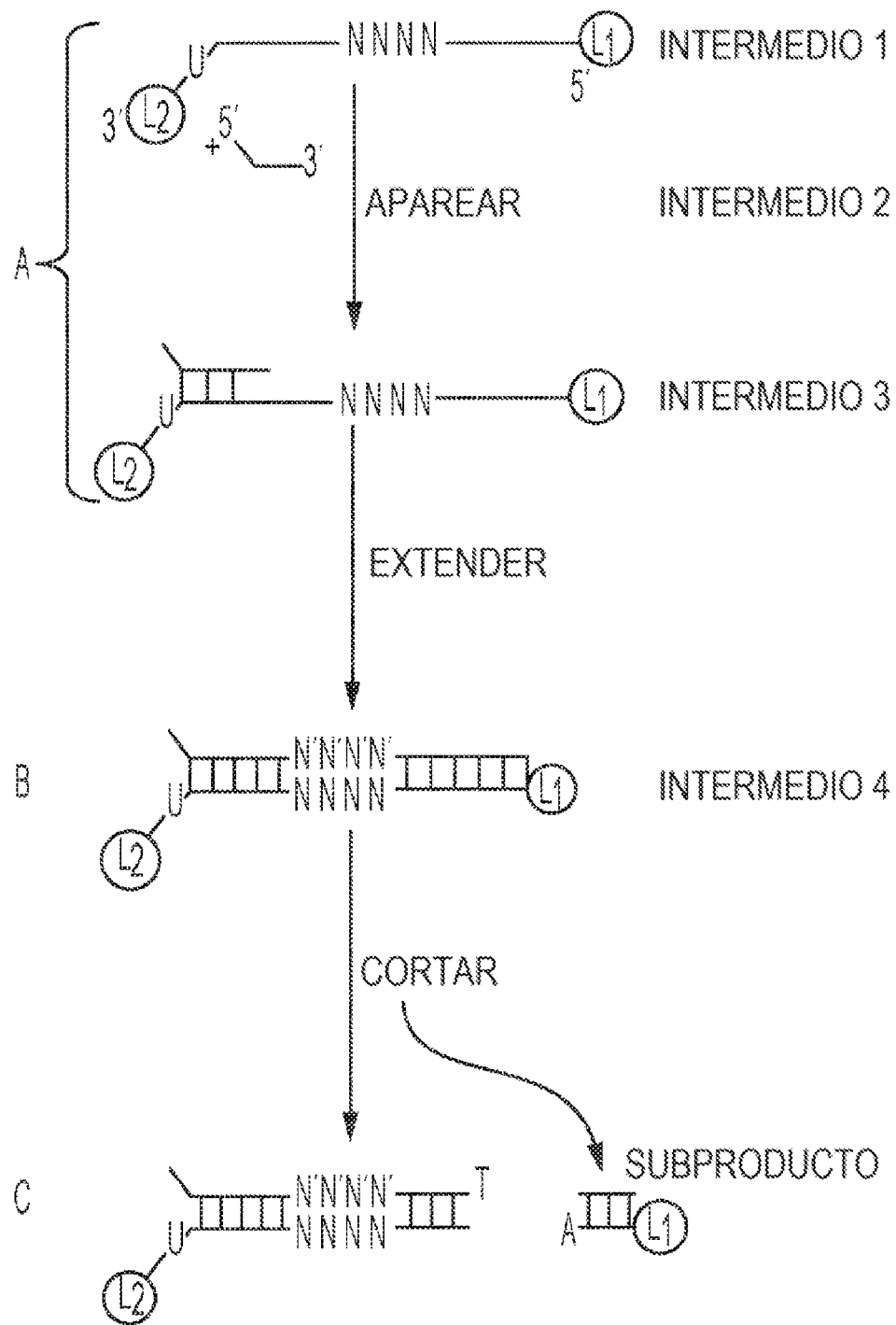


FIG. 6

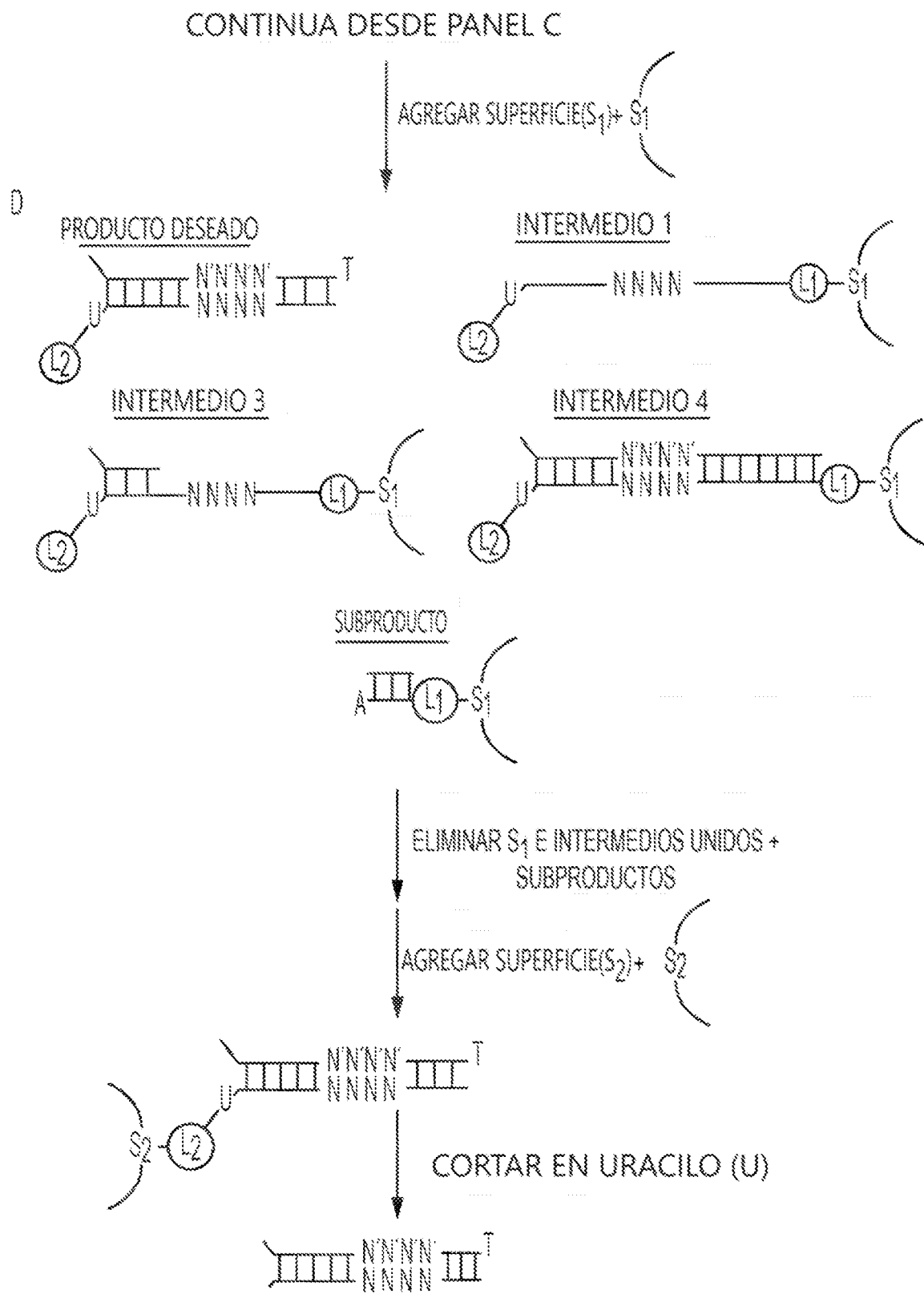


FIG. 6 Continuación

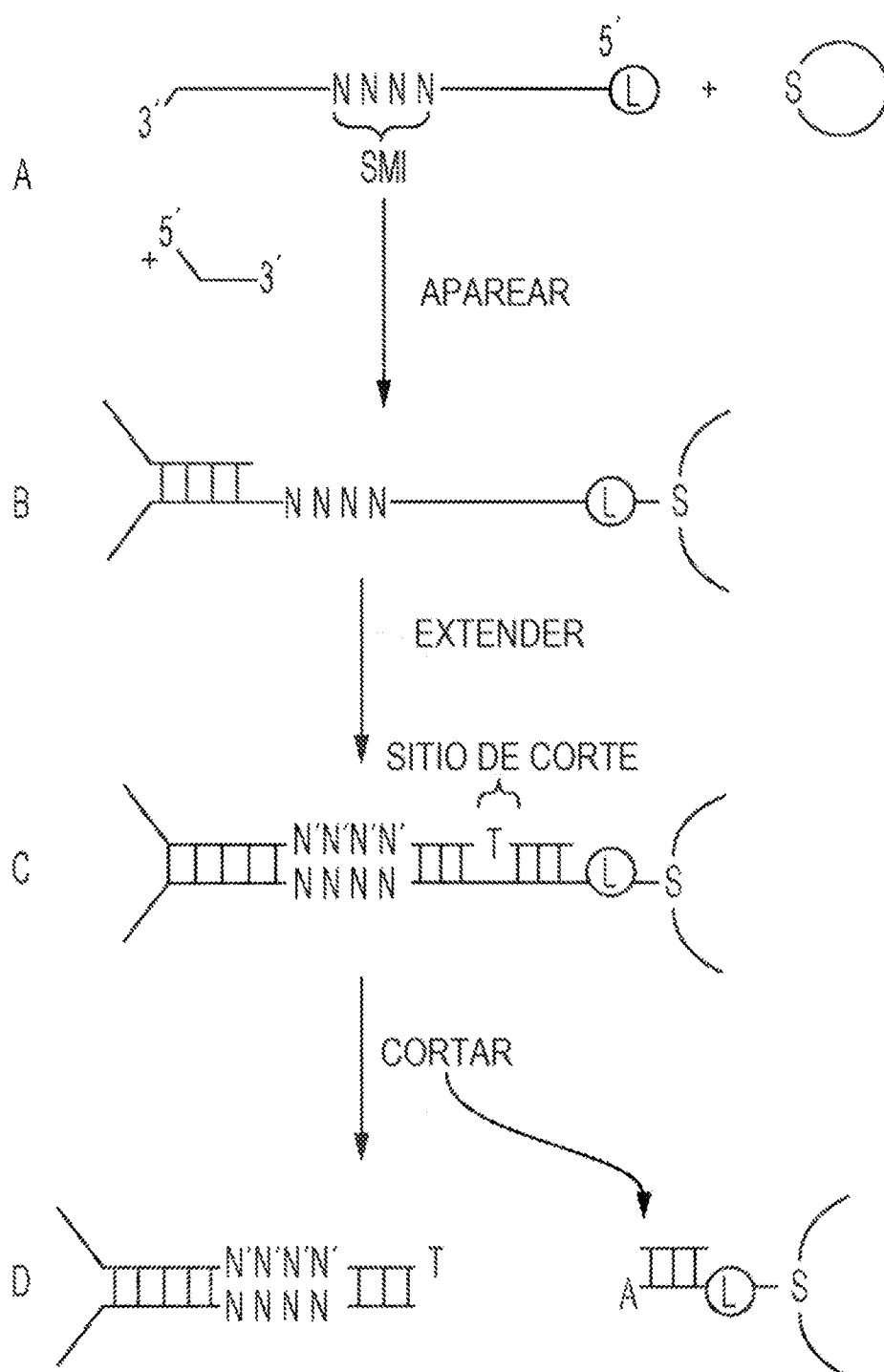


FIG. 7

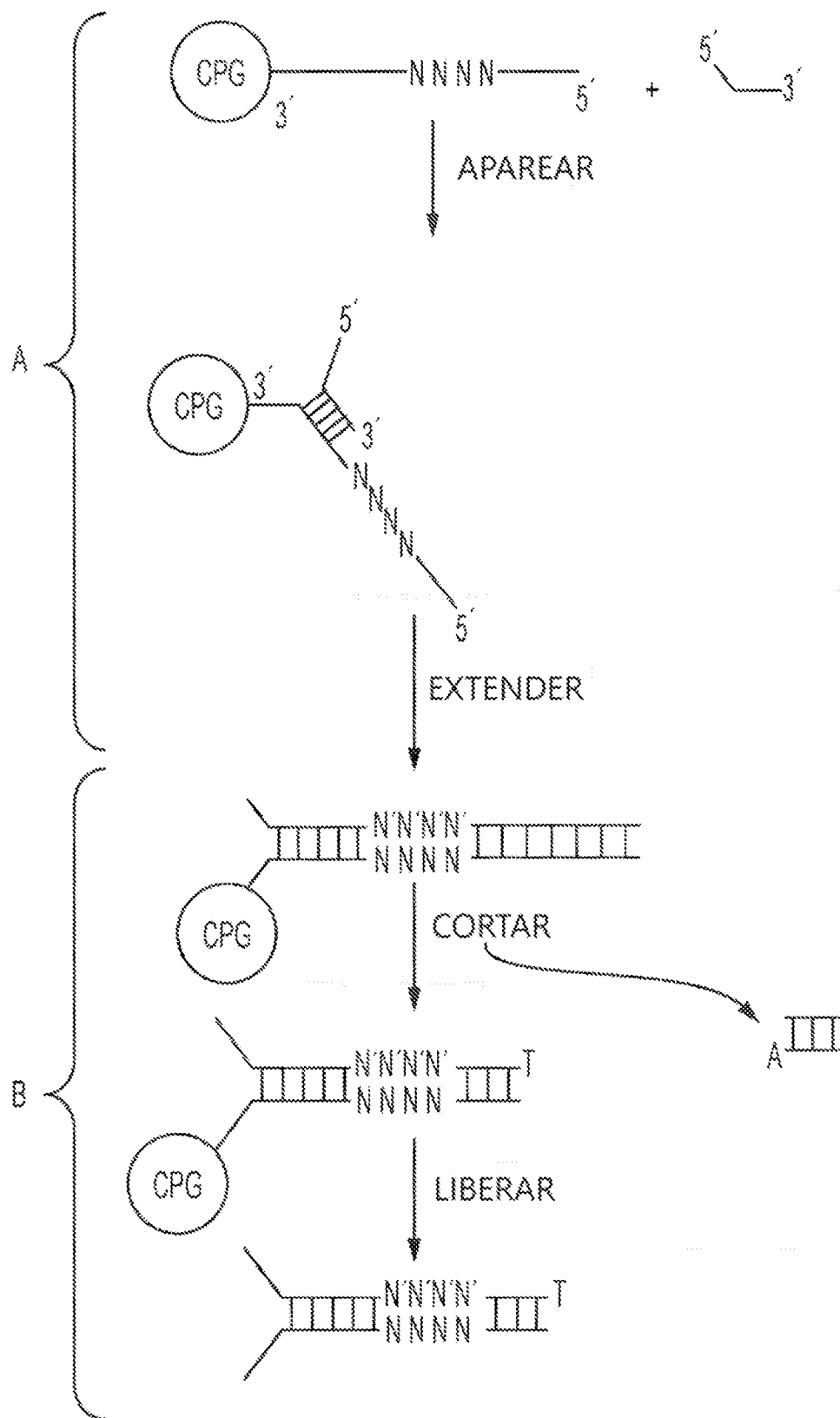


FIG. 8

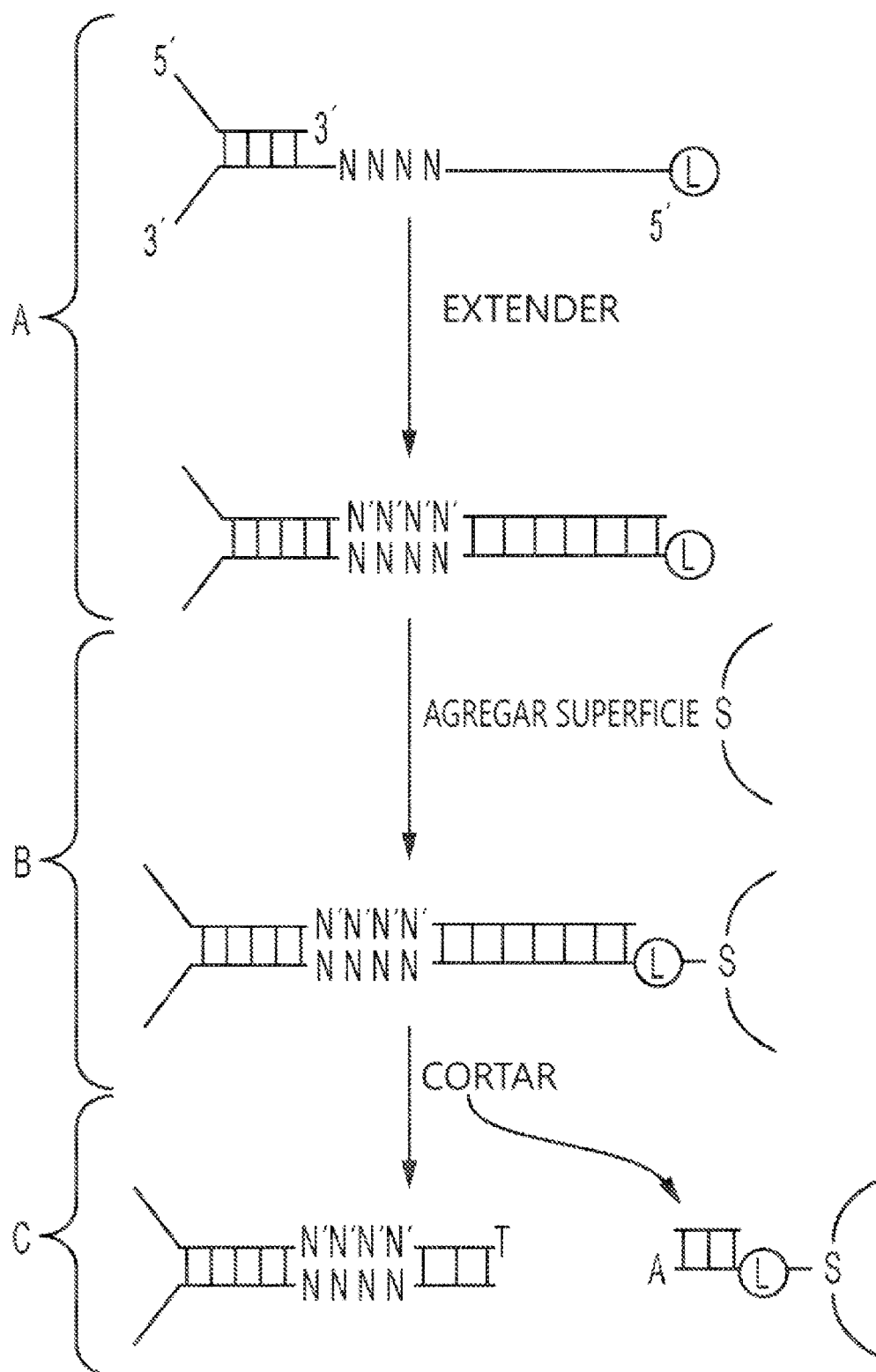


FIG. 9

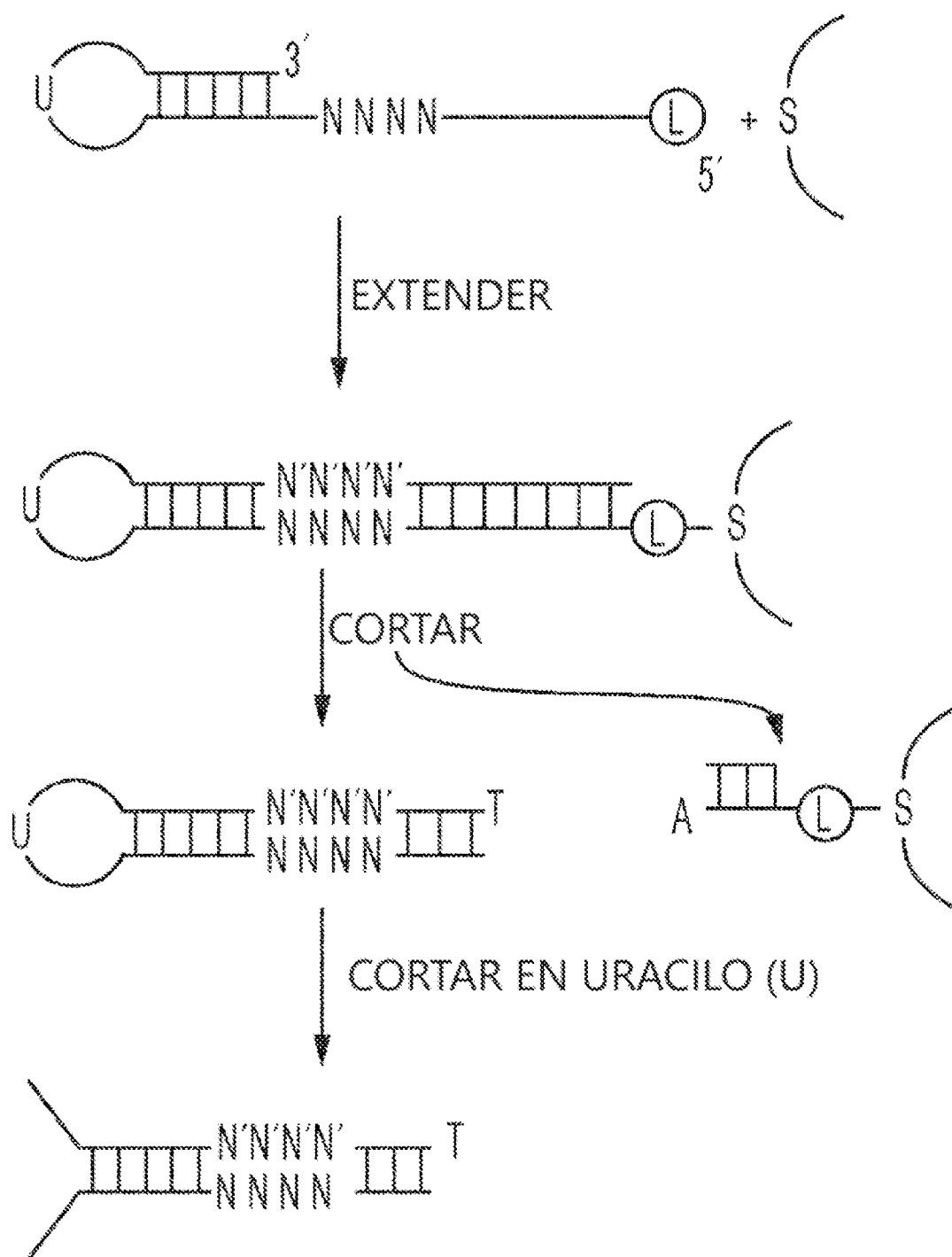


FIG. 10A

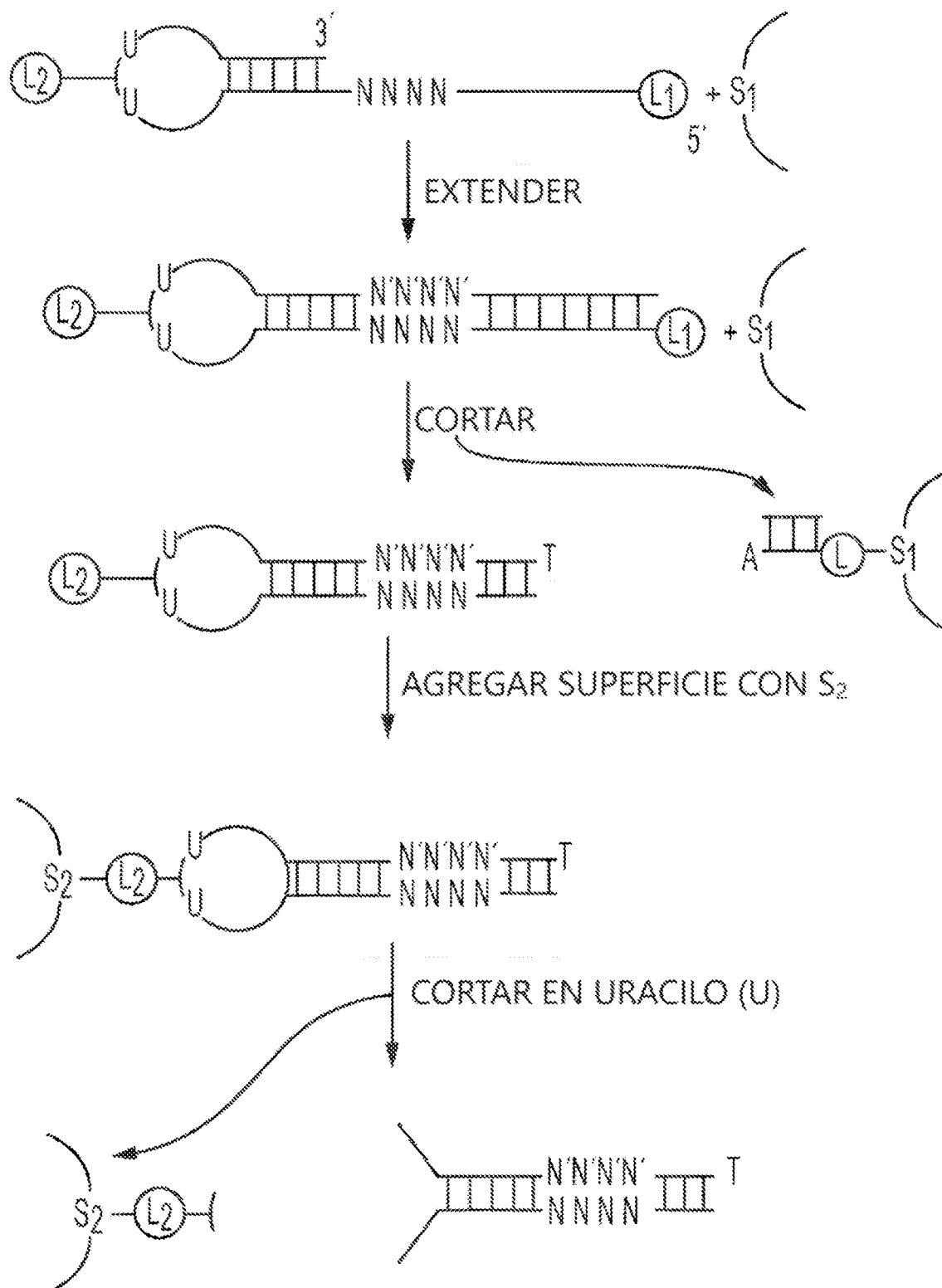


FIG. 10B

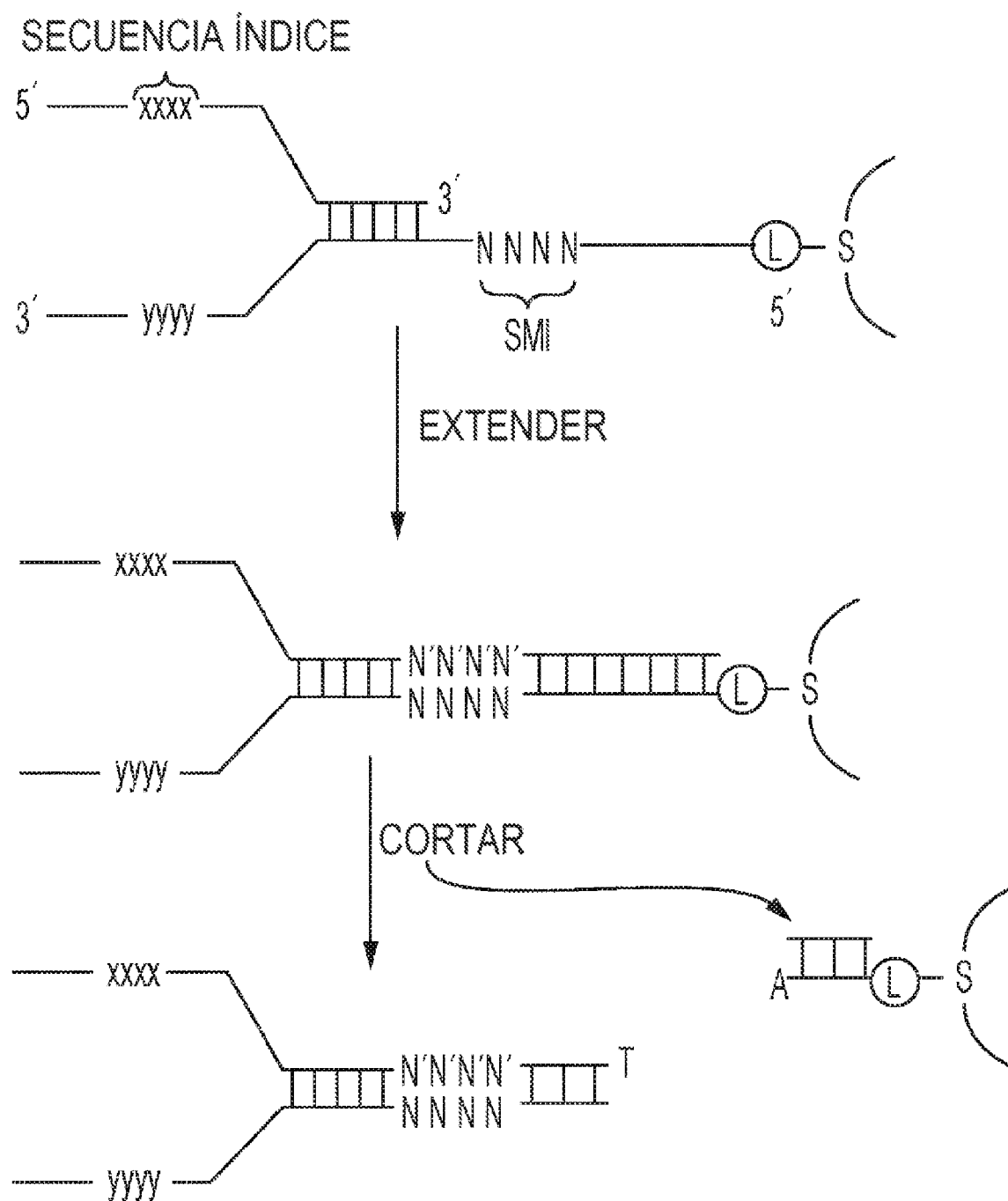


FIG. 11A

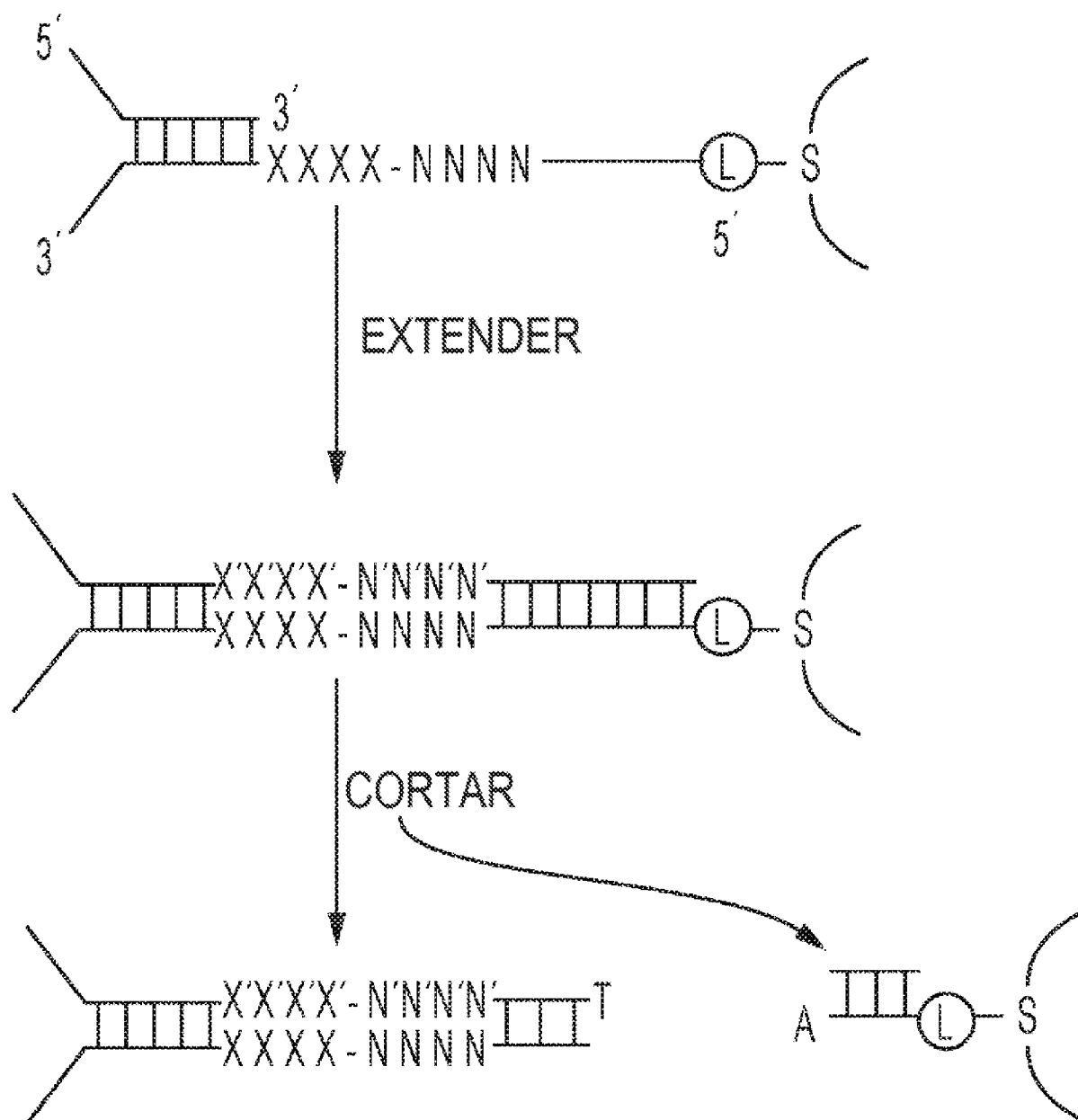


FIG. 11B

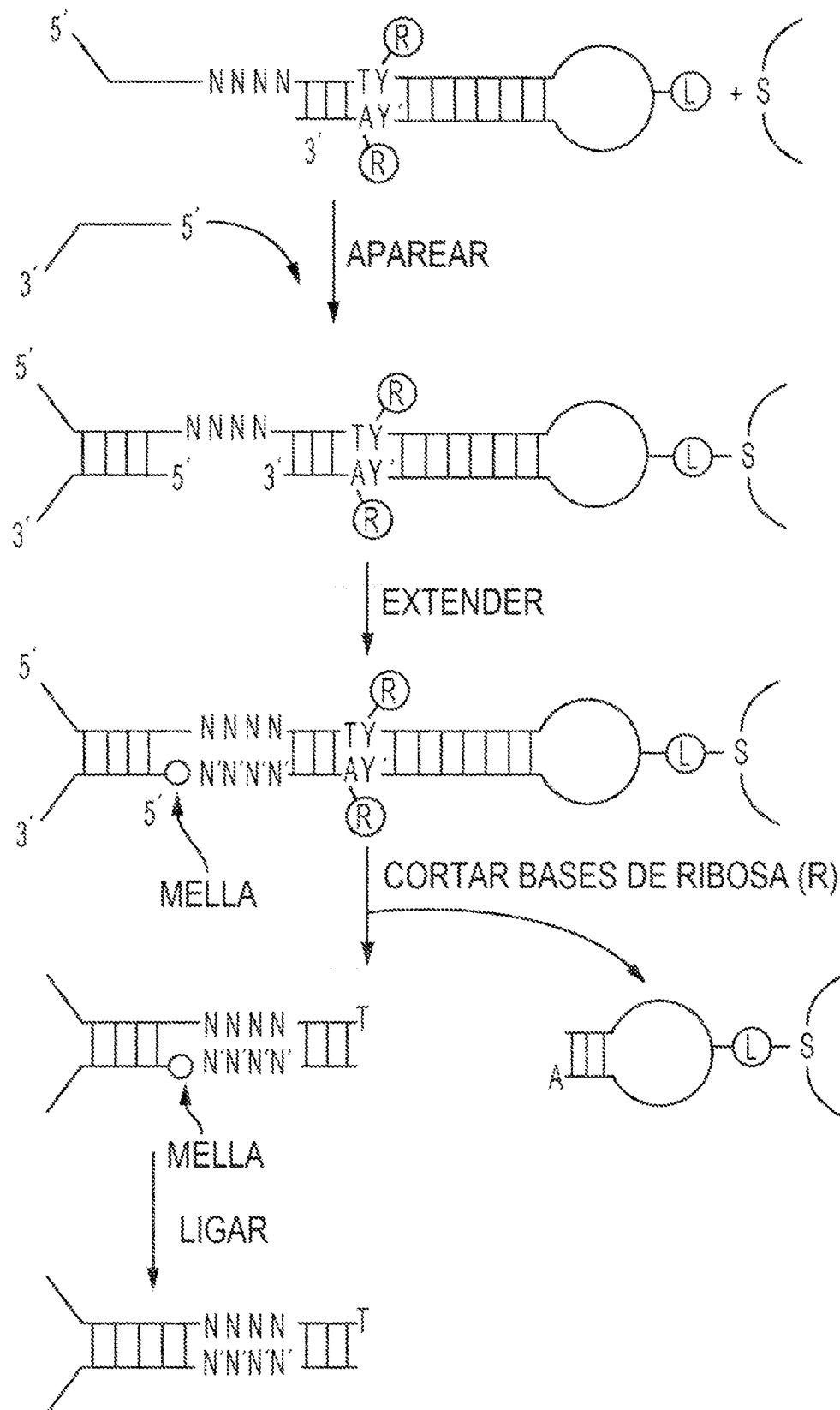


FIG. 12

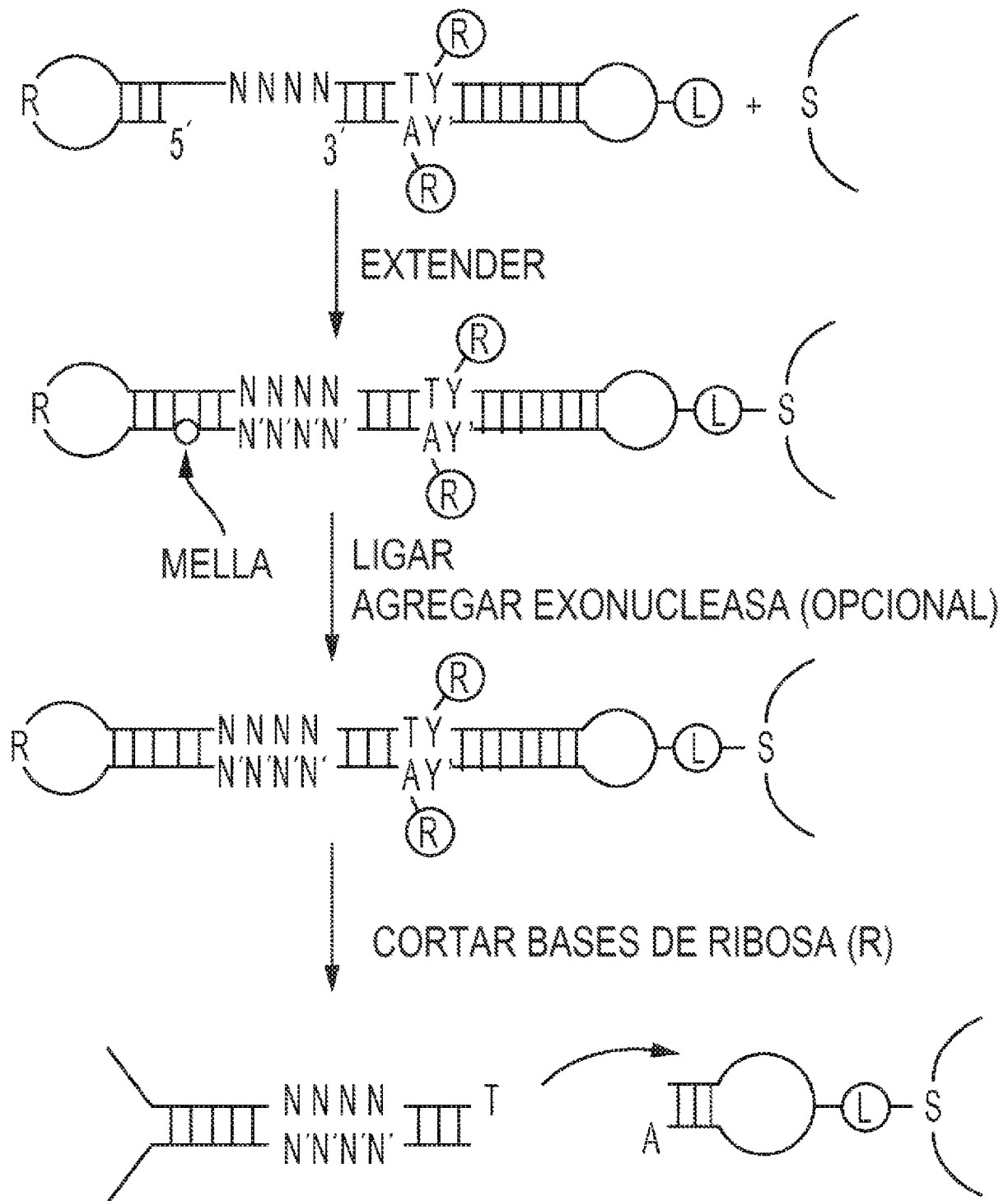


FIG. 13A

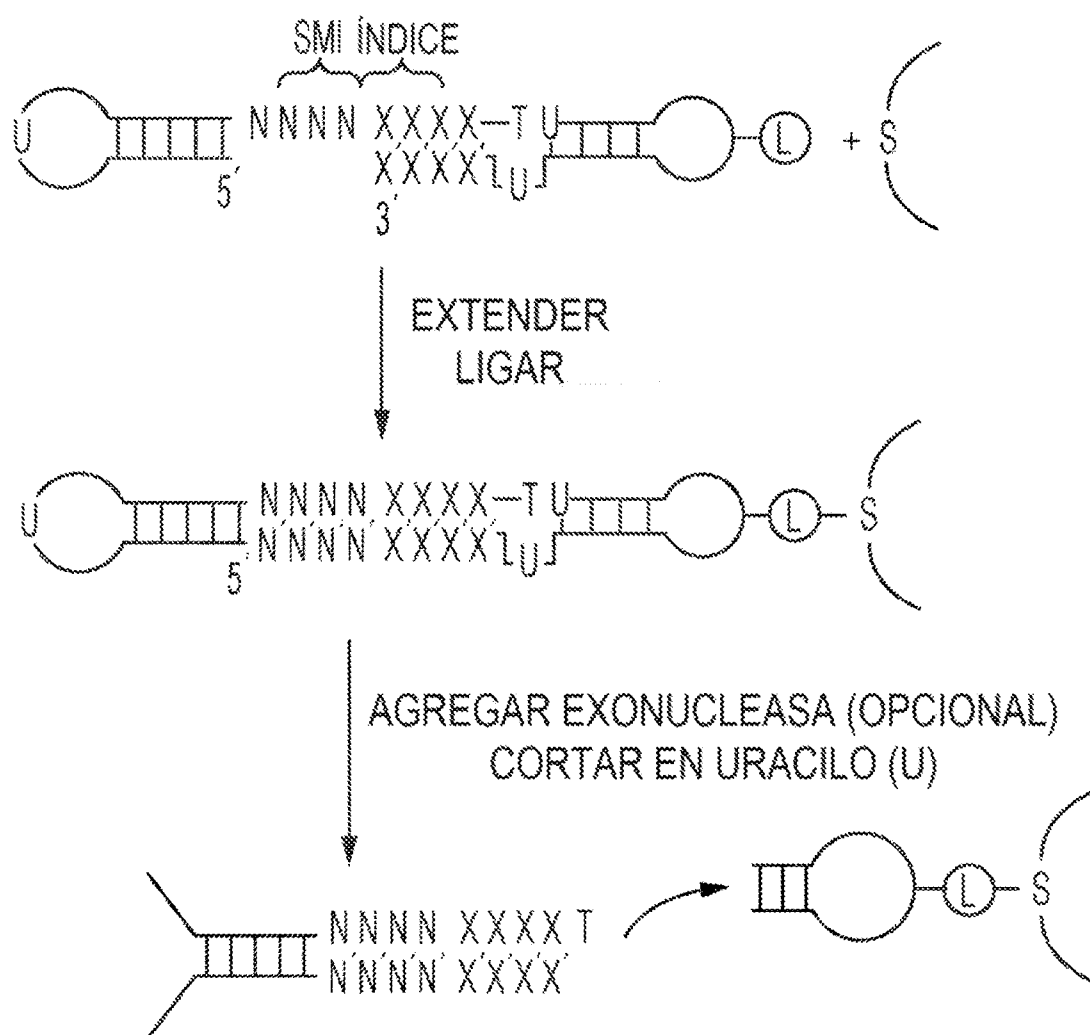


FIG. 13B

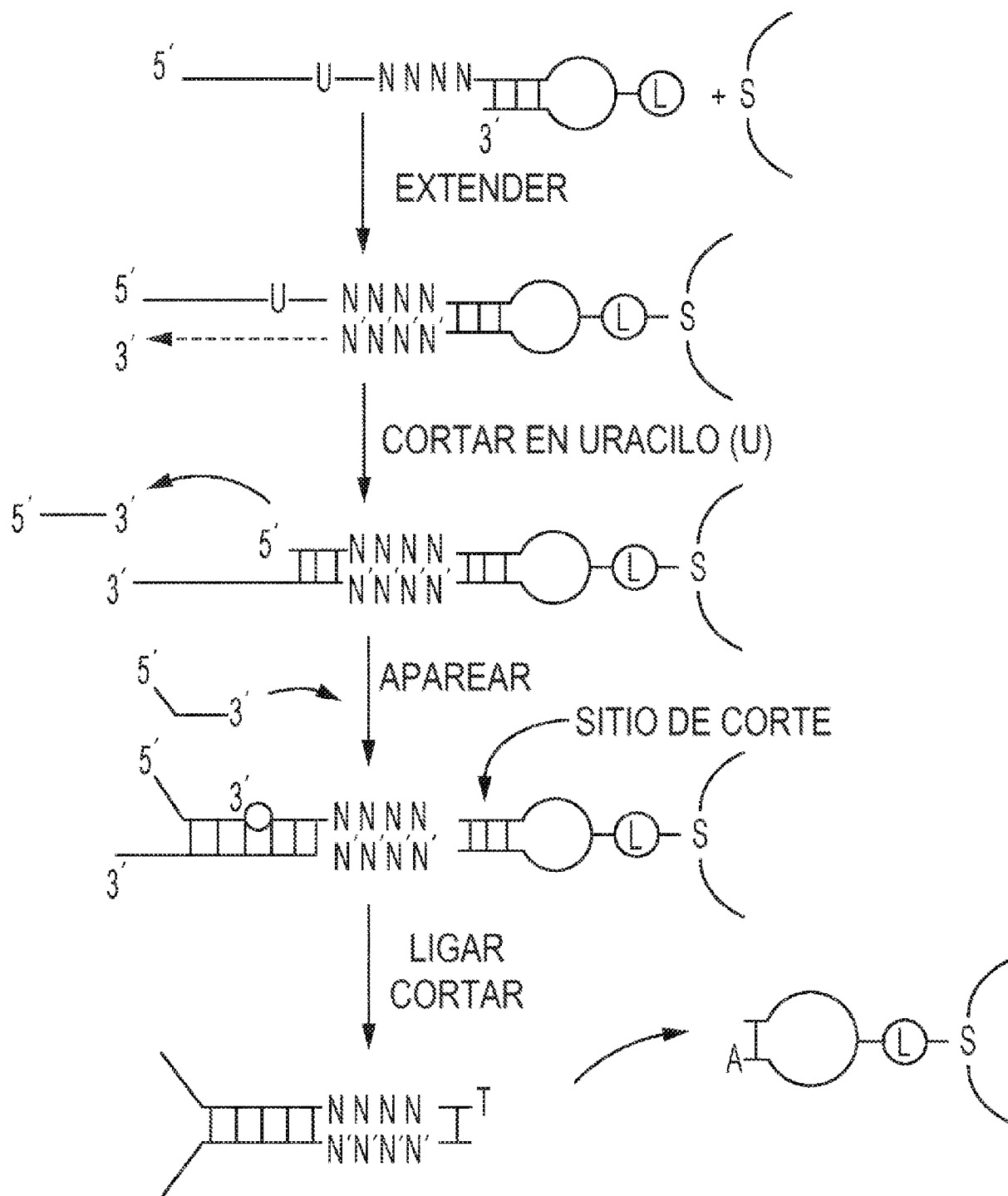


FIG. 13C

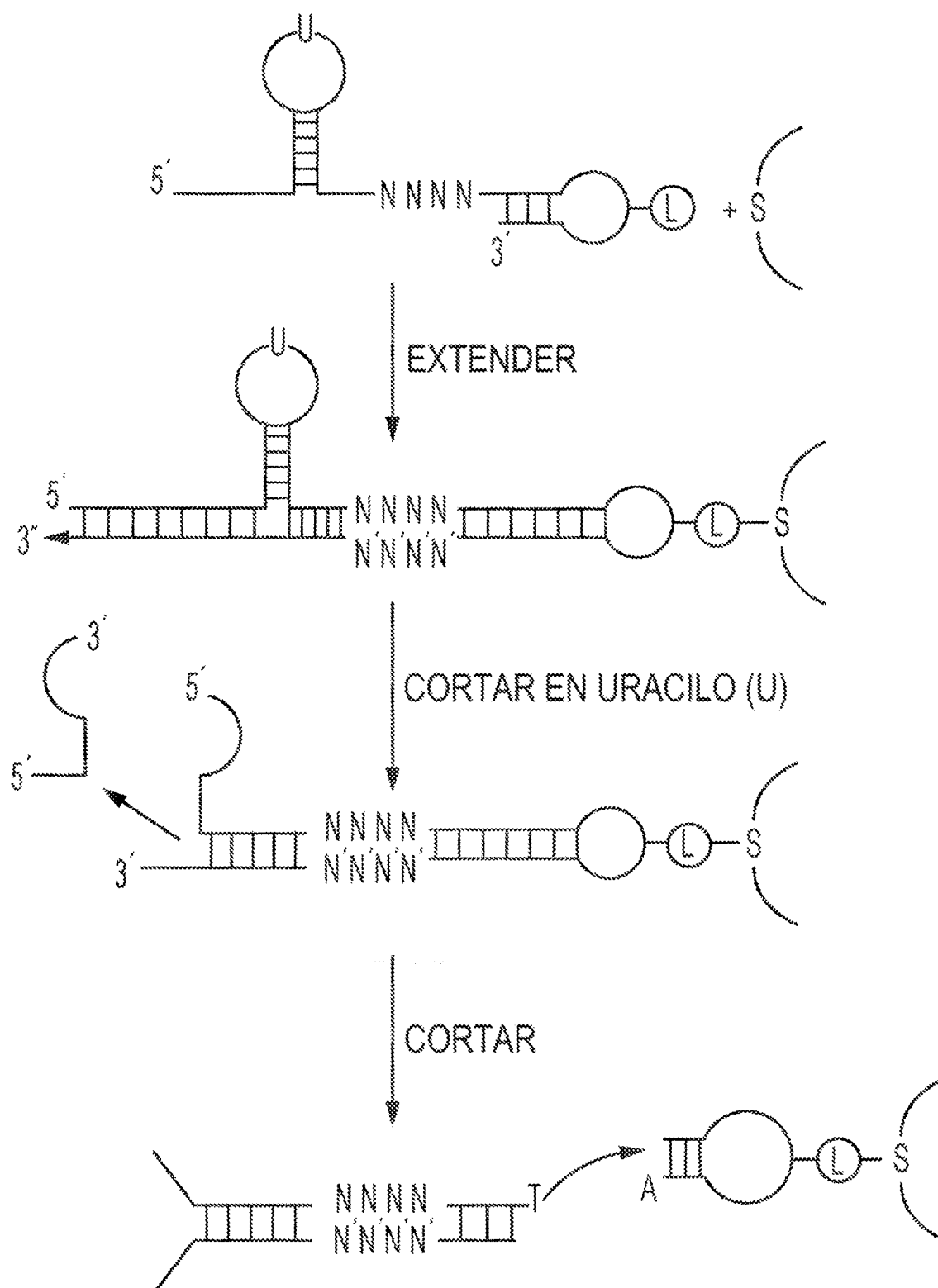


FIG. 13D