

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 743 029**

51 Int. Cl.:

**C12Q 1/68** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **24.10.2009 E 17171687 (1)**

97 Fecha y número de publicación de la concesión europea: **07.08.2019 EP 3272879**

54 Título: **Composiciones de extremo del transposón y métodos para modificar ácidos nucleicos**

30 Prioridad:

**24.10.2008 US 108321 P**

**24.10.2008 US 108326 P**

**24.10.2008 US 108329 P**

**16.02.2009 US 152868 P**

**25.02.2009 US 155431 P**

**05.06.2009 US 184530 P**

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

**18.02.2020**

73 Titular/es:

**EPICENTRE TECHNOLOGIES CORPORATION  
(100.0%)**

**5602 Research Park Boulevard, Suite 200  
Middleton, WI 53562-2345, US**

72 Inventor/es:

**JENDRISAK, JEROME;  
DAHL, GARY;  
GRUNENWALD, HAIYING LI y  
CARUCCIO, NICHOLAS**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

**ES 2 743 029 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

Composiciones de extremo del transposón y métodos para modificar ácidos nucleicos

Esta solicitud reivindica la prioridad de las solicitudes provisionales U. S. Nros. de serie 61/108.321, presentada el 24 de octubre de 2008; 61/108.326, presentada el 24 de octubre de 2008; 61/108.329, presentada el 24 de octubre de 2008; 61/155.431, presentada el 25 de febrero de 2009; y 61/184.530, presentada el 5 de junio de 2009.

Campo de la invención

La presente invención se refiere a métodos, composiciones y kits para usar transposasa y composiciones de extremo del transposón para generar una biblioteca de fragmentos de ADN rotulados de ADN objetivo. Los fragmentos de ssADN generados son de utilidad como plantillas, por ejemplo, para una variedad de aplicaciones, incluyendo, por ejemplo, secuenciación de ADN de alto rendimiento, masivamente paralela y/o múltiple.

Antecedentes de la invención

Hay una variedad de métodos y aplicaciones para los que se desea generar una biblioteca de moléculas de ADN fragmentadas y rotuladas de moléculas objetivo de ADN bicatenario (dsADN). A menudo, la finalidad consiste en generar moléculas de ADN monocatenario más pequeñas (ssADN) (por ejemplo, fragmentos de ADN) a partir de moléculas de dsADN más grandes para usar como plantillas en reacciones de ADN o ARN polimerasa (por ejemplo, para usar como plantillas en reacciones de secuenciación de ADN o en reacciones de amplificación de ADN o ARN en las que un cebador se funde con el rótulo y se extiende por una polimerasa).

Hasta recientemente, la mayor parte de las secuenciaciones de ADN se realizó usando el método de secuenciación de terminación de cadena de didesoxi de Sanger, en donde un cebador se extiende por medio de una polimerasa usando el ADN por secuenciar como una plantilla. Se realizan cuatro reacciones, cada una con una mezcla de todos los nucleótidos canónicos (dATP, dCTP, dGTP y dTTP) y uno de los cuatro didesoxinucleótidos que terminan en cadena (ddATP, ddCTP, ddGTP o ddTTP) y cada reacción produce un grupo anidado de fragmentos terminados en cadena que comienzan con el cebador y terminan con el didesoxinucleótido. Cuando estas moléculas de ADN que terminan en cadena se separan por tamaño después de electroforesis, el orden en que los ddNTPs se incorporaron refleja la secuencia del ADN de plantilla. Usando estos métodos, la secuencia se podría determinar para algunos cientos o miles de bases del sitio del cebador. La determinación de secuencias más grandes requería el empalme de la secuencia más larga junto con la información superpuesta de numerosos clones.

Como estos métodos tradicionales requieren de grandes cantidades de plantilla de ADN y como estos métodos producían pobres resultados si están presentes grandes cantidades de ADN no plantilla, la secuenciación de didesoxi de Sanger se lleva a cabo a menudo usando un ADN clonado o amplificado. Por ejemplo, la mayor parte de la secuenciación llevada a cabo durante el proyecto de genoma humano, que comenzó formalmente en 1990 y culminó con el anuncio de la terminación de un borrador en bruto de la secuencia del genoma humano en 2000 y la publicación de la secuencia del último cromosoma humano en 2006, se basó en el uso de bibliotecas genómicas que consistían en una población de bacterias huésped, cada una de las cuales portaba una molécula de ADN que estaba clonada en un vector de ADN, de modo que la colección de todos los clones de ADN, que llevaba cada uno una parte del ADN genómico, representaba el genoma completo. Este era un proceso tedioso y altamente iterativo, que implicaba la construcción y banca de grandes cantidades de clones de ADN (por ejemplo, clones BAC), que, a su vez, se subclonaron a menudo para generar bibliotecas de clones de ADN más pequeños que se usaron como plantillas de secuenciación. A menudo, los cebadores usados en estos métodos se diseñaron para fusionarse con el vector de modo tal que se extenderían en el ADN clonado desconocido durante las reacciones de secuenciación. Este enfoque permitió usar el mismo grupo de cebadores para analizar muchos clones diferentes.

A fin de reducir la cantidad de subclonación requerida para el proyecto de secuenciación del genoma humano, un método que a veces se usó era la "transposición in vitro". El método de transposición in vitro comprende el uso de elementos genéticos móviles denominados transposones para insertar una pequeña porción de ADN de secuencia conocida en el medio del ADN desconocido. El método comprende la incubación de un clon de ADN de una biblioteca genómica con un transposón en condiciones en las que se produce una simple inserción del transposón en el clon de ADN, luego se transforman las células de E. coli con una alícuota de la reacción de transposición in vitro y se seleccionan células que contenían un marcador como un marcador de resistencia a antibióticos, codificados por el transposón. Así, la reacción de transposición in vitro genera una biblioteca de "clones de inserción de transposones" del clon de ADN principal, cada uno de los cuales contiene el transposón insertado en una ubicación diferente en el clon de ADN. Cada clon de inserción se secuencia luego hacia fuera de cada extremo del transposón usando un diferente cebador para cada cadena de ADN. Tal como se describió con anterioridad, la secuencia completa del clon de ADN principal se construye superponiendo las secuencias obtenidas de diferentes clones de inserción. Los ejemplos del uso de este método de inserción de transposones para el proyecto de genoma humano fueron descritos por Butterfield, YSN et al., Nucleic Acids Res 30: 2460-2468, 2002; Shevchenko, Y et al., Nucleic Acids Res 30: 2469-2477, 2002; y Haapa, S et al., Genome Res 9: 308-315, 1999. El uso del proceso de transposición in vitro para el proyecto de genoma humano facilitó la secuenciación completa tanto de clones genómicos de ADN como de clones de cADN generados de mARN codificado por el ADN genómico. Sin embargo, una desventaja de este método de

transposición in vitro era que no era totalmente in vitro, dado que requería las 2 etapas de transformación de células de E. coli, selección de colonias de E. coli que contenían inserciones de transposones y luego aislamiento del ADN de los clones de inserción de transposones para secuenciación.

5 A fin de eliminar el requerimiento de transformar células de E. coli con una alícuota de la reacción de transposición in vitro y cultivo de las células de E. coli en medio selectivo para obtener clones de inserción de transposones, Teknanen et al. (Patente U. S. No. 6.593.113 desarrolló métodos totalmente a base de transposones in vitro que comprendían una reacción de transposición in vitro y una reacción de amplificación por PCR para seleccionar plantillas de secuenciación. De acuerdo con Teknanen et al., el ADN examinado o el ADN objetivo usado en sus métodos puede variar de algunos pares de bases hasta 40 pares de kilobases, siendo el único factor limitativo para no usar segmentos de ADN aún más largos como ADN objetivo la incapacidad de reacciones de amplificación, como PCR, para amplificar segmentos más largos. Así, en algunas formas de realización, para generar plantillas de secuenciación usando este método, el ADN examinado o ADN objetivo de hasta aproximadamente 40 Kb se somete primero a una reacción de transposición in vitro y luego se amplifica por PCR usando, como un primer cebador de PCR, un cebador fijo que es complementario a una secuencia conocida en el ADN objetivo o, si el ADN objetivo es clonado en un vector, un cebador fijo que es complementario a una secuencia en el vector y como un segundo cebador de PCR, un cebador selectivo que es complementario a una secuencia del extremo del transposón a la que se une el ADN objetivo, más, opcionalmente, uno a diez nucleótidos adicionales de identidad conocida en su extremo 3'. En otra forma de realización, dos cebadores selectivos se usan para la etapa de amplificación por PCR, donde al menos uno de ellos tiene uno a diez nucleótidos adicionales de identidad conocida en su extremo 3'. Los métodos de Teknanen et al. proporcionan ciertos beneficios para la secuenciación de Sanger porque eliminan la necesidad de utilizar células de E. coli para seleccionar moléculas de ADN que tienen inserciones de transposones. Sin embargo, estos métodos están limitados al ADN objetivo de un tamaño de hasta aproximadamente 40 Kb y, debido al uso de cebadores fijos o selectivos, los métodos seleccionan moléculas de ADN que exhiben sólo una parte de las secuencias exhibidas por el ADN objetivo. En consecuencia, a pesar de que estos métodos eran de utilidad para la secuenciación de Sanger, no son apropiados para generar plantillas de secuenciación para novedosos métodos de secuencias de ADN de "próxima generación", que son capaces de generar datos de secuencia de hasta millones de plantillas de secuenciación en una simple corrida de secuenciación usando un formato masivamente paralelo o múltiple.

Las plataformas de secuenciación de próxima generación incluyen 454 FLX™ o 454 TITANIUM™ (Roche), en analizador genómico SOLEXA™ (Illumina), el secuenciador molecular simple HELISCOPE™ (Helicos Biosciences) y el secuenciador de ADN SOLID™ (Life Technologies/Applied Biosystems) instruments), así como otras plataformas aún bajo desarrollo por compañías tales como Intelligent Biosystems y Pacific Biosystems. A pesar de que la química por la cual se genera información de secuencias varía para las diferentes plataformas de secuenciación de próxima generación, todas ellas comparten la característica común de generación de los datos de secuenciación de una gran cantidad de plantillas de secuenciación, en las que las reacciones de secuenciación se corren de forma simultánea. En general, los datos de todas estas reacciones de secuenciación se recolectan usando un escáner y luego ensamblándolos y analizándolos con ordenadores y potentes programas de bioinformática. Las reacciones de secuenciación se llevan a cabo, se leen, se ensamblan y se analizan de una manera "masivamente paralela" o "múltiple". La naturaleza masivamente paralela de estos instrumentos requirió un cambio de pensamiento acerca de qué tipo de plantillas de secuenciación son necesarias y de cómo generarlas a fin de obtener la máxima cantidad posible de datos de secuenciación de estos poderosos instrumentos. Así, más que requerir bibliotecas genómicas de clones de ADN en E. coli, ahora es necesario pensar en términos de sistemas in vitro para generar bibliotecas de fragmentos de ADN que comprendan una colección o población de fragmentos de ADN generados de ADN objetivos en una muestra, en donde la combinación de todos los fragmentos de ADN en la colección o población exhibe secuencias que son cualitativa y/o cuantitativamente representativas de la secuencia del ADN objetivo de la que se generaron los fragmentos de ADN. De hecho, en algunos casos, es necesario pensar en términos de generación de bibliotecas de fragmentos de ADN que consisten en múltiples bibliotecas de fragmentos de ADN genómicos múltiples, cada uno de los cuales está rotulado con un rótulo de dirección o código de barras para permitir la identificación de la fuente de cada fragmento secuenciado.

En general, estos métodos de secuenciación requieren la fragmentación de ADN genómico o cADN bicatenario (preparado a partir de ARN) en fragmentos de ssADN más pequeños y adición de rótulos a al menos una cadena o preferentemente ambas cadenas de los fragmentos de ssADN. En algunos métodos, los rótulos proporcionan sitios de cebado para secuenciación de ADN usando una ADN polimerasa. En algunos métodos, los rótulos también proporcionan sitios para capturar los fragmentos sobre una superficie, como una perla (por ejemplo, antes de la amplificación por PCR en emulsión para algunos de estos métodos; por ejemplo, usando métodos tal como se describen en la patente U. S. No 7.323.305). En la mayoría de los casos, las bibliotecas de fragmentos de ADN usados como plantillas para la secuenciación de próxima generación comprenden fragmentos de ADN 5' y 3' rotulados o "fragmentos de ADN dirrotulados". En general, los métodos corrientes para generar bibliotecas de fragmentos de ADN para la secuenciación de próxima generación comprenden fragmentar el ADN objetivo que se desea secuenciar (por ejemplo, ADN objetivo que comprende ADN genómico o cADN bicatenario después de la transcripción inversa de ARN) usando un sonicador, nebulizador o una nucleasa y uniendo (por ejemplo, por ligación) oligonucleótidos consistentes en adaptadores o rótulos con los extremos 5' y 3' de los fragmentos.

Hay una cantidad de problemas e ineficiencias con los métodos corrientes para generar plantillas de secuenciación de próxima generación, tal como se ilustra por medio del flujo de trabajo usado en el Wellcome Trust Sanger Institute,

uno de los centros genómicos más grandes del mundo (por ejemplo, descrito en Quail, MA et al., Nature Methods 5: 1005-1010, 2008). Por ejemplo, Quail et al. hallaron que la nebulización de ADN genómico para la secuenciación dio la pérdida de aproximadamente la mitad del ADN en masa y sólo aproximadamente el 5% del ADN original consistía en fragmentos en aproximadamente 200-bp de rango de tamaño deseado para la secuenciación usando el analizador genómico Illumina. Hallaron que un método alternativo, denominado "adapted focused acoustics" daba mejores rendimientos de ADN fragmentado y aproximadamente el 17% del ADN original consistía en fragmentos de 200-bp de rango de tamaño deseado, pero incluso este proceso es un desperdicio en términos de la muestra o del ADN objetivo. Además, los fragmentos de ADN resultantes a menudo requieren de la selección de tamaño por electroforesis en gel y etapas adicionales para rotular los fragmentos de ADN seleccionados por tamaño, que es difícil, laborioso, insume tiempo y es costoso.

Así, muchos de los métodos usados actualmente para la fragmentación y la rotulación del ADN bicatenario para usar en la secuenciación de próxima generación son un desperdicio del ADN, requieren instrumentos costosos para la fragmentación y los procedimientos para la fragmentación, rotulación y recuperación de fragmentos de ADN rotulados son difíciles, tediosos, laboriosos, insumen tiempo, son costosos, requieren cantidades relativamente grandes de ácidos nucleicos de muestra. Además, muchos de estos métodos generan fragmentos de ADN rotulados que no son completamente representativos de las secuencias contenidas en los ácidos nucleicos de muestra de los que se generaron. Así, lo necesario en el arte son métodos para generar bibliotecas de fragmentos de ADN dirrotulados de una manera masivamente paralela que superan las limitaciones de los métodos actuales.

Algunos de los métodos de secuenciación de próxima generación usan sustratos de ssADN circulares en su proceso de secuenciación. Por ejemplo, las solicitudes de patentes U. S. Nros. 20090011943; 20090005252; 20080318796; 20080234136; 20080213771; 20070099208; y 20070072208 de Drmanac et al., revela la generación de plantillas de ssADN circulares para secuenciación de ADN masivamente paralela. La solicitud de patente U. S. No. 20080242560 de Gunderson y Steemers revela métodos que comprenden: preparación de bolas de ADN digitales (ver, por ejemplo, FIG. 8 en la solicitud de patente U. S. No. 20080242560); y/o clivaje específico del sitio y amplificación de ADN, como ADN genómico, incluyendo amplificación por amplificación de desplazamiento múltiple o amplificación de genoma completo (por ejemplo, su FIG 17) o por RCA hiperramificado (por ejemplo, su FIG 18) para generar matrices de ácidos nucleicos amplificados (por ejemplo, ILLUMINA BeadArrays™; ILLUMINA, San Diego CA, Estados Unidos).

Lo que se necesitan son métodos mejorados, composiciones y kits para preparar fragmentos de ssADN circulares rotulados de ADN de una muestra biológica (por ejemplo, de ADN genómico o ADN mitocondrial o ADN episomal, incluyendo ADN clonado en un plásmido, BAC, fósido u otro vector episomal) para usar en métodos de amplificación o de secuenciación de ADN (como los métodos descritos en las solicitudes de patentes U. S. Nros. 20090011943; 20090005252; 20080318796; 20080234136; 20080213771; 20070099208; y 20070072208 de Drmanac et al.; o en la solicitud de patente U. S. No. 20080242560 de Gunderson y Steemers o por Turner et al. de Pacific Biosciences y publicados en su sitio Web en [www.pacificbiosciences.com](http://www.pacificbiosciences.com)).

Además, algunos métodos para amplificación, como amplificación de genoma completo, también requieren fragmentación y rotulación de ADN genómico. Algunos de estos métodos se reseñan en: Whole Genome Amplification, ed. por S. Hughes y R. Lasken, 2005, Scion Publishing Ltd (en la Web mundial en [scionpublishing.com](http://scionpublishing.com)).

Se necesitan métodos mejorados para generar bibliotecas de fragmentos de ADN de moléculas de ADN objetivo para amplificación, incluyendo amplificación de genomas completos o parciales de un organismo (por ejemplo, de una muestra clínica) o de múltiples organismos (por ejemplo, ADN metagenómico objetivo de una muestra medioambiental), para posterior análisis (por ejemplo, por PCR en tiempo real, PCR en emulsión, hibridación genómica comparativa (CGH), secuenciación genómica comparativa (CGS) o para preparar sondas rotuladas específicas de ADN (por ejemplo, sondas específicas de cromosomas, por ejemplo, pinturas cromosómicas o, por ejemplo, sondas específicas de genes, por ejemplo, por hibridación fluorescente in situ (FISH), para una variedad de fines (por ejemplo, con fines investigativos, diagnósticos e industriales).

Así, lo que se necesita en la técnica son métodos mejores y más eficientes para preparar bibliotecas de fragmentos de ADN rotulados de ADN objetivo para usar en métodos de análisis de ácidos nucleicos tales como métodos de secuenciación de próxima generación y amplificación. Se necesitan métodos para generar bibliotecas de fragmentos de ADN que no requieran instrumentos especializados y que sean más simples, más rápidos, que requieran menos manos a la vez, se puedan realizar con menores muestras de ADN y menores volúmenes, sean eficientes en la rotulación de uno o ambos extremos de los fragmentos y generen fragmentos de ADN rotulados que sean cualitativa y cuantitativamente representativos de los ácidos nucleicos objetivo en la muestra de la que se generan.

#### Síntesis de la invención

La presente invención se refiere a métodos, composiciones y kits para tratar ácido nucleico y en particular, métodos y composiciones para fragmentar y rotular ADN usando composiciones de transposón. Los métodos y composiciones de la presente invención son de utilidad, por ejemplo, para generar bibliotecas de fragmentos de ADN rotulados para uso, por ejemplo, en métodos de secuenciación de próxima generación, hibridación fluorescente in situ, y similares. En algunas formas de realización de preferencia, la presente invención se refiere a la preparación de fragmentos de ssADN lineales o fragmentos de ssADN circulares rotulados (y sus productos de amplificación) de ADN objetivo que

comprende cualquier dsADN de interés (incluyendo cADN bicatenario preparado a partir de ARN), de cualquier fuente, para el análisis genómico, subgenómico, transcriptómico o metagenómico o el análisis de la expresión de ARN.

5 En algunas formas de realización, la presente invención proporciona métodos para generar una biblioteca de fragmentos de ADN rotulados de un ADN objetivo, que comprende la incubación del ADN objetivo con una transposasa y un extremo de transposón o composición de extremo de transposón que comprende una cadena transferida que tiene un dominio de rótulo en su parte 5', en condiciones en las que una reacción de transposición es catalizada por la transposasa y en donde el ADN objetivo se fragmenta para generar una pluralidad de fragmentos de ADN objetivo y una cadena transferida del extremo del transposón o composición de extremo del transposón se une con los extremos 5' de cada uno de una pluralidad de los fragmentos de ADN objetivo, para producir una pluralidad de fragmentos de ADN objetivo rotulado en 5'. En algunas realizaciones, los métodos también comprenden la incubación de la pluralidad de fragmentos de ADN objetivo con rótulo en 5' con al menos una enzima modificadora del ácido nucleico en condiciones en las que un rótulo 3' se une con un extremo 3' lejos del fragmento de ADN objetivo 5' para producir fragmentos de ADN objetivo dirrotulados.

10 En algunas realizaciones, la presente invención proporciona métodos de rotulación de un fragmento de un ADN objetivo, que comprende la incubación de ADN objetivo con una transposasa y un extremo del transposón o composición de extremo del transposón que comprende una cadena transferida que comprende un dominio de rótulo en su parte 5', en condiciones en las que una reacción de transposición es catalizada por la transposasa, en donde el ADN objetivo se fragmenta y la cadena transferida del extremo del transposón o la composición de extremo del transposón se une con un extremo 5' de un fragmento del ADN objetivo para producir un fragmento de ADN objetivo con rótulo en 5'. En algunas realizaciones, los métodos también comprenden la incubación del fragmento de ADN objetivo con rótulo en 5' con una enzima modificadora de ácido nucleico en condiciones en las que un rótulo 3' se une con un extremo 3' del fragmento de ADN objetivo con rótulo en 5' para producir un fragmento ADN objetivo dirrotulado. Los métodos se limitan al uso de cualquier enzima modificadora de ácido nucleico particular. Por ejemplo, enzimas modificadoras de ácido nucleico comprenden polimerasas, nucleasas, ligasas, y similares. En algunas realizaciones, la enzima modificadora de ácido nucleico comprende una ADN polimerasa y el rótulo 3' se forma por extensión del extremo 3' del fragmento de ADN objetivo de rótulo en 5'. En algunas formas de realización, la ADN polimerasa comprende una ADN polimerasa dependiente de la plantilla y en algunas formas de realización, la ADN polimerasa comprende una ADN polimerasa independiente de la plantilla. En algunas formas de realización de preferencia, la ADN polimerasa es una ADN polimerasa dependiente de la plantilla que tiene un desplazamiento de la cadena y/o actividad de 5' nucleasa.

15 En algunas formas de realización, una enzima modificadora del ácido nucleico usada en los presentes métodos es una ligasa y el rótulo 3' se forma por ligamiento de un oligonucleótido con el extremo 3' del fragmento de ADN objetivo de rótulo en 5'. En algunas formas de realización, la ligasa comprende una ligasa dependiente de la plantilla, si bien en algunas formas de realización, la ligasa comprende una ligasa independiente de la plantilla.

20 En algunas formas de realización, los extremos transferidos comprenden dominios de rótulo. En realizaciones preferidas, los dominios de rótulo que comprenden uno o varios de un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de amplificación, un dominio de rótulo de detección, un dominio de rótulo de direccionamiento y un dominio del promotor de la transcripción. En algunas formas de realización, los dominios de rótulo son dominios de rótulo de secuenciación que comprenden o consisten en rótulos de secuenciación seleccionados de rótulos de secuenciación de Roche 454A y 454B, rótulos de secuenciación ILLUMINA™ SOLEXA™, rótulos de secuenciación SOLID™ de Applied Biosystems, rótulos de secuenciación SMRT™ de Pacific Biosciences, rótulos de secuenciación Pollonator Polony o rótulos de secuenciación de Complete Genomics.

25 Algunas formas de realización de la divulgación también comprenden la amplificación de uno o varios fragmentos de ADN objetivo con rótulo en 5' y/o fragmentos de ADN objetivo dirrotulados. En algunas realizaciones preferidas de la divulgación, la amplificación comprende el uso de uno o varios de una reacción de amplificación por PCR, una reacción de amplificación de desplazamiento de cadena, una reacción de amplificación de círculo de rodadura, una reacción de ligasa en cadena, una reacción de amplificación mediada por transcripción o una reacción de amplificación mediada por bucle. En ciertas formas de realización preferidas de la divulgación, la amplificación comprende la amplificación no selectiva de fragmentos de ADN objetivo con rótulo en 5' que comprende una biblioteca de fragmentos de ADN o fragmentos de ADN objetivo dirrotulados que comprenden una biblioteca de fragmentos de ADN.

30 En algunas formas de realización de la presente invención, la composición de extremo del transposón usada en la rotulación de un fragmento o biblioteca comprende una pluralidad de cadenas transferidas que difieren en la secuencia de ácidos nucleicos en al menos un nucleótido y la amplificación comprende la amplificación selectiva de fragmentos de ADN dirrotulados en base a las secuencias de ácidos nucleicos de los rótulos de extremo 5' o dominios de rotulación. En otras formas de realización de la divulgación, la amplificación comprende una reacción de polimerasa en cadena usando un único cebador de oligonucleótido que es complementario al rótulo 3' de los fragmentos de ADN objetivo dirrotulados.

35 En algunas formas de realización de la divulgación, la amplificación comprende una reacción de amplificación de desplazamiento de cadena usando un único cebador de oligonucleótido, en donde el cebador de oligonucleótido

consiste sólo en ribonucleótidos o consiste sólo en ribonucleótidos de purina y sólo 2'-F-2'-desoxirribonucleótidos de pirimidina y la reacción de amplificación de desplazamiento de cadena comprende una ADN polimerasa con cadena desplazada y una ribonucleasa H.

5 En algunas formas de realización de la divulgación, la amplificación comprende una reacción de polimerasa en cadena usando un primer y un segundo cebador de oligonucleótido, donde cada uno comprende porciones terminales 3', en donde al menos la porción terminal 3' del primer cebador de PCR es complementario al rótulo 3' de los fragmentos de ADN objetivo dirrotulados y en donde al menos una porción terminal 3' del segundo cebador de PCR exhibe la secuencia de al menos una porción de un rótulo en 5' o dominio de rótulo de los fragmentos de ADN objetivo dirrotulados. En ciertas formas de realización preferidas de la divulgación, el primer o segundo cebador de oligonucleótido comprende una porción terminal 5' en donde al menos la porción terminal 5' del primer cebador no es complementaria al rótulo 3' de los fragmentos de ADN objetivo dirrotulados o en donde la porción 5' del segundo cebador no exhibe la secuencia de al menos una porción del rótulo en 5' o dominio de rótulo de los fragmentos de ADN objetivo dirrotulados. En formas de realización particularmente preferidas de la divulgación, el primer y el segundo cebador de oligonucleótido comprenden cada uno porciones terminales 5', en donde al menos la porción terminal 5' del primer cebador de PCR no es complementaria al rótulo 3' de los fragmentos de ADN objetivo dirrotulados, y/o en donde la porción terminal 5' del segundo cebador de PCR no exhibe la secuencia de al menos una porción del dominio de rótulo en 5' de los fragmentos de ADN objetivo dirrotulados.

20 En algunas formas de realización, la presente invención proporciona métodos de generación de una población de fragmentos de ADN monocatenarios circulares rotulados de un ADN objetivo. En determinadas formas de realización, esto comprende la incubación del ADN objetivo con una transposasa y un extremo del transposón o una composición de extremo del transposón que comprende una cadena transferida que tiene un dominio de rótulo en su parte 5' y un extremo del transposón en su porción 3', en condiciones en las que una reacción de transposición es catalizada por la transposasa, de modo que el ADN objetivo se fragmenta para generar una pluralidad de fragmentos de ADN objetivo y la cadena transferida del extremo del transposón o la composición de extremo del transposón se une con el extremo 5' de cada uno de la pluralidad de los fragmentos de ADN objetivo para producir una población de fragmentos de ADN objetivo con rótulo en 5'. Estos métodos también comprenden etapas de desnaturalización de los fragmentos de ADN objetivo con rótulo en 5' para producir fragmentos de ADN objetivo monocatenarios con rótulo en 5' e incubación de los fragmentos de ADN objetivo monocatenarios con rótulo en 5' con una ligasa de ácido nucleico en condiciones en las que los fragmentos de ADN objetivo monocatenarios con rótulo en 5' se ligan intramolecularmente para formar fragmentos de ADN monocatenarios circulares rotulados, donde cada uno exhibe secuencias de la cadena transferida y una porción del ADN objetivo.

35 En algunas formas de realización divulgadas, se desea clivar este ADN monocatenario circular. Así, en algunas formas de realización de los métodos de la presente invención, el dominio de rótulo exhibe una secuencia o estructura de un sitio de clivaje y el método también comprende: incubación de los fragmentos de ADN monocatenarios circulares rotulados con al menos una enzima que comprende una composición de enzima de clivaje en donde la composición de enzima de clivaje cliva los fragmentos de ADN monocatenarios circulares rotulados para producir fragmentos de ADN monocatenarios lineales dirrotulados. En ciertas formas de realización preferidas, la composición de enzima de clivaje comprende una enzima de restricción. En algunas formas de realización, el dominio de rótulo exhibe una secuencia de un sitio de restricción y el método también comprende la fusión con los fragmentos de ADN monocatenarios circulares rotulados de oligonucleótidos complementarios al dominio de rótulo y la incubación de los fragmentos de ADN monocatenarios circulares rotulados con una endonucleasa de restricción que reconoce el sitio de restricción, en donde la endonucleasa de restricción cliva los fragmentos de ADN monocatenarios circulares rotulados para producir fragmentos de ADN monocatenarios lineales dirrotulados.

45 En algunas formas de realización de la divulgación, es útil amplificar los fragmentos y bibliotecas de la invención. Así, algunas formas de realización también comprenden la amplificación de uno o varios de los fragmentos de ADN monocatenarios circulares rotulados y/o fragmentos de ADN monocatenarios lineales dirrotulados. En ciertas formas de realización preferidas, la amplificación comprende una reacción de polimerasa en cadena usando un primer y un segundo cebador de oligonucleótido, donde cada uno comprende porciones terminales 3', en donde al menos la porción terminal 3' del primer cebador de PCR es complementario a al menos una porción de la secuencia de la cadena transferida en los fragmentos de ADN monocatenarios circulares rotulados o en los fragmentos de ADN monocatenarios lineales dirrotulados y en donde al menos la porción terminal 3' del segundo cebador de PCR es complementaria al menos una porción del complemento de la cadena transferida en los fragmentos de ADN monocatenarios circulares rotulados o en los fragmentos de ADN monocatenarios lineales dirrotulados.

55 En algunas formas de realización en los que los fragmentos de ADN monocatenarios circulares rotulados o los fragmentos de ADN monocatenarios lineales dirrotulados se amplifican, los primeros y segundos cebadores de oligonucleótido comprenden cada uno porciones terminales 5', en donde la porción terminal 5' del primer cebador de PCR no es complementaria a la secuencia de la cadena transferida en los fragmentos de ADN monocatenarios circulares rotulados o en los fragmentos de ADN monocatenarios lineales dirrotulados y en donde la porción terminal 5' del segundo cebador de PCR no es complementaria al complemento de la cadena transferida en los fragmentos de ADN monocatenarios circulares rotulados o en los fragmentos de ADN monocatenarios lineales dirrotulados.

60 En algunas formas de realización, la divulgación proporciona métodos de generación de una población de fragmentos

de ADN circulares rotulados de un ADN objetivo, que comprende la incubación de ADN objetivo con una transposasa y una composición de extremo del transposón en horquilla que comprende un oligonucleótido que exhibe una secuencia de cadena no transferida en su extremo 5', una secuencia de cadena transferida en su extremo 3' y una secuencia de bucle interviniente que comprende un dominio de rótulo, en condiciones en las que el oligonucleótido puede formar un tallo-bucle intramolecular y en donde una reacción de transposición es catalizada por la transposasa, de modo tal que el ADN objetivo se fragmenta para generar una pluralidad de fragmentos de ADN objetivo y el oligonucleótido de la composición de extremo del transposón en horquilla se une con el extremo 5' de cada uno de la pluralidad de fragmentos de ADN objetivo para producir una población de fragmentos de ADN objetivo con rótulo en 5'. En algunas formas de realización, el método también comprende el llenado de brechas y muescas de ligamiento en las moléculas del fragmento. En algunas formas de realización, esto comprende la incubación de la población de fragmentos de ADN objetivo con rótulo en 5' con una ligasa dependiente de la plantilla y una ADN polimerasa que carece de 5' a 3'-exonucleasa, 3' a 5'-exonucleasa y actividades de desplazamiento de cadena o uno o varios tamaños de oligonucleótidos de secuencia aleatoria que, solos o en combinación, tienen la misma longitud que las brechas monocatenarias en los fragmentos de ADN con rótulo en 5' que resultan después de una reacción de transposición con la transposasa y la composición de extremo del transposón en horquilla, en condiciones en las que las brechas monocatenarias en los fragmentos de ADN objetivo con rótulo en 5' se completan y el extremo 3' de cada fragmento de ADN con rótulo en 5' se une con el extremo 5' de otro fragmento de ADN con rótulo en 5' que comprende una porción complementaria del ADN objetivo para formar fragmentos de ADN circulares rotulados que comprenden el dominio de rótulo en estructuras de bucle y ambas cadenas de una porción del ADN objetivo. En algunas formas de realización preferidas, el llenado y la unión comprenden la incubación de fragmentos de ADN con rótulo en 5' con la ADN polimerasa en condiciones en las que el extremo 3' de cada fragmento de ADN con rótulo en 5' se extiende para formar una población de productos de extensión de fragmentos de ADN con rótulo en 5' y la incubación de productos de extensión de fragmentos de ADN con rótulo en 5' con la ligasa dependiente de la plantilla en condiciones en las que los productos de extensión de fragmentos de ADN con rótulo en 5' se ligan, generando así los fragmentos de ADN circulares rotulados. En formas de realización particularmente preferidas, la ADN polimerasa y la ligasa se proporcionan en una mezcla y el llenado y el ligamiento se llevan a cabo en una única mezcla de reacción.

En algunas formas de realización, las etapas de llenado y ligamiento comprenden la incubación de los fragmentos de ADN rotulados en 5' con uno o varios tamaños de oligonucleótidos de secuencia aleatoria y la ligasa dependiente de la plantilla en condiciones en las que los oligonucleótidos de secuencia aleatoria se fusionan y se rellenan las brechas monocatenarias y se ligan entre sí o con extremos adyacentes de fragmentos de ADN rotulados en 5' para formar fragmentos de ADN circulares rotulados.

En formas de realización preferidas, el método también comprende la separación de los fragmentos de ADN circulares rotulados de ADN lineal, oligonucleótidos de secuencia aleatoria no ligados, y/o la composición de extremo del transposón en horquilla no unida con el ADN objetivo. En formas de realización particularmente preferidas, la mezcla de reacción que contiene los fragmentos de ADN circulares rotulados se trata con T5 exonucleasa para eliminar el ADN lineal, como fragmentos no ligados y oligonucleótidos de secuencia aleatoria.

En algunas formas de realización, se desea clivar las moléculas de ADN circulares rotuladas. Por ejemplo, en algunas formas de realización, el método también comprende una etapa de: clivaje de fragmentos de ADN circulares rotulados en las estructuras de bucle para generar fragmentos de ADN bicatenarios de cola de milano, donde cada cadena tiene una porción del rótulo en su extremo 5' y una porción del rótulo en su extremo 3'. Así, el dominio de rótulo en las estructuras de bucle exhibe una secuencia o estructura de un sitio de clivaje y el método también comprende: incubación de los fragmentos de ADN monocatenarios circulares rotulados con al menos una enzima que comprende una composición de enzima de clivaje en donde la composición de enzima de clivaje cliva los fragmentos de ADN monocatenarios circulares rotulados para producir fragmentos de ADN bicatenarios de cola de milano.

En algunas formas de realización, la composición de enzima de clivaje comprende una N-glicosilasa y una AP endonucleasa. En ciertas formas de realización preferidas, la enzima de clivaje es una N-glicosilasa seleccionada de entre uracil-N-glicosilasa y una AP endonucleasa y proteína FPG y la AP endonucleasa se selecciona de entre endonucleasa III o endonucleasa IV de *E. coli*.

En ciertas formas de realización preferidas, la composición de enzima de clivaje comprende una enzima de restricción. En algunas formas de realización, el dominio de rótulo exhibe una secuencia de un sitio de restricción y el método también comprende la fusión con los fragmentos de ADN monocatenarios circulares rotulados oligonucleótidos complementarios al dominio de rótulo y la incubación de los fragmentos de ADN monocatenarios circulares rotulados con una endonucleasa de restricción que reconoce el sitio de restricción, en donde la endonucleasa de restricción cliva los fragmentos de ADN monocatenarios circulares rotulados para producir fragmentos de ADN bicatenarios de cola de milano, donde cada cadena tiene una porción del rótulo en su extremo 5' y una porción del rótulo en su extremo 3'.

Algunas formas de realización comprenden adicionalmente la desnaturalización de los fragmentos de ADN bicatenarios de cola de milano para generar fragmentos de ADN monocatenarios lineales dirrotulados.

Las formas de realización circular y de cola de milano se usan en métodos que comprenden el uso de los fragmentos de ADN como plantillas en un método de secuenciación de ADN o una reacción de amplificación. Así, en algunas formas de realización, los métodos también comprenden la amplificación de fragmentos de ADN circulares rotulados,

fragmentos de ADN bicatenarios de cola de milano y/o fragmentos de ADN monocatenarios lineales dirrotulados. En formas de realización preferidas, la amplificación comprende el uso de uno o varios de una reacción de amplificación por PCR, una reacción de amplificación de desplazamiento de cadena, una reacción de amplificación de círculo de rodadura, una reacción de ligasa en cadena, una reacción de amplificación mediada por transcripción o una reacción de amplificación mediada por bucle. En formas de realización particularmente preferidas, la divulgación, la amplificación comprende una reacción de polimerasa en cadena usando un primer y un segundo cebador de oligonucleótido, donde cada uno comprende porciones terminales 3', en donde al menos la porción terminal 3' del primer cebador de PCR es complementario a al menos una porción del dominio de rótulo y en donde al menos una porción 3' del segundo cebador de PCR exhibe la secuencia de al menos una porción del dominio de rótulo. En algunas formas de realización, los primeros y segundos cebadores de oligonucleótido comprenden cada uno porciones terminales 5', en donde la porción terminal 5' del primer cebador de PCR no es complementaria a la secuencia de rótulo y en donde la porción terminal 5' del segundo cebador de PCR no exhibe la secuencia del dominio de rótulo.

Formas de realización predferidas de cualquiera de las amplificaciones por PCR descritas con anterioridad comprenden amplificaciones en las que las porciones terminales 5' del primero y/o el segundo cebador de PCR exhiben dominios de rótulo. En formas de realización todavía más preferidas, los dominios de rótulo comprenden uno o varios de un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de amplificación, un dominio de rótulo de detección, un dominio de rótulo de direccionamiento y un dominio de promotor de la transcripción.

En formas de realización de particular preferencia de los métodos descritos en la presente, los dominios de rótulo son dominios de rótulo de secuenciación que comprenden o consisten en rótulos de secuenciación seleccionados de rótulos de secuenciación de Roche 454A y 454B, rótulos de secuenciación ILLUMINA™ SOLEXA™, rótulos de secuenciación SOLID™ de Applied Biosystems, rótulos de secuenciación SMRT™ de Pacific Biosciences, rótulos de secuenciación de Pollonator Polony o rótulos de secuenciación de Complete Genomics.

En algunas formas de realización, la presente invención proporciona métodos de rotulación de un fragmento de un ADN objetivo, que comprende la incubación de ADN objetivo con una transposasa y una composición de extremo del transposón que comprende una cadena transferida que exhibe, en su parte 5', la secuencia de un dominio de rótulo que no es un extremo del transposón, y, en su porción 3', la secuencia del extremo del transposón transferido, en donde el ADN objetivo se fragmenta y la cadena transferida de la composición de extremo del transposón se une con un extremo 5' de un fragmento del ADN objetivo para producir un fragmento de ADN objetivo de rótulo en 5'.

En algunas formas de realización, la presente invención proporciona métodos de producción de una biblioteca de fragmentos de ADN rotulados en 5', que comprende la incubación del ADN objetivo con una transposasa y una composición de extremo del transposón que comprende cadenas transferidas que exhiben, en sus porciones 5', la secuencia de uno o varios dominios de rotulación para un fin particular y, en sus porciones 3', la secuencia del extremo del transposón transferido, en condiciones en las que una reacción de transposición es catalizada por la transposasa, en donde el ADN objetivo se fragmenta y las cadenas transferidas de la composición de extremo del transposón se unen con los extremos 5' de un fragmento del ADN objetivo para producir fragmentos de ADN objetivo con rótulo en 5', de modo que la reacción de transposición produce una pluralidad de fragmentos de ADN objetivo con rótulo en 5' que comprende una biblioteca de fragmentos de ADN del ADN objetivo.

La presente invención también describe composiciones adecuadas para fragmentación mediante transposición in vitro. Por ejemplo, en algunas formas de realización, la presente invención proporciona una composición que comprende una molécula de ácido nucleico sintético que tiene una porción 5' que comprende un dominio de rótulo y una porción 3' que comprende una cadena transferida de un extremo del transposón. En algunas formas de realización, invención proporciona una composición que comprende una pluralidad de moléculas de ácidos nucleicos sintéticos, en donde las moléculas de ácidos nucleicos comprenden porciones 5' que comprenden dominios de rotulación que difieren en al menos un nucleótido y porciones 3' que comprenden una cadena transferida de un extremo del transposón.

En algunas formas de realización de las composiciones descritas con anterioridad, al menos la porción 3' de la molécula de ácido nucleico es ADN bicatenario.

En ciertas formas de realización preferidas, el extremo del transposón es un extremo de transposón Tn5, mientras que en otras formas de realización, el extremo del transposón es un extremo de transposón Mu.

En algunas forms de realización, el dominio de rótulo de la composición de molécula de ácido nucleico comprende uno o más de un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de amplificación, un dominio de rótulo de detección, un dominio de rótulo de direccionamiento y un dominio de promotor de la transcripción. En formas de realización de particular preferencia, los dominios de rótulo comprenden rótulos de secuenciación que comprenden o consisten en rótulos de secuenciación seleccionados de rótulos de secuenciación de Roche 454A y 454B, rótulos de secuenciación ILLUMINA™ SOLEXA™, rótulos de secuenciación SOLID™ de Applied Biosystems, rótulos de secuenciación SMRT™ de Pacific Biosciences, rótulos de secuenciación de Pollonator Polony o rótulos de secuenciación de Complete Genomics.

En algunas formas de realización, la composición que comprende una molécula de ácido nucleico también comprende

una transposasa purificada. En formas de realización preferidas, la transposasa es una transposasa Tn5 y una Mu transposasa. En formas de realización preferidas, la molécula de ácido nucleico y la transposasa se proporcionan en una mezcla. En formas de realización particularmente preferidas, la mezcla también comprende un detergente no iónico. En formas de realización preferidas de mayor preferencia, el detergente no iónico comprende Nonidet P-40 y/o Tween-20.

En algunas formas de realización, la presente divulgación proporciona un kit que comprende cualquiera de las composiciones descritas con anterioridad o en otro lado de la presente. En algunas formas de realización, el kit también comprende uno o varios de una ligasa, una polimerasa y/o reactivos para una reacción de amplificación. En formas de realización particularmente preferidas, los reactivos para una reacción de amplificación comprenden reactivos para una reacción de polimerasa en cadena. En formas de realización preferidas, los reactivos o una reacción de amplificación y/o reacción de polimerasa en cadena comprenden al menos un cebador que comprende una porción 3' que es complementaria al complemento del dominio de rótulo de la porción 5' de la molécula de ácido nucleico. En formas de realización preferidas, el dominio de rótulo comprende uno o varios de un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rotulación de amplificación, un dominio de rótulo de detección, un dominio de rótulo de direccionamiento y un dominio de promotor de la transcripción y en ejemplos de particular preferencia, el dominio de rótulo comprende un dominio de rótulo de secuenciación que comprende o consiste en un rótulo de secuenciación seleccionado de rótulos de secuenciación de Roche 454A y 454B, rótulos de secuenciación ILLUMINA™ SOLEXA™, rótulos de secuenciación SOLID™ de Applied Biosystems, rótulos de secuenciación SMRT™ de Pacific Biosciences, rótulos de secuenciación de Pollonator Polony o rótulos de secuenciación de Complete Genomics.

En algunas formas de realización, un kit de la presente divulgación también comprende reactivos para una reacción de secuenciación de ADN.

En algunas formas de realización, la presente invención comprende una mezcla de reacción que comprende un ADN objetivo bicatenario y cualquiera de las moléculas de ácidos nucleicos del extremo del transposón rotulado y/o las composiciones de transposasa descritas con anterioridad.

En algunas formas de realización, la presente invención proporciona una composición que comprende una transposasa purificada y una pluralidad de extremos del transposón sintético o composiciones de extremo del transposón. En algunas formas de realización, las composiciones de extremo del transposón comprenden extremos del transposón en horquilla, mientras que, en algunas formas de realización, los extremos del transposón sintético o las composiciones de extremo del transposón comprenden cadenas transferidas separadas y cadenas no transferidas. En algunas formas de realización preferidas, las cadenas transferidas comprenden dominios de rótulo 5', por ejemplo, que comprende uno o varios de un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de amplificación, un dominio de rótulo de detección, un dominio de rótulo de direccionamiento y un dominio de promotor de la transcripción. En formas de realización particularmente preferidas, la presente invención los dominios de rótulo comprenden rótulos de secuenciación que comprenden o consisten en un rótulo de secuenciación seleccionado de rótulos de secuenciación de Roche 454A y 454B, rótulos de secuenciación ILLUMINA™ SOLEXA™, rótulos de secuenciación SOLID™ de Applied Biosystems, rótulos de secuenciación SMRT™ de Pacific Biosciences, rótulos de secuenciación de Pollonator Polony o rótulos de secuenciación de Complete Genomics.

En algunas formas de realización, la composición que comprende una transposasa purificada comprende una pluralidad de extremos del transposón sintético compuestos por al menos dos cadenas transferidas que difieren entre sí por al menos un nucleótido y en formas de realización preferidas, las cadenas transferidas comprenden porciones 5' y porciones 3', en donde al menos dos de las porciones 5' de las cadenas transferidas comprenden rótulos que difieren entre sí por al menos un nucleótido y en donde las porciones 3' de las cadenas transferidas comprenden la misma secuencia del extremo del transposón.

En algunas formas de realización los extremos del transposón comprenden extremos del transposón Mu y la transposasa es transposasa Mu y en algunas formas de realización de preferencia, las porciones 3' de las cadenas transferidas comprenden una secuencia de un extremo del transposón Mu y en donde las porciones 5' de las cadenas transferidas no son de un transposón Mu.

En algunas formas de realización, los extremos del transposón comprenden extremos del transposón Tn5 y la transposasa es transposasa Tn5 y en algunas formas de realización de preferencia, las porciones 3' de las cadenas transferidas comprenden una secuencia de un extremo del transposón Tn5 y en donde las porciones 5' de las cadenas transferidas no son de un transposón Tn5.

En algunas formas de realización, la presente invención proporciona composiciones que comprenden una biblioteca de fragmentos de ADN, en donde la biblioteca de fragmentos de ADN comprende fragmentos del ADN objetivo que 5 tiene extremos 5' que comprenden secuencias de cadenas transferidas de extremos del transposón o composiciones de extremo del transposón. Las secuencias de las cadenas transferidas comprenden dominios de rótulo 5' y en formas de realización preferidas, la biblioteca de fragmentos de ADN comprende fragmentos de ADN objetivo que comprende

rótulos 3' complementarios a una cadena transferida de un extremo del transposón o composición de extremo del transposón. En algunas formas de realización, la biblioteca de fragmentos de ADN comprende fragmentos bicatenarios del ADN objetivo.

5 En algunas formas de realización, la presente divulgación proporciona composiciones que comprenden un fragmento de ADN circular rotulado de un ADN objetivo, que comprende una porción que comprende la secuencia de cadena no transferida en el extremo 5' de la porción, una secuencia de cadena transferida en el extremo 3' o la porción, una secuencia de bucle interviniente que comprende un dominio de rótulo, secuencias de ambas cadenas de una porción de un ADN objetivo.

10 En algunas formas de realización, la presente invención proporciona composiciones que comprenden un fragmento de ADN monocatenario circular rotulado de un ADN objetivo, que comprende una secuencia de cadena transferida de un extremo del transposón o la composición de extremo del transposón y una porción monocatenaria de un ADN objetivo. La secuencia de cadena transferida comprende un dominio de rótulo.

15 En algunas formas de realización, la presente invención proporciona composiciones que comprenden un fragmento de ADN bicatenario de cola de milano de un ADN objetivo, que comprende una porción bicatenaria de un ADN objetivo en donde cada cadena tiene un extremo 5' que comprende al menos una porción de una secuencia de cadena transferida y un extremo 3' que comprende al menos una porción de una secuencia de cadena no transferida.

Las formas de realización de la invención se describen en este compendio y en la siguiente descripción detallada de la invención.

20 Aunque la invención se ha descrito en relación con formas de realización específicas, debe entenderse que la invención reivindicada no debe estar indebidamente limitada a tales formas de realización específicas.

Breve descripción de los dibujos

25 Las siguientes figuras forman parte de la presente memoria descriptiva y están incluidas para demostrar también ciertos aspectos de la presente divulgación. La invención se puede comprender mejor por referencia a una o varias de estas figuras en combinación con la descripción detallada de formas de realización específicas presentadas en la presente.

La FIG. 1 proporciona un diagrama esquemático que muestra la inserción de un transposón en un ADN objetivo en una reacción de transposición.

La FIG. 2 proporciona un diagrama esquemático que muestra la fragmentación y la rotulación de ADN objetivo por inserción de extremos del transposón en una reacción de transposición.

30 La FIG. 3 proporciona un diagrama esquemático de los productos de dos eventos de inserción de la composición de extremo del transposón catalizada con transposasa. Así, el producto de la inserción de la composición de extremo del transposón catalizada con transposasa representa a la izquierda de la figura muestra una orientación del extremo del transposón en donde la cadena transferida de la composición de extremo del transposón (es decir, en donde el extremo del transposón transferido exhibe la secuencia 5' AGATGTGTATAAGAGACAG 3' (SEQ ID NO: 1), con su extremo 3' unido con el ADN objetivo) está en la cadena superior y el producto de la inserción de la composición de extremo del transposón catalizada con transposasa representada a la derecha de la figura muestra una orientación de extremo del transposón en donde la cadena transferida está en la cadena inferior. La cadena no transferida es la SEQ ID NO:2.

35 La FIG. 4 ilustra ejemplos de dos diferentes extremos de transposón rotulados, donde cada uno comprende una cadena transferida de oligonucleótido con un rótulo diferente en la porción 5' para usar en la generación de una biblioteca de fragmentos de ADN rotulados. La extensión de los extremos 3' de cada cadena usando, por ejemplo, ADN polimerasa con 5' nucleasa o actividad de desplazamiento de cadena, produce fragmentos de ssADN dirrotulados. La secuencia de cadena transferida mostrada es la SEQ ID NO:1; la cadena no transferida es la SEQ ID NO:2.

45 La FIG. 5 muestra una imagen en gel de agarosa que muestra el rango de tamaño de fragmento de ADN con rótulo en productos de transposición 5' producidos usando diferentes concentraciones de transposoma.

La FIG. 6 muestra una imagen en gel de agarosa que muestra el rango de tamaño de productos de transposición de fragmentos de ADN con rótulo en 5' producidos en reacciones de cinco minutos a diferentes temperaturas, con distintos tampones de reacción, en presencia o ausencia de dimetilformamida (DMF).

50 La FIG. 7 ilustra un ejemplo del método en donde una ADN polimerasa que tiene una actividad de ADN polimerasa de desplazamiento de cadena y/o que tiene una actividad de 5'- a 3'-exonucleasa se usa para ligar el complemento de la cadena transferida con los fragmentos de ADN con rótulo en 5' de la reacción de transposición in vitro para generar una biblioteca de fragmentos de ADN que comprende fragmentos de ssADN dirrotulados. Tal como se muestra, la actividad de desplazamiento de cadena y/o 5'- a 3'-exonucleasa de la ADN polimerasa desplaza o digiere el ADN que se fusiona corriente debajo del producto de extensión de la ADN polimerasa y la extensión por la ADN polimerasa une

un segundo rótulo que comprende o consisten en una secuencia de ADN que es complementaria al primer rótulo insertado en la cadena opuesta. En algunas formas de realización, los productos de fragmentos de ADN dirrotulados son amplificados por PCR usando oligonucleótidos que son complementarios al complemento de la cadena transferida. La secuencia de cadena transferida mostrada es la SEQ ID NO: 1; la cadena no transferida es la SEQ ID NO:2.

5 La FIG. 8 ilustra un ejemplo del método en donde una ADN polimerasa tiene actividad de ADN polimerasa de desplazamiento de cadena y/o que tiene actividad de 5'- a 3'-exonucleasa se usa para unir el segundo rótulo con los fragmentos de ADN con rótulo en 5' de la reacción de transposición in vitro para generar una biblioteca de fragmentos de ADN que comprende fragmentos de ADN dirrotulados. Tal como se muestra, la actividad de desplazamiento de  
10 cadena y/o de 5'- a 3'-exonucleasa de la ADN polimerasa desplaza o digiere el ADN que se fusiona corriente abajo del producto de extensión de ADN polimerasa y la extensión por la ADN polimerasa une un segundo rótulo que comprende o consiste en una secuencia de ADN que complementaria al primer rótulo insertado en la cadena opuesta. En algunas formas de realización, los productos de fragmentos de ssADN dirrotulados se amplifican por PCR usando oligonucleótidos que son complementarios a las diferentes secuencias en los respectivos primer o segundo rótulo como cebadores de PCR. La secuencia de cadena transferida mostrada es la SEQ ID NO:1; la cadena no transferida  
15 es la SEQ ID NO:2.

La FIG 9 proporciona una comparación de longitud de lectura de secuenciación, precisión y cobertura de un cóntigo simple usando una biblioteca de fragmentos de ADN producidos de acuerdo con formas de realización de la presente invención, en comparación con una biblioteca de control producida usando nebulización.

20 La FIG. 10 proporciona un diagrama esquemático que muestra fragmentación y rotulación de ADN objetivo para producir una biblioteca compatible con Roche/454 por inserción de extremos rotulados de transposón en una reacción de transposición, seguido de códigos de barras +/- de PCR específicos del rótulo.

La FIG 11 muestra una imagen en gel de agarosa que muestra ADN de entrada (carril 2), el rango de tamaño del fragmento de ADN con rótulo en productos de transposición 5' (carril 3), el rango de tamaño de productos de reacción de PCR (carril 4) y una reacción de control (carril 5) para la preparación de una biblioteca de secuenciación compatible  
25 con Roche/454 FLX tal como se ilustra en la FIG. 10.

La FIG 12 muestra una imagen en gel de agarosa que muestra el rango de tamaño de fragmento de ADN con rótulo en productos de transposición 5' (carril 2), el rango de tamaño de productos de reacción de PCR (carril 3) y una reacción de control (carril 5) para la preparación de una biblioteca de secuenciación compatible con Roche/454 FLX Titanium de amplicón ADN, similar al método ilustrado en la FIG. 10.

30 La FIG. 13 proporciona un diagrama esquemático que muestra fragmentación y rotulación de ADN objetivo para producir una biblioteca compatible con Illumina/Solexa por inserción de extremos rotulados de transposón en una reacción de transposición, seguido de códigos de barras +/- de PCR específicos de rotulación.

La FIG. 14 muestra una imagen en gel de agarosa que muestra ADN de entrada (carril 2), el rango de tamaño del fragmento de ADN con rótulo en productos de transposición 5' (carril 3), el rango de tamaño de productos de reacción de PCR (carril 4) y una reacción de control (carril 5) para la preparación de una biblioteca de secuenciación compatible  
35 con Illumina GAII con código de barras tal como se ilustra en la Figura 13.

La FIG. 15 compara el proceso y la complejidad de los métodos de la técnica anterior de la preparación de bibliotecas con preparación de bibliotecas de fragmentos de ADN de acuerdo con formas de realización de la presente invención.

40 La FIG. 16. muestra un ejemplo de un producto del método de la invención después de la incubación de una transposasa (por ejemplo, aquí EZ-Tn5<sup>TM</sup> Transposase) y una composición de extremo del transposón en horquilla (por ejemplo, la composición de extremo del transposón en horquilla EZ-Tn5<sup>TM</sup> "pMETS-N-MENTS" representada aquí) en una reacción in vitro de transposasa en presencia del ADN objetivo bicatenario (por ejemplo, ADN genómico o cADN bicatenario).

45 FIG. 17. Fragmentación y rotulación de ADN objetivo con una composición de extremo del transposón en horquilla. La FIG 17A muestra un diagrama esquemático de un producto (en una población de muchos de estos productos) resultantes de dos eventos de inserción catalizados con transposasa de la composición de extremo del transposón en horquilla en ADN objetivo. Brevemente, el ADN objetivo (por ejemplo, que comprende ADN genómico bicatenario o cADN) se incuba en una reacción in vitro de transposasa con una transposasa y una composición de extremo del transposón en horquilla (por ejemplo, EZ-Tn5<sup>TM</sup> Transposase y una composición de extremo del transposón en horquilla EZ-Tn5<sup>TM</sup>). La secuencia terminal transferida de cada la composición de extremo del transposón en horquilla insertada se une por medio de una estructura en bucle con la secuencia de cadena no transferida del extremo del transposón. El bucle puede tener cualquier dominio de rótulo arbitrario, tales como un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de detección, un dominio de rótulo de direccionamiento, un dominio de promotor de la transcripción o un dominio de rótulo de amplificación. Por  
50 ejemplo, el dominio de rótulo de secuenciación puede exhibir la secuencia de un rótulo de secuenciación Roche 454A o 454B. Por ejemplo, en esta figura, el dominio de rótulo de secuenciación exhibe una o varias secuencias en el bucle entre las secuencias terminales de transposón complementarias (el tallo). La FIG 17B muestra una electroforesis en gel de agarosa al 1% teñido con SYBR Gold de los productos de rótulo 5' y fragmentación de 1 µg de dsADN genómico  
55

T7 D111 usando 0, 0,5, 1, 2 o 3  $\mu\text{M}$  de la composición de extremo del transposón en horquilla pMETS-N-MENTS y cantidades equimolares de EZ-Tn5<sup>TM</sup> transposasa (EPICENTRE Biotechnologies, Madison, WI), después de incubación en 33 mM de acetato de Tris pH 7,6, 66 mM de KOAc y 10 mM de Mg(OAc)<sub>2</sub> durante 2 horas a 37 °C.

5 FIG. 18. Generación de plantillas de ADN circulares con rótulo 5'. La FIG 18A es un diagrama esquemático de una forma de realización del método. Brevemente, se rellenan brechas de 9-nucleótido generadas por dos inserciones de la composición de extremo del transposón en horquilla en un ADN objetivo por extensión de los extremos 3' de los fragmentos de ADN con rótulo en 5' generados usando una ADN polimerasa que carece de 5'- a 3' exonucleasa y actividades de desplazamiento de cadena (por ejemplo, T4 ADN polimerasa) y el ADN monocatenario en las regiones de brecha como plantillas y luego los productos de extensión de ADN polimerasa de los fragmentos de ADN con rótulo en 5' se ligan usando una ligasa dependiente de la plantilla (por ejemplo, ADN ligasa de E. coli) para generar fragmentos de ADN circulares rotulados. Los fragmentos de ADN circulares rotulados son resistentes a T5 exonucleasa, que se usa en la forma de realización en la figura para eliminar ADN monocatenario lineal no ligado y ADN bicatenario. La FIG 18B muestra un gel de electroforesis de agarosa al 1% teñido con SYBR Gold de los productos de reacción generados en presencia o ausencia de T4 ADN polimerasa y/o ADN ligasa de E. coli. Tal como se mostró, los fragmentos de ADN circulares rotulados, que son resistentes al tratamiento con T5 exonucleasa, se generaron sólo en presencia tanto de T4 ADN polimerasa como de ADN ligasa de E. coli.

La FIG. 19 ilustra el uso de una transferasa terminal para unir un segundo rótulo (3') con fragmentos de ADN rotulados en 5', para generar una biblioteca de fragmentos de ADN que comprende fragmentos de ssADN con rótulos en 5' y 3' ("dirrotulados"). En la forma de realización ilustrada, un dominio de rótulo de secuenciación que comprende un rótulo de secuenciación (SEQ) se añade usando una ADN polimerasa dependiente de la plantilla.

La FIG. 20 ilustra el uso de un oligonucleótido de rótulo terminal como una plantilla para añadir un segundo rótulo (3') con fragmentos de ssADN rotulados en 5' para generar una biblioteca de fragmentos de ADN que comprende fragmentos de ssADN dirrotulados.

La FIG. 21 ilustra una forma de realización en donde fragmentos de ADN con rótulo en 5' se incuban en presencia de una ADN ligasa y un ligamiento de oligonucleótido de rotulación que comprende una porción 5' que tiene un grupo fosfato en su extremo 5' y que exhibe una secuencia aleatoria que se fusiona con la brecha de 9-base o la región de ADN monocatenario que resulta de la inserción catalizada con EZ-Tn5 transposasa del extremo de transposón EZ-Tn5 ME en el ADN objetivo y una porción 3' que exhibe una segunda secuencia de rótulo (rótulo No. 2). En este ejemplo, el oligonucleótido de rótulo de ligamiento tiene una porción 5' que exhibe una secuencia aleatoria de 6-nucleótidos. La secuencia aleatoria se fusiona con el ADN objetivo fusionado monocatenario en las regiones de brecha de 9-bases que resulta de la inserción de los extremos del transposón (por ejemplo aquí, el extremo mosaico EZ-Tn5 de 19 pares de bases o el extremo del transposón ME) en el ADN objetivo bicatenario. Aquellos oligonucleótidos de rotulación de ligamiento que se fusionan con el ADN monocatenario objetivo en las regiones de brecha de modo que un extremo 5'-fosforilado se apoya en el extremo 3' de un fragmento de ADN con rótulo en 5' se unen luego por medio de la ligasa de ácido nucleico en una reacción de ligamiento dependiente de la plantilla, generando así fragmentos de ADN rotulados en 5' y 3' con el primer rótulo en el extremo 5' y el segundo rótulo en el extremo 3'. Así, si las inserciones del extremo del transposón se producen en ambas cadenas del ADN objetivo en proximidad cercana (por ejemplo, en sitios en el ADN objetivo que están dentro de aproximadamente 50 Kb, aproximadamente 40 Kb, aproximadamente 30 Kb, aproximadamente 20 Kb, aproximadamente 10 Kb, aproximadamente 5 Kb, aproximadamente 1 Kb, aproximadamente 500 bp o, con preferencia, dentro de aproximadamente 150 hasta aproximadamente 500 entre sí), las cadenas con rótulo doble se pueden purificar, amplificar por PCR, opcionalmente, rotular con una tintura detectable (por ejemplo, para usar como objetivo para fusionar con una micromatriz, por ejemplo, para análisis de expresión), o, primero capturar en una superficie (por ejemplo, en una perla; por ejemplo una perla para secuenciación de próxima generación) y luego amplificar (por ejemplo, usando PCR en emulsión, por ejemplo, para usar como plantillas de secuenciación de próxima generación; o usando PCR de ciclo limitado, por ejemplo, para uso en experimentos de variación de cantidad de copias (CNV), por ejemplo, por hibridación genómica comparativa en una micromatriz).

La FIG. 22 (A) muestra un diagrama esquemático de una forma de realización de fragmentación de ADN, rotulación y circularización de ADN genómico usando el método de la invención. La casilla marcada con línea negra representa la composición de extremo del transposón p454.1 MEDS. Las líneas de puntos representan fragmentos de dsADN genómicos T7 D111.

La FIG. 22 (B) muestra un gel de agarosa gel de los productos de rótulo 5' y fragmentación de dsADN genómico T7 D111 usando la composición de extremo del transposón p454.1MEDS y EZ-Tn5<sup>TM</sup> transposasa (EPICENTRE Biotechnologies). 1  $\mu\text{g}$  de dsADN genómico T7 D111 se incubó con o sin 0,5  $\mu\text{M}$  de la composición de extremo del transposón p454.1MEDS en 33 mM de acetato de Tris pH 7,6, 66 mM de KOAc y 10 mM de Mg(OAc)<sub>2</sub> durante 1 hora a 37 °C, ya sea en presencia o ausencia de EZ-Tn5 transposasa tal como se indicó. Los productos de reacción de ssADN lineal de rótulo 5' se resolvieron por electroforesis en un gel de agarosa al 1% y se tiñeron con SYBR Gold.

La FIG. 23 (A) muestra un esquema de amplificación por PCR de los fragmentos de ssADN circulares rotulados usando los oligonucleótidos pMETS y p454.1 como cebadores de PCR.

La FIG. 23 (B) muestra un gel de agarosa de los productos de amplificación por PCR obtenidos cuando los fragmentos

de ssADN circulares rotulados obtenidos usando el método de la invención se amplificaron por PCR usando los oligonucleótidos pMETS y pc454.1 como cebadores de PCR. Primero, se desnaturalizó el fragmentado de rótulo 5', obtenido tal como se mostró en la FIG. 22 (B), calentando a 95 °C durante 3 minutos y enfriando rápidamente en hielo. Se incubó una porción de los fragmentos de ssADN lineales desnaturalizados de rótulo 5' obtenidos en 33 mM de acetato de Tris pH 7,6, 66 mM de KOAc, 2,5 mM de MnCl<sub>2</sub> y 1 M de betaína durante 2 horas a 60 °C en presencia o ausencia de 400 unidades de CIRCLIGASE™ ssADN ligasa tal como se describe en el Ejemplo 17, a fin de dar circularización a los fragmentos de ssADN lineales de rótulo 5'. Los productos de reacción se incubaron con exonucleasa I y exonucleasa III durante 1 hora a 37 °C para digerir los fragmentos de ssADN lineales antes de la amplificación por PCR. Los fragmentos de ssADN circulares rotulados se amplificaron luego por PCR usando los oligonucleótidos pMETS y pc454.1 como cebadores de PCR. Los productos de PCR se resolvieron por electroforesis en gel de agarosa al 1% y se visualizaron por tinción SYBR Gold.

#### Definiciones

A menos que se defina o describa específicamente de manera diferente en otro lado de la presente, los siguientes términos y descripciones relacionados con la invención se han de entender como sigue más abajo.

15 Cuando los términos “por ejemplo”, “por ejemplo”, “tales como”, “incluyen”, “incluyendo” o variaciones de ellos se usan en la presente, estos términos no se considerarán términos de limitación y se interpretarán como “pero sin limitación” o “sin limitación”.

20 El uso de términos “un”, “una”, “el” y “la” y referentes similares en el contexto de la descripción de la invención (en especial en el contexto de las siguientes reivindicaciones) se han de construir para cubrir tanto el singular como el plural, a menos que se indique otra cosa en la presente o claramente se contradiga por contexto.

25 Tal como se usan en la presente, los términos “aislado”, “aislar”, “aislamiento”, “purificado”, “purificar”, “purificación” y sus equivalentes gramaticales tal como se usan en la presente, a menos que se especifique otra cosa, se refieren a la reducción en la cantidad de al menos un contaminante (como secuencia de proteínas y/o ácidos nucleicos) a partir de una muestra o una fuente (por ejemplo, una célula) de la que se aísla el material. Así, los resultados de purificación dan como resultado un “enriquecimiento”, es decir, un aumento en la cantidad de una secuencia de proteínas y/o ácidos nucleicos deseable en la muestra.

30 Tal como se usan en la presente, un “rótulo” se refiere a un componente de ácido nucleico no objetivo, en general, ADN, que proporciona un medio de direccionamiento a un fragmento de ácido nucleico al que se une. Por ejemplo, en formas de realización preferidas, un rótulo comprende una secuencia de nucleótidos que permite la identificación, el reconocimiento y/o la manipulación molecular o bioquímica del ADN al que el rótulo se une (por ejemplo, proporcionando un sitio para fusión de un oligonucleótido, como un cebador para extensión por una ADN polimerasa, o un oligonucleótido para captura o para una reacción de ligamiento). El proceso de unión del rótulo con la molécula de ADN a veces se menciona en la presente como “rotulación” y el ADN que se somete a la rotulación o que contiene un rótulo se menciona como “rotulado” (por ejemplo, “ADN rotulado”).

35 Tal como se usa en la presente, el término “ligasa” se refiere a una enzima modificadora del ácido nucleico que cataliza la formación intra- e intermolecular de enlaces fosfodiéster entre los términos de 5'-fosfato y 3'-hidroxilo cadenas de ácido nucleico. Las ligasas incluyen, por ejemplo, ligasas independientes de la plantilla tales como CIRCLIGASE™ ssADN ligasa, que se puede unir con extremos de ARN y ADN monocatenarios y ligadas dependientes de la plantilla u homólogas, que sella muescas en ADN bicatenario (ejemplo descrito más abajo).

40 Tal como se usa en la presente, una “ligasa homóloga” o “ligasa dependiente de la plantilla” implica una ADN ligasa que cataliza la formación intra- e intermolecular de enlaces fosfodiéster entre los términos de 5'-fosfato y 3'-hidroxilo de cadenas de ADN que son adyacentes entre sí cuando se fusionan con un polinucleótido complementario. Algunas formas de realización de ligamiento intramolecular producen una molécula circular y se mencionan como “circularización”. El polinucleótido al que ambos extremos del ADN que se han de ligar se fusionan de modo adyacente se menciona en la presente como una “plantilla de ligamiento” y el ligamiento se menciona como “ligamiento homólogo” o “ligamiento dependiente de la plantilla”. La plantilla de ligamiento puede ser una secuencia de ADN complementaria en ADN genómico u otro ADN en una muestra biológica (en cuyo caso se menciona a menudo como una “secuencia objetivo”) o la plantilla de ligamiento puede ser “oligodexirribonucleótido en puente” u “oligodesoxirribonucleótido de tablilla de ligamiento” (o “tablilla de ligamiento”) que se sintetiza y/o proporciona específicamente para usar en un ensayo o método particular. Los ejemplos de ADN ligasas homólogas o dependientes de la plantilla incluyen ADN ligasas de tipo NAD tales como ADN ligasa de E. coli, Tt ADN ligasa, Tfl ADN ligasa y AMPLIGASE® ADN ligasa (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos), que catalizan el ligamiento intramolecular de moléculas de ssADN sólo en presencia de una plantilla de ligamiento y ADN ligasas de tipo ATP tales como T4 ADN ligasa o FASTLINK™ ADN ligasa (EPICENTRE Biotechnologies), que, si bien no requieren de una plantilla de ligamiento para ligamiento de extremos romos, catalizan el ligamiento dependiente de la plantilla de modo mucho más eficiente.

En algunas formas de realización de preferencia, la ligasa dependiente de la plantilla es de una bacteria psicrófila o un bacteriófago psicrófilo, de modo que el ligamiento se puede llevar a cabo a menores temperaturas (por ejemplo, cuando las secuencias de los oligonucleótidos o polinucleótidos que forman la unión de ligamiento exhiben menores

T<sub>m</sub>). Una ADN ligasa se selecciona para usar en el método que es activa a una temperatura a la cual las moléculas de ADN usadas para unión (por ejemplo, productos de extensión de fragmentos de ADN con rótulo en 5' o los fragmentos de ADN con rótulo en 5' y los oligonucleótidos de secuencia aleatoria) se fusionan durante un tiempo suficiente para ser ligadas por la ligasa.

- 5 Una etapa importante en las formas de realización del método de la presente invención es el uso de una reacción de transposición in vitro para fragmentar y rotular el ADN objetivo para generar fragmentos de ADN rotulados. La reacción de transposición in vitro requiere de una transposasa, una composición de extremo del transposón y condiciones de reacción apropiadas.

10 Una "transposasa" implica una enzima que es capaz de formar un complejo funcional con una composición con contenido de extremo del transposón (por ejemplo, transposones, extremos del transposón, composiciones de extremo del transposón) y catalizar la inserción o transposición de la composición que contiene el extremo del transposón en el ADN bicatenario objetivo con el que se incuba en una reacción de transposición in vitro.

15 La expresión "extremo del transposón" significa un ADN bicatenario que exhibe sólo las secuencias de nucleótidos (las "secuencias del extremo del transposón") que son necesarias para formar el complejo con la transposasa o la enzima integrasa que es funcional en una reacción de transposición in vitro. Un extremo de transposón forma un "complejo" o un "complejo sináptico" o un "complejo de transposoma" o una "composición de transposoma con una transposasa o integrasa que reconoce y se une con el extremo de transposón y cuyo complejo es capaz de insertar o transponer el extremo del transposón en ADN objetivo con el que se incuba en una reacción de transposición in vitro.

20 Un extremo de transposón exhibe dos secuencias complementarias que consisten en una "secuencia de extremo de transposón transferida" o "una cadena transferida" y una "secuencia de extremo de transposón no transferida", o "cadena no transferida". Por ejemplo, un extremo de transposón que forma un complejo con una Tn5 transposasa hiperactiva (por ejemplo, EZ-Tn5™ Transposase, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos) que es activa en una reacción de transposición in vitro comprende una cadena transferida que exhibe una "secuencia de extremo de transposón transferida" como sigue:

25 5' AGATGTGTATAAGAGACAG 3' (SEQ ID NO:1)

y una cadena no transferida que exhibe una "secuencia de extremo del transposón no transferida" como sigue:

5' CTGTCT CTTATACACATCT 3' (SEQ ID NO:2).

30 El extremo 3' de una cadena transferida se une o transfiere al ADN objetivo en una reacción de transposición in vitro. La cadena no transferida, que exhibe una secuencia de extremo del transposón que es complementaria a la secuencia de extremo del transposón transferida, no se une o transfiere al ADN objetivo en una reacción de transposición in vitro.

35 En algunas formas de realización, la cadena transferida y la cadena no transferida se unen de modo covalente. Por ejemplo, en algunas formas de realización, las secuencias de cadena transferidas y no transferidas se proporcionan en un único oligonucleótido, por ejemplo, en una configuración de horquilla. Como tal, a pesar de que el extremo libre de la cadena no transferida no se une con el ADN objetivo directamente por la reacción de transposición, la cadena no transferida se vuelve unida con el fragmento de ADN de forma indirecta, porque la cadena no transferida se liga con la cadena transferida por el bucle de la estructura de horquilla.

40 Una "composición de extremo del transposón" implica una composición que comprende un extremo del transposón (es decir, el segmento mínimo de ADN bicatenario que es capaz de actuar con una transposasa para someterse a una reacción de transposición), opcionalmente más secuencia o secuencias adicionales. 5' de la secuencia del extremo del transposón transferida y/o 3' de la secuencia del extremo del transposón no transferida. Por ejemplo, un extremo del transposón unido con un rótulo es una "composición de extremo del transposón". En algunas formas de realización, la composición de extremo del transposón comprende o consiste en dos oligonucleótidos del extremo del transposón que consisten en el "oligonucleótido del extremo del transposón transferido" o "cadena transferida" y el "oligonucleótido del extremo del transposón no transferido", o "cadena no transferida" que, en combinación, exhiben las secuencias del extremo del transposón y en donde una o ambas cadenas comprenden una secuencia adicional.

45

50 Las expresiones "oligonucleótido del extremo del transposón transferido" y "cadena transferida" se usan indistintamente y se refieren a la porción transferida tanto de "extremos del transposón" como "composiciones de extremo del transposón", es decir, respecto de si el extremo del transposón está unido con un rótulo o con otro resto. De modo similar, las expresiones "oligonucleótido del extremo del transposón no transferido" y "cadena no transferida" se usan indistintamente y se refieren a la porción no transferida tanto de "extremos del transposón" como "composiciones de extremo del transposón". En algunas formas de realización, una composición de extremo del transposón es una "composición de extremo del transposón en horquilla". Tal como se usa en la presente, una 5 "composición de extremo del transposón en horquilla" significa una composición de extremo del transposón que consiste en un oligodesoxirribonucleótido simple que exhibe una secuencia del extremo del transposón no transferida en su extremo 5', una secuencia del extremo de transposón transferida en su extremo 3' y una secuencia arbitraria interviniente entre la secuencia del extremo de transposón no transferida y la secuencia del extremo del transposón transferida que es suficientemente larga para permitir una formación intramolecular de tallo-bucle, de modo tal que la porción de extremo de transposón pueda funcionar en una reacción de transposición. En algunas formas de

55

realización, el extremo 5' de la composición de extremo del transposón en horquilla tiene un grupo fosfato en la posición 5' del 5'-nucleótido. En algunas formas de realización, la secuencia arbitraria interviniente entre la secuencia del extremo del transposón no transferida y la secuencia del extremo del transposón transferida de una composición de extremo del transposón en horquilla proporciona un rótulo (por ejemplo, incluyendo uno o varios dominios de rótulo) para un uso o aplicación particular.

En algunas formas de realización, los métodos de la presente invención producen fragmentos de ssADN circulares rotulados. En algunas formas de realización, los fragmentos de ssADN circulares rotulados exhiben únicamente la secuencia de la cadena transferida de la composición del extremo del transposón y los fragmentos de ssADN circulares rotulados no exhiben la secuencia de la cadena no transferida de la composición del extremo del transposón.

En algunas formas de realización, la composición de extremo del transposón usada en el método de la presente invención comprende oligonucleótidos del extremo del transposón que exhiben únicamente las secuencias del extremo del transposón que forman un complejo con la transposasa o integrasa y que se necesitan para la reacción de transposición; en estas formas de realización, el rótulo en los fragmentos de ssADN circulares rotulados generados usando el método exhibe únicamente la secuencia del extremo del transposón transferida.

Sin embargo, en algunas formas de realización, la composición de extremo del transposón comprende o consiste en al menos un oligonucleótido del extremo del transposón que exhibe una o varias otras secuencias de nucleótidos además de las secuencias del extremo del transposón. Así, en algunas formas de realización, la composición de extremo del transposón comprende una cadena transferida que exhibe una o varias otras secuencias de nucleótidos 5' de la secuencia del extremo del transposón transferida, cuyas una o varias otras secuencias de nucleótidos también son exhibidas por el rótulo. Así, además de la secuencia del extremo del transposón transferida, el rótulo puede tener una o varias otras porciones de rótulo o dominios de rótulo.

Tal como se usa en la presente, una "porción de rótulo" o un "dominio de rótulo" implican una porción o dominio de un rótulo que exhibe una secuencia para un fin o aplicación pretendido deseado. Una porción de rótulo o dominio de rótulo es el "dominio del extremo del transposón", cuya porción de rótulo o dominio de rótulo exhibe la secuencia del extremo del transposón transferida. En algunas formas de realización en donde la cadena transferida también exhibe una o varias otras secuencias de nucleótidos 5' de la secuencia del extremo del transposón transferida, el rótulo también tiene uno o varios otros "dominios de rótulo" en dicha porción 5', cada uno de los cuyos dominios de rótulo se proporciona para cualquier finalidad deseada. Por ejemplo, algunas formas de realización de la invención comprenden o consisten en una composición de extremo del transposón que comprende o consiste en: (i) una cadena transferida que exhibe una o varias secuencias 5' de la secuencia del extremo del transposón transferida que comprende o consiste en un dominio de rótulo seleccionado de entre uno o varios de un sitio de restricción, un dominio de rótulo, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de amplificación, un dominio de rótulo de detección, un dominio de rótulo de direccionamiento y un dominio de promotor de transcripción; y (ii) una cadena no transferida que exhibe la secuencia del extremo del transposón no transferida. La invención comprende formas de realización del método que usa cualquiera o varias de dichas composiciones de extremo del transposón.

Tal como se usa en la presente, un "dominio de clivaje" se refiere a una secuencia de ácidos nucleicos que es susceptible.

Tal como se usa en la presente, un "dominio del sitio de restricción" implica un dominio de rótulo que exhibe una secuencia con el fin de facilitar el clivaje usando una endonucleasa de restricción. Por ejemplo, en algunas formas de realización, el dominio de sitio de restricción se usa para generar fragmentos de ssADN lineales dirrotulados. En algunas formas de realización, el dominio de sitio de restricción se usa para generar un extremo 5' bicatenario compatible en el dominio de rótulo de modo que su extremo se pueda ligar con otra molécula de ADN usando una ADN ligasa dependiente de la plantilla. En algunas formas de realización de preferencia, el dominio de sitio de restricción en el rótulo exhibe la secuencia de un sitio de restricción que está presente sólo raramente, si lo está, en el ADN objetivo (por ejemplo, un sitio de restricción para una endonucleasa de restricción de corte raro como NotI o AscI). En algunas formas de realización de preferencia, el sitio de restricción en el dominio de sitio de restricción es para una endonucleasa de restricción de tipo II, como endonucleasa de restricción FokI.

En algunas formas de realización en donde la composición de cadena transferida del extremo del transposón comprende uno o varios dominios del sitio de restricción 5' de la secuencia del extremo del transposón transferida, donde el método también comprende: fusión de un oligodesoxirribonucleótido que es complementario al sitio de restricción monocatenario de los fragmentos de ssADN circulares rotulados y luego clivaje los fragmentos de ssADN circulares rotulados en el sitio de restricción usando la endonucleasa de restricción que reconoce el sitio de restricción. Así, en algunas formas de realización, el método comprende linealizar los fragmentos de ssADN circulares rotulados para generar fragmentos de ssADN lineales dirrotulados.

En algunas otras formas de realización en donde la cadena transferida de la composición del extremo del transposón comprende uno o varios dominios del sitio de restricción 5' de la secuencia del extremo del transposón transferida, la cadena transferida de la composición del extremo del transposón comprende una horquilla bicatenaria que comprende el sitio de restricción y el método también comprende las etapas de clivaje de los fragmentos de ssADN lineales

rotulados en el sitio de restricción usando la endonucleasa de restricción que reconoce el sitio de restricción; sin embargo, en algunas formas de realización, este método no se prefiere porque la horquilla bicatenaria proporciona un sitio de dsADN en el que la composición de extremo del transposón puede ser trasladada por la transposasa o la integrasa.

- 5 En algunas formas de realización de preferencia que comprenden (i) generación de un sitio de restricción bicatenario, ya sea por fusión de un oligodesoxirribonucleótido que es complementario al sitio de restricción monocatenario o uso de una cadena transferida que comprende una horquilla bicatenaria y (ii) luego clivaje del sitio de restricción usando la endonucleasa de restricción que reconoce el sitio de restricción bicatenario, donde el método también comprende la etapa de ligamiento de fragmentos de ssADN lineales rotulados clivados con la endonucleasa de restricción con otra molécula de ADN que tiene un extremo 3' compatible.

10 Tal como se usan en la presente, un "dominio de rótulo de captura" o un "rótulo de captura" implican un dominio de rótulo que exhibe una secuencia a fin de facilitar la captura del fragmento de ssADN al que se une el dominio de rótulo (por ejemplo, para proporcionar un sitio de fusión o un rótulo de afinidad para una captura de los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados en una perla u otra superficie, por ejemplo, en donde el sitio de fusión de la secuencia de dominio de rótulo permite la captura por fusión con una secuencia específica que está en una superficie como una sonda en una perla o en un microchip o micromatriz o una perla de secuenciación). En algunas formas de realización del método, después de capturar los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados por fusión con una sonda complementaria sobre una superficie, el dominio de rótulo de captura proporciona un sitio para cebar la síntesis del ADN usando dichos fragmentos de ssADN circulares rotulados o dichos fragmentos de ssADN lineales dirrotulados (o los complementos de dichos fragmentos de ssADN circulares rotulados o fragmentos de ssADN lineales dirrotulados) como plantillas. En algunas otras formas de realización, el dominio de rótulo de captura comprende una porción 5' de la cadena transferida que se une con un grupo o resto químico que comprende o consiste en una molécula de unión por afinidad (por ejemplo, en donde la porción 5' de la cadena transferida se une con una primera molécula de unión por afinidad como biotina, estreptavidina, un antígeno o un anticuerpo que se une con el antígeno, que permite la captura de los fragmentos de ssADN rotulados circulares o los fragmentos de ssADN lineales dirrotulados sobre una superficie a la que se une una segunda molécula de unión por afinidad que forma un par de unión específica con la primera molécula de unión por afinidad).

15 Tal como se usan en la presente, un "dominio de rótulo de secuenciación" o un "rótulo de secuenciación" implica un dominio de rótulo que exhibe una secuencia con fines de facilitar la secuenciación de los fragmentos de ssADN a los que el rótulo se une usando el método para sintetizar fragmentos de ssADN circulares rotulados (por ejemplo, para proporcionar un sitio cebador para secuenciación por síntesis o para proporcionar sitios de fusión para la secuenciación por ligamiento o para proporcionar sitios de fusión para la secuenciación por hibridación). Por ejemplo, en algunas formas de realización, el dominio de rótulo de secuenciación proporciona un sitio para cebar la síntesis del ADN de dicho fragmento de ssADN o el complemento de dicho fragmento de ssADN.

20 Tal como se usa en la presente, un "dominio de rótulo de amplificación" implica un dominio de rótulo que exhibe una secuencia con fines de facilitar la amplificación de un ácido nucleico al que se anexa dicho rótulo. Por ejemplo, en algunas formas de realización, el dominio de rótulo de amplificación proporciona un sitio cebador para una reacción de amplificación de ácido nucleico usando una ADN polimerasa (por ejemplo, una reacción de amplificación por PCR o una reacción de amplificación de desplazamiento de cadena o una reacción de amplificación de círculo de rodadura) o una plantilla de ligamiento para el ligamiento de sondas usando una ligasa dependiente de la plantilla en una reacción de amplificación de ácido nucleico (por ejemplo, una reacción en cadena de ligamiento).

25 Tal como se usan en la presente, un "dominio de rótulo de detección" o un "rótulo de detección" implica un dominio de rótulo que exhibe una secuencia o un resto químico o bioquímico detectable con fines de facilitar la detección de los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados (por ejemplo, en donde la secuencia o resto químico comprende o se une con una molécula detectable; como una molécula detectable seleccionada de entre: una tintura visible, fluorescente, quimioluminiscente u otra tintura detectable; una enzima que es detectable en presencia de un sustrato, por ejemplo, una fosfatasa alcalina con NBT más BCIP o una peroxidasa con un sustrato apropiado); una proteína detectable, por ejemplo, una proteína fluorescente verde; y una molécula de unión por afinidad que se liga a un resto detectable o que puede formar un par de unión por afinidad o un par de unión específica con otra molécula de unión por afinidad detectable; o cualquiera de las muchas moléculas detectables o sistemas conocidos en la técnica).

30 Tal como se usan en la presente, un "dominio de rótulo de direccionamiento" o un "rótulo de direccionamiento" implican un dominio de rótulo que exhibe una secuencia que permite la identificación de una muestra específica (por ejemplo, en donde la cadena transferida tiene un diferente dominio de rótulo de direccionamiento que exhibe una secuencia distinta para cada muestra).

35 Tal como se usan en la presente, un "dominio de promotor de transcripción" o un "dominio de promotor" implica un dominio de rótulo que exhibe una secuencia para una secuencia promotora con sentido o para una secuencia promotora antisentido de un promotor de ARN polimerasa. Tal como se usa en la presente, una "secuencia promotora con sentido" implica la secuencia de un promotor de ARN polimerasa que se une con la cadena de ADN que sirve

como la plantilla para la transcripción por una ARN polimerasa que se une con el promotor de la ARN polimerasa e inicia la transcripción en condiciones de reacción apropiadas para la transcripción. Tal como se usan en la presente, una "secuencia promotora antisentido" implica la secuencia de un promotor de ARN polimerasa que es complementario a la secuencia promotora con sentido. En algunas formas de realización, la secuencia promotora con sentido exhibida por el dominio de promotor de transcripción es para una ARN polimerasa que se une con un promotor de ARN polimerasa monocatenario e inicia la transcripción, en cuyas formas de realización la secuencia promotora con sentido es suficiente para funcionar como el promotor de ARN polimerasa (por ejemplo, para N4 ARN polimerasa de bacteriófago). En algunas formas de realización, la secuencia promotora con sentido es para una ARN polimerasa que se une con un promotor de ARN polimerasa bicatenario e inicia la transcripción, en cuyas formas de realización el método comprende preparar el promotor de ARN polimerasa bicatenario (por ejemplo, por fusión con la secuencia promotora con sentido de oligodesoxirribonucleótido que exhibe una secuencia promotora antisentido que es complementaria a la secuencia promotora con sentido o usando los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados como plantillas para la síntesis de dsADN que comprende o que consiste en la secuencia promotora con sentido) antes de la transcripción usando una ARN polimerasa a la que se une e inicia la transcripción a partir del promotor de ARN polimerasa bicatenario. En algunas formas de realización, la secuencia promotora con sentido es para una ARN polimerasa de tipo T7 (por ejemplo, seleccionada de entre T7 ARN polimerasa, T3 ARN polimerasa y SP6 ARN polimerasa). Un dominio de promotor de transcripción que exhibe una secuencia promotora con sentido permite la síntesis de ARN que es complementario al ADN monocatenario objetivo al que la cadena transferida de la composición del extremo del transposón se liga usando el método. Los fragmentos de ssADN circulares rotulados generados usando una composición de extremo del transposón que comprende una cadena transferida que tiene un dominio de promotor de la transcripción que exhibe una secuencia promotora antisentido no pueden ser transcritos por una ARN polimerasa. Sin embargo, en algunas formas de realización, el dsADN sintetizado por extensión de un cebador que se fusiona con los fragmentos de ssADN circulares rotulados se usa para la transcripción por una ARN polimerasa que se une e inicia la transcripción a partir de un promotor de ARN polimerasa bicatenario; en estas formas de realización, el ARN sintetizado exhibe la misma secuencia que los fragmentos de ssADN circulares rotulados.

Los nombres y las descripciones de diferentes dominios de rótulo son por conveniencia, de modo de comprender con mayor facilidad y tratar los fines y aplicaciones pretendidos de las diferentes porciones o dominios del rótulo en distintas formas de realización. Sin embargo, estos nombres y descripciones no pretenden limitar el uso o las aplicaciones del rótulo o de cualquiera de sus dominios de rótulo de modo alguno. Así, cualquier rótulo o dominio de rótulo particular se puede usar para cualquier finalidad además o en lugar de la finalidad o aplicación pretendida o primaria. Además, un dominio de rótulo puede comprender dos o más distintos dominios de rótulo (por ejemplo, un dominio de rótulo de secuenciación puede comprender tanto el dominio del extremo del transposón como otro dominio de rótulo 5' del dominio del extremo del transposón) o un dominio de rótulo puede proporcionar las funciones o fines o aplicaciones de dos o más diferentes dominios de rótulo (por ejemplo, el dominio del extremo de transposón puede proporcionar la finalidad del extremo del transposón transferido y también proporciona la función o la finalidad de un dominio de rótulo de secuenciación y/o un dominio de rótulo de captura para una aplicación particular). Más aún, el rótulo no necesita ser descrito en términos de uno o varios diferentes dominios a fin de ser usado para cualquier fin o aplicación o función particular.

Tal como se usan en la presente, los términos "amplificar" o "amplificado", "amplificación" tal como se usan en referencia a un ácido nucleico o reacciones de ácido nucleico, se refieren a métodos in vitro de producción de copias de un ácido nucleico particular, como un ácido nucleico objetivo o un ácido nucleico rotulado producido, por ejemplo, por medio de una forma de realización de la presente invención. Numerosos métodos de amplificación de ácidos nucleicos se conocen en la técnica y las reacciones de amplificación incluyen reacciones de polimerasa en cadena, reacciones en cadena de ligasa, reacciones de amplificación de desplazamiento de cadena, reacciones de amplificación de círculo de rodadura, métodos de amplificación mediada por transcripción como NASBA (por ejemplo, patente U.S. No. 5.409.818), métodos de amplificación mediada por bucle (por ejemplo, amplificación "LAMP" usando secuencias formadoras de bucles, por ejemplo, tal como se describe en la patente U. S. No. 6.410.278). El ácido nucleico que se amplifica puede ser ADN que comprende, consiste o se deriva de ADN o ARN o una mezcla de ADN y ARN, incluyendo ADN y/o ARN modificados. Los productos resultantes de la amplificación de una molécula o moléculas de ácido nucleico (es decir, "productos de amplificación"), ya sea que el ácido nucleico inicial es ADN, ARN o ambos, pueden ser ADN o ARN o una mezcla de nucleósidos o nucleótidos de ADN y ARN o pueden comprender nucleósidos o nucleótidos de ADN o ARN modificados. Una "copia" no significa necesariamente una perfecta complementariedad o identidad de secuencia con la secuencia objetivo. Por ejemplo, las copias pueden incluir análogos de nucleótido tales como desoxiinosina o desoxiuridina, alteraciones de secuencia intencionales (tales como alteraciones de la secuencia introducidas a través de un cebador que comprende una secuencia que es hibridable, pero no complementaria, a la secuencia objetivo, y/o errores de secuencia que se producen durante la amplificación.

"Sustancias de unión por afinidad" o "moléculas de unión por afinidad" o "moléculas por afinidad" en la presente significan moléculas que tienen afinidad y "se unen" entre sí en determinadas condiciones, mencionadas como "condiciones de unión", para formar un "par de unión específica". Por ejemplo, biotina y estreptavidina, biotina y aavidina o digoxigenina y un anticuerpo específico que se une con digoxigenina son ejemplos de "pares de unión específica", con los miembros de cada par de unión específica que comprende "moléculas de unión por afinidad" o "sustancias de unión por afinidad" o "moléculas por afinidad". Las moléculas de unión por afinidad (por ejemplo, biotina

y/o estreptavidina) se pueden unir covalentemente o conjugar o unir no covalentemente con otras moléculas (por ejemplo, con ARN o ADN) o con una superficie sólida usando métodos conocidos en la técnica (por ejemplo, usando reactivos y métodos tal como se describe en *Avidin-Biotin Chemistry: A Handbook*, de D. Savage et al., Pierce Chemical Company, 1992 y en *Handbook of Fluorescent Probes and Research Products*, Ninth Edition, de R. P. Hoagland, Molecular Probes, Inc. y en *BIOCONJUGATE Techniques*, de Greg T. Hermanson, publicado por Academic Press, Inc., San Diego, CA, 1996). Las moléculas por afinidad que se conjugan con ADN o ARN también se pueden sintetizar usando un sintetizador de oligonucleótido usando reactivos y métodos conocidos en la técnica.

El término "unión" de acuerdo con la presente invención implica la interacción entre una molécula de afinidad y una sustancia de unión por afinidad como resultado de enlaces no covalentes, tales como, pero no limitadas a, enlaces de hidrógeno, interacciones hidrofóbicas, enlaces de Van der Waals y enlaces iónicos. Sin quedar ligado por la teoría, se cree en la técnica que estos tipos de enlaces no covalentes resultan en la unión, en parte debido a las formas o estructuras complementarias de las moléculas implicadas en el par de unión específica. En base a la definición para "unión", y la amplia variedad de moléculas de unión por afinidad o partes de unión específica, queda claro que las condiciones de unión varían para los diferentes pares de unión específica. Los expertos en la técnica pueden hallar o determinar con facilidad las condiciones en las que, en una muestra, se produce la unión entre las moléculas de unión por afinidad. En particular, los expertos en la técnica pueden determinar con facilidad las condiciones en las que la unión entre las moléculas de unión por afinidad que se consideraría en la técnica como "unión específica" se puede producir. Tal como se entiende en la técnica, esta especificidad se debe usualmente a la mayor afinidad entre las moléculas de unión por afinidad que para otras sustancias y componentes (por ejemplo, paredes de vasos, soportes sólidos) en una muestra. En ciertos casos, la especificidad también puede implicar o se puede deber a una asociación significativamente más rápida de moléculas de unión por afinidad que con otras sustancias y componentes en una muestra.

Los términos "fusionar" o "hibridar" y "fusión" o "hibridación" se refieren a la formación de complejos entre secuencias de nucleótidos que son suficientemente complementarias para formar complejos por medio de la formación de pares de bases de Watson-Crick. Respecto de la presente invención, las secuencias de ácidos nucleicos que son "complementarias a" o "complementarias con" o que "hibridan" o "se fusionan" con cada una de las otras son capaces de formar "híbridos" o "complejos" que son suficientemente estables para servir al fin pretendido. No se requiere que toda base de ácido nucleico dentro de una secuencia exhibida por una molécula de ácido nucleico sea capaz de formar pares de base o forma pares o se complejice con cada base de ácido nucleico con una secuencia exhibida por una segunda molécula de ácido nucleico a fin de que las dos moléculas de ácidos nucleicos o las respectivas secuencias exhibidas allí sean "complementarias" o "fusionadas" o "hibridadas" con otra. Tal como se usan en la presente, los términos "complementario" o "complementariedad" se usan con referencia a una secuencia de nucleótidos relacionada con las reglas de la formación de pares de bases. Por ejemplo, la secuencia 5'-A-G-T-3', es complementaria a la secuencia 3'-T-C-A-5'. La complementariedad puede ser "parcial", en donde sólo algunas de las bases de ácidos nucleicos coinciden de acuerdo con las reglas de formación de pares de bases. O bien, puede haber una complementariedad "completa" o "total" entre los ácidos nucleicos. El grado de complementariedad entre cadenas de ácido nucleico tiene efectos significativos sobre la eficacia y la longitud de la hibridación entre cadenas de ácido nucleico. Esto es de particular importancia en las reacciones de amplificación, así como los métodos de detección que dependen de la hibridación de ácidos nucleicos. La expresión "homología" se refiere a un grado de complementariedad de una secuencia de ácidos nucleicos con otra secuencia de ácidos nucleicos. Puede haber una homología parcial o una homología completa (es decir, complementariedad). Una secuencia de parcialmente complementaria es una que al menos inhibe parcialmente una secuencia completamente complementaria de la hibridación en un ácido nucleico objetivo y se menciona el uso de la expresión funcional "sustancialmente homóloga". La inhibición de la hibridación de la secuencia completamente complementaria a la secuencia objetivo se puede examinar usando un ensayo de hibridación (Southern o Northern blot, hibridación de solución, y similares) en condiciones de baja rigurosidad. Una secuencia sustancialmente homóloga o sonda competirá e inhibirá la unión (es decir, la hibridación) de una secuencia completamente homóloga a un objetivo en condiciones de baja rigurosidad. Esto no quiere decir que las condiciones de baja rigurosidad sean tales que no se permita la unión no específica; las condiciones de baja rigurosidad requieren que la unión de dos secuencias con otra sea una interacción específica (es decir, selectiva). La ausencia de unión no específica puede ser ensayada con el uso de un segundo objetivo que carece de complementariedad o que tiene sólo un bajo grado de complementariedad (por ejemplo, menos de aproximadamente el 30% de complementariedad). En el caso en que la unión específica sea baja o no existente, la sonda no hibridará en un ácido nucleico objetivo. Cuando se usa en referencia a una secuencia de ácidos nucleicos bicatenarios como un cADN o un clon genómico, la expresión "sustancialmente homólogo" se refiere a cualquier oligonucleótido o sonda que puede hibridar en una o ambas cadenas de la secuencia de ácidos nucleicos bicatenarios en condiciones de baja rigurosidad tal como se describió en la presente. Tal como se usan en la presente, los términos "fusión" o "hibridación" se usan en referencia a la formación de pares de cadenas complementarias de ácidos nucleicos. La hibridación y la longitud de la hibridación (es decir, la longitud de la asociación entre cadenas de ácido nucleico) es impactada por muchos factores bien conocidos en la técnica incluyendo el grado de complementariedad entre los ácidos nucleicos, la rigurosidad de las condiciones implicadas afectadas por tales condiciones como la concentración de sales, la  $T_m$  (temperatura de fusión) del híbrido formado, la presencia de otros componentes (por ejemplo, la presencia o ausencia de polietilenglicol o betaina), la molaridad de las cadenas de hibridación y el contenido G:C de cadenas de ácido nucleico.

En general, "cADN" o una "molécula de cADN" se refiere a un "ADN complementario" que es sintetizado por extensión

de un cebador catalizada por ADN polimerasa dependiente de ARN o transcriptasa inversa que se fusiona con una molécula de ARN de interés usando al menos una porción de la molécula de ARN de interés como una plantilla (cuyo proceso también se denomina “transcripción inversa”). Las moléculas de cADN sintetizadas son “homólogas” o “en par de bases con” o “forman un complejo con” al menos una porción de la plantilla.

5 Tal como se usa en la presente, una “población de fragmentos de ADN” se refiere a una pluralidad o colección de fragmentos de ADN, por ejemplo, de un ADN objetivo. En algunas formas de realización, una población de fragmentos de ADN comprende una biblioteca de fragmentos de ADN que comprende secuencias que son cualitativa y/o cuantitativamente representativas de la secuencia del ADN objetivo, si bien en algunas formas de realización, una población de fragmentos de ADN contiene un subgrupo de una biblioteca de ADN, por ejemplo, puede no ser representativa de la secuencia del ADN objetivo.

10 Tal como se usa en la presente, una “biblioteca de fragmentos de ADN” implica una colección o “población de fragmentos” de ADN rotulados (por ejemplo, fragmentos de ADN dirrotulados o fragmentos de ssADN circulares rotulados) generados de ADN objetivo, en donde la combinación de los fragmentos de ADN rotulados en la colección o población exhibe secuencias que son cualitativa y/o cuantitativamente representativas de la secuencia del ADN objetivo del que se generan los fragmentos de ADN rotulados y en donde los fragmentos de ADN rotulados que están en la colección o población no fueron seleccionados o seleccionados contra el uso intencional de un método que incluye o excluye fragmentos de ADN rotulados en base a la composición del nucleótido o la secuencia del ADN objetivo.

20 Por una variedad de motivos, es posible que una biblioteca de fragmentos de ADN pueda no contener un fragmento de ADN rotulado que represente cada secuencia que es exhibida por el ADN objeto. Por ejemplo, en algunas formas de realización, la biblioteca de fragmentos de ADN rotulados puede no contener fragmentos de ADN rotulados que exhiben secuencias de los extremos de un ADN objetivo que comprende dsADN lineal (por ejemplo, debido a una baja frecuencia de inserción de dos composiciones de extremo del transposón en las porciones terminales del ADN objetivo). En general, una baja frecuencia o falta de fragmentos de ADN rotulados que exhiben secuencias de ciertas porciones o regiones del ADN objetivo es aceptable para el fin o la aplicación pretendidos. Sin embargo, la invención también comprende formas de realización del método adicional para aquellas situaciones en las que se considera importante o deseable para un fin o aplicación particular para generar una biblioteca de fragmentos de ADN en donde hay una mayor probabilidad de que los fragmentos de ADN rotulados exhiban cada secuencia que es exhibida por el ADN objetivo del que se generaron los fragmentos (por ejemplo, ver la sección de la memoria descriptiva intitulada (“Métodos para generar bibliotecas de fragmentos de ADN con mejor representación de secuencias en los extremos del ADN objetivo”). Más aún, en algunos casos, la probabilidad de que la biblioteca de fragmentos de ADN contenga un fragmento de ADN rotulado que exhibe cada secuencia del ADN objetivo se incrementará si muchas moléculas de ADN objetivo están presentes en la etapa de reacción de transposición del método, generando así más moléculas de fragmentos de ADN con rótulo en 5' usando el método. Así, otro método más para incrementar la probabilidad de que una biblioteca de fragmentos de ADN contenga un fragmento de ADN rotulado que exhibe cada secuencia que es exhibida por el ADN objetivo consiste en amplificar el ADN objetivo y luego usar el ADN objetivo amplificado en vez del ADN objetivo para generar la biblioteca de fragmentos de ADN. En formas de realización más, en donde el ADN objetivo comprende dsADN preparado de ARN usando una reacción de transcripción inversa, la cantidad de ADN objetivo se amplifica por amplificación del ARN antes de convertirlo en dsADN usando la etapa de transcripción inversa. Algunos métodos de amplificación de ARN y moléculas de ADN que se pueden usar para proporcionar ADN objetivo amplificado se divulgan en la presente. Sin embargo, la divulgación no está limitada respecto del método usado para amplificar el ADN objetivo. En algunas formas de realización, el ADN objetivo se amplifica usando uno de los métodos revelados en la presente, mientras que en algunas otras formas de realización, se usa otro método conocido en la técnica.

45 Tal como se usa en la presente, la expresión “enzima modificadora del ácido nucleico” se refiere a cualquier enzima que actúa sobre el ADN para efectuar una modificación, por ejemplo, clivaje, ligamiento, polimerización, fosforilación, etc. Las enzimas modificadoras del ácido nucleico incluyen, por ejemplo, polimerasas, nucleasas, transferasas, ligasas, fosforilasas, fosfatasa, metilasas, transosasas, etc. “Enzimas modificadoras del ADN” comprenden cualquier enzima que actúa sobre el ADN, incluyendo enzimas que actúan también sobre otros sustratos, como ARN.

50 Tal como se usa en la presente, una “ADN polimerasa” se refiere a una enzima que cataliza la polimerización de desoxirribonucleótidos en una cadena de ADN. Las ADN polimerasas comprenden “ADN polimerasa dependiente de la plantillas”, que requieren un ácido nucleico de plantilla para determinar el orden en el que se añaden los desoxirribonucleótidos al polímero o pueden ser “independientes de la plantilla” de modo que catalicen la polimerización sin referencia a una secuencia de plantilla.

55 Una “ADN polimerasa dependiente del ADN” es una enzima que sintetiza una copia de ADN complementario (“cADN”) por extensión de un cebador que se fusiona con una plantilla de ADN. Algunas ADN polimerasas dependientes del ADN también pueden sintetizar una copia de ADN complementario de una plantilla de ARN, un proceso también mencionado como “transcripción inversa”. Las ADN polimerasas que pueden transcribirse a la inversa también se pueden mencionar como “transcriptasas inversas”.

60 Además de sintetizar los polímeros de ADN, las ADN polimerasas pueden comprender otras características o

actividades. Por ejemplo, una ADN polimerasa se puede caracterizar por tener o carecer de actividad de 5' a 3' exonucleasa (también mencionada como una actividad de 5' exonucleasa o 5' nucleasa), actividad de 3' a 5' exonucleasa, actividad de desplazamiento de cadena y se pueden caracterizar con respecto al grado en que son procesivos o distributivos, tal como se trata más adelante con mayor detalle.

5 Algunas ADN polimerasas son capaces de desplazar la cadena complementaria a la cadena de la plantilla cuando una nueva cadena de ADN es sintetizada por la polimerasa. Este proceso se denomina "desplazamiento de cadena" y las ADN polimerasas que tienen esta actividad se mencionan en la presente como "ADN polimerasas de desplazamiento de cadena". La plantilla para la síntesis de ADN de desplazamiento de cadena puede ser un ADN lineal o circular monocatenario (ssADN) o ADN bicatenario (dsADN). Si la plantilla de ADN es un círculo monocatenario, la síntesis del ADN cebado procede alrededor del círculo, con desplazamiento continuo de la cadena delante de la cadena de replicación, un proceso denominado "replicación del círculo de rodadura". La replicación del círculo de rodadura resulta en la síntesis de copias en tándem de la plantilla circular. En general, se prefiere que una ADN polimerasa específica de una plantilla de ADN usada para un método de la invención sintetice eficazmente el ADN de una longitud apropiada para el fin pretendido sin salir de la plantilla (o terminar la síntesis del ADN), que se menciona como la procesividad de la enzima. La capacidad de una ADN polimerasa de desplazar la cadena puede ser determinada con facilidad usando la polimerasa en un ensayo de replicación del círculo de rodadura tal como describen Fire y Xu (Proc. Natl. Acad. Sci. USA 92: 4641-4645, 1995). El desplazamiento de la cadena y la procesividad de la ADN polimerasa también se pueden ensayar usando métodos descritos en Kong et al. (J. Biol. Chem. 268: 1965-1975, 1993). La transferasa terminal también se define como una ADN polimerasa en la presente, cuya ADN polimerasa se usa como una composición en algunas formas de realización de los kits relacionados con, y métodos de la presente invención. La transferasa terminal se prefiere en algunas formas de realización porque cataliza la adición independiente de la plantilla de dNTPs a los términos de 3'-hidroxilo de ADN.

Algunas formas de realización comprenden un método que usa una composición de ADN polimerasa que tiene actividad de 5' a 3' exonucleasa para liberar un nucleótido que está rotulado con un resto detectable (por ejemplo, un resto que comprende una molécula visible, fluorescente, quimioluminiscente u otra molécula detectable) como un medio para ensayar la polimerización del ADN y así, detectar y/o cuantificar la presencia en la muestra de la molécula de ácido nucleico que sirve como la plantilla (por ejemplo, de una manera similar a los ensayos TagMan® de Applied Biosystems, Inc.). En algunas formas de realización, la presente invención comprende una composición de ADN polimerasa que carece de actividad de 5' a 3' exonucleasa.

30 Algunas formas de realización comprenden un método que usa una composición de ADN polimerasa que carece de actividad de 5' a 3' exonucleasa. Por ejemplo, en algunas formas de realización, una composición de ADN polimerasa que carece de actividad de 5' a 3' exonucleasa se usa para la secuenciación de ADN. Por ejemplo, en algunas otras formas de realización, una composición de ADN polimerasa que carece de actividad de 5' a 3' exonucleasa se usa para la amplificación genómica completa.

35 En algunas formas de realización, la presente invención comprende una composición de ADN polimerasa que tiene actividad de 5' a 3' exonucleasa. En algunas formas de realización de preferencia (por ejemplo, en donde se usa una ADN polimerasa, además de una ligasa dependiente de la plantilla, para unión en uno de los métodos descritos en la presente), el método usa una composición de ADN polimerasa que carece de actividad de 5' nucleasa (incluyendo tanto la actividad de 5' a 3' exonucleasa como la actividad de nucleasa dependiente de la estructura 5'). Por ejemplo, en algunas otras formas de realización, una composición de ADN polimerasa que carece de actividad de 5' a 3' exonucleasa se usa para rellenar una brecha. Así, en algunas formas de realización de métodos o kits, la presente invención comprende una composición de ADN polimerasa que carece de actividad de 5' nucleasa. Sin embargo, una composición de ADN polimerasa que tiene actividad de 5' nucleasa para liberar un nucleótido o un oligonucleótido que está rotulado con un resto detectable (por ejemplo, un resto que comprende una molécula visible, fluorescente, quimioluminiscente u otra molécula detectable) como un medio para ensayar la polimerización de ADN y así, detectar y/o cuantificar la presencia en la muestra de la molécula de ácido nucleico que sirve como la plantilla (por ejemplo, de una manera similar a los ensayos TagMan® de Applied Biosystems, Inc.) se podría usar para cuantificar moléculas de ADN generadas usando un método de la invención.

Los ejemplos de ADN polimerasas que desplazan la cadena que se pueden usar incluyen, pero no se limitan a, RepliPHI™ phi29 ADN polimerasa, DisplaceAce™ ADN polimerasa, rGka ADN polimerasa, SequiTherm™ ADN polimerasa, Taq ADN polimerasa, Tfl ADN polimerasa y MMLV transcriptasa inversa (todas disponibles de EPICENTRE Biotechnologies, Madison, WI, Estados Unidos). En algunas formas de realización, se usa una mezcla de una ADN polimerasa que carece de actividad de verificación de 3' a 5' exonucleasa con una ADN polimerasa que tiene esta actividad, como FAILSAFE™ ADN polimerasa como la ADN polimerasa que desplaza la cadena. La mezcla de enzimas es de utilidad en algunas formas de realización porque exhibe una mejor fidelidad durante la síntesis del ADN (es decir, sintetiza el ADN con menor cantidad de nucleótidos que no son complementarios a la plantilla). Las tasas de fidelidad y/o error de muchas ADN polimerasas en condiciones particulares son conocidas, cuando son métodos de medición de la fidelidad (por ejemplo, por secuenciación).

En general, es deseable en un método de amplificación de desplazamiento de cadena de la presente divulgación que la cantidad de la ADN polimerasa de desplazamiento de cadena usada en el método es lo más alta posible sin inhibir o afectar adversamente la reacción. Por ejemplo, REPLIPHI™ phi29 ADN polimerasa (EPICENTRE) se puede usar

en aproximadamente un microgramo de proteína de una reacción de 20 microlitros y DISPLACE™ ADN polimerasa (EPICENTRE) se puede usar en aproximadamente 50 unidades a aproximadamente 300 unidades en una reacción de 50 microlitros. Cuando las definiciones para unidades varían para las diferentes ADN polimerasas e incluso para similares ADN polimerasas de diversos vendedores o fuentes y también porque la actividad para cada enzima varía a diferentes temperaturas y en diversas condiciones de reacción, se desea optimizar la cantidad de ADN polimerasa que desplaza la cadena y condiciones de reacción para cada plantilla de ADN y el cebador usados.

El desplazamiento de cadena se puede facilitar por medio del uso de un factor de desplazamiento de cadena, como helicasa, pero como se puede usar una variedad de ADN polimerasas para la presente invención, tal como un factor de desplazamiento de cadena no se requiere usualmente. Se considera que cualquier ADN polimerasa que puede realizar la replicación de círculo de rodadura en presencia de un factor de desplazamiento de cadena es apropiada para usar en formas de realización de la invención que comprenden desplazamiento de cadena incluso si la ADN polimerasa no lleva a cabo la replicación del círculo de rodadura en ausencia de otro factor. Los factores de desplazamiento de cadena que permiten la replicación del círculo de rodadura incluyen, pero no se limitan a, la subunidad accesoria de BMRFI polimerasa (Tsurumi et al., *J. Virology*, 67: 7648-7653, 1993), proteína de unión a ADN de adenovirus (Zijderveld y van der Vliet, *J. Virology*, 68: 1158-1164, 1994), proteína del virus herpes simplex ICP8 (Boehmer y Lehman, *J. Virology*, 67: 711-715, 1993); Skaliter y Lehman, *Proc. Natl. Acad. Sci. USA*, 91: 10.665-10.669, 1994), proteínas de unión de ADN monocatenario (SSB; Rigler y Romano, *J. Biol. Chem.*, 270: 8910-8919, 1995) y helicasa de timo de ternero (Siegel et al., *J. Biol. Chem.*, 267: 13.629-13.635, 1992).

Un "mononucleósido" o "nucleósido", tal como se usa en la presente, se refiere a un compuesto que consiste en una base de purina (guanina (G) o adenina (A)) o pirimidina (timina (T), uridina (U) o citidina (C)) ligada covalentemente con un azúcar pentosa, mientras que "nucleótido" se refiere a un nucleósido fosforilado en uno de los grupos hidroxilo del azúcar pentosa. El término "canónico" se usa para referirse a las cuatro bases de ácidos nucleicos comunes adenina, citosina, guanina y timina que se hallan comúnmente en el ADN o a los respectivos desoxirribonucleósidos, desoxirribonucleótidos o 2'-desoxirribonucleósido-5'-trifosfatos que contienen una base canónica. La expresión "no canónica" se usa para referirse a bases de ácido nucleico en ADN distintas de las cuatro bases canónicas o los respectivos desoxirribonucleósidos, desoxirribonucleótidos o 2'-desoxirribonucleósido-5'-trifosfatos que contienen una base no canónica. Por ejemplo, a pesar de que el uracilo es una base de ácido nucleico común en ARN, el uracilo es una base no canónica en ADN. "Las bases no canónicas" se hallan en ácidos nucleicos como un resultado de la incorporación de nucleótido no canónico (por ejemplo, por síntesis usando un sintetizador de oligonucleótidos o por síntesis usando una ADN polimerasa) o como resultado de modificación de bases existentes (canónicas o no canónicas).

Un "ácido nucleico" o "polinucleótido" implica una molécula polimérica que comprende una serie de "mononucleósidos", también mencionados como "nucleósidos", en donde la posición 3' del azúcar pentosa de un nucleósido está ligado por un ligamiento de internucleósido, tales como, pero no limitados a, un enlace fosfodiéster, con la posición 5' del azúcar pentosa del siguiente nucleósido. Un nucleósido ligado con un grupo fosfato se menciona como un "nucleótido". El nucleótido que está ligado a la posición 5' del siguiente nucleótido en la serie se menciona como el "5' de" o el "5' nucleótido" y el nucleótido que está ligado con la posición 3' del 5' nucleótido se menciona como "3' de" o el "3' nucleótido". Tal como se usan en la presente, los términos "5' de" y "3' de" se refieren a la posición u orientación de un grupo químico particular, nucleótido, secuencia de nucleótidos o elemento genético (por ejemplo, una secuencia promotora de ARN polimerasa) respecto de otro grupo químico, nucleótido, secuencia de nucleótidos o elemento genético dentro de una cadena simple de un ácido nucleico. Si una primera secuencia de ácidos nucleicos es 3' de una segunda secuencia en una cadena, el complemento de la primera secuencia estará 5' del complemento de la segunda secuencia en la cadena complementaria. La descripción de la invención se comprenderá con respecto a la posición relativa 5' o 3' y la orientación de una secuencia o elemento genético dentro de una cadena de ácido nucleico particular.

Se dice que las moléculas de ácidos nucleicos lineales tienen un "término 5'" (extremo 5') y un "término 3'" (extremo 3') porque los ligamientos de fosfodiéster del ácido nucleico se producen en el carbono 5' y carbono 3' de los restos de azúcar de los mononucleótidos sustituyentes. El extremo de un polinucleótido en el que estará un nuevo ligamiento con un carbono 5' es su nucleótido 5' terminal. El extremo de un polinucleótido en el que estará un nuevo ligamiento con un 3' carbono es su nucleótido 3' terminal. Un nucleótido terminal, tal como se usa en la presente, es el nucleótido en la posición final del término 3' o 5'.

El azúcar pentosa del ácido nucleico puede ser ribosa, en cuyo caso, el ácido nucleico o polinucleótido se menciona como "ARN", o puede ser 2'-desoxirribosa, en cuyo caso, el ácido nucleico o polinucleótido se menciona como "ADN". Alternativamente, en especial si el ácido nucleico se sintetiza químicamente, el ácido nucleico puede estar compuesto por mononucleótidos tanto de ADN como de ARN. En ambos ARN y ADN, cada azúcar pentosa está ligada covalentemente con una de cuatro bases comunes o "canónicas" de ácido nucleico (cada una mencionada también como una "base"). Tres de las bases naturales predominantes que se ligan con los azúcares (adenina, citidina y guanina) son comunes tanto para ADN como para ARN, mientras una base es diferente; el ADN tiene la base adicional timina, mientras que el ARN tiene la base adicional uridina. En algunos casos, la uridina puede estar presente como una base en ADN. Los expertos en la técnica comúnmente piensan en un pequeño polinucleótido como un "oligonucleótido". El término "oligonucleótido" tal como se usa en la presente se define como una molécula que comprende dos o más desoxirribonucleótidos o ribonucleótidos, con preferencia de aproximadamente 6 a 100

nucleótidos, pero no hay un límite definido para la longitud de un oligonucleótido. El tamaño exacto dependerá de muchos factores que, a su vez, depende de la última función o el uso del oligonucleótido.

Además, por una variedad de razones, un ácido nucleico o polinucleótido de la invención puede comprender una o varias bases de ácido nucleico modificadas, restos de azúcar o ligaciones de internucleósido. A modo de ejemplo, algunas razones para usar ácidos nucleicos o polinucleótidos que contienen bases modificadas, restos de azúcar o ligaciones de internucleósido incluyen, pero no se limitan a: (1) modificación de la  $T_m$ ; (2) cambio de la susceptibilidad del polinucleótido a una o varias nucleasas; (3) suministro de un resto para unir una etiqueta; (4) suministro de una etiqueta o un neutralizador para una etiqueta; o (5) suministro de un resto como biotina, para unirse con otra molécula que está en solución o ligada a una superficie. Por ejemplo, en algunas formas de realización, un oligonucleótido como un cebador puede ser sintetizado, de modo que una porción aleatoria contenga uno o varios análogos de ácido ribonucleico de conformación restringida tales como, pero no limitados a uno o varios análogos de ácido ribonucleico en los que un anillo ribosa está "bloqueado" con un puente de metileno que conecta el átomo 2'-O con el átomo 4'-C (por ejemplo, disponible de Exiqon, Inc. bajo la marca registrada de "LNA<sup>TM</sup>"); estos nucleótidos modificados resultan en un aumento de la  $T_m$  o temperatura de fusión en aproximadamente 2 grados a aproximadamente 8 grados centígrados por monómero de nucleótido. Si la  $T_m$  está aumentada, puede ser posible reducir la cantidad de nucleótidos aleatorios en la posición 3' aleatoria del oligoribonucleótido de rotulación terminal. Sin embargo, un nucleótido modificado como un LNA debe ser validado para funcionar en el método para su fin pretendido, así como satisfacer otros criterios del método. Por ejemplo, en algunas formas de realización en las que se usa un cebador de oligonucleótido que comprende ribonucleótidos, un criterio para usar el nucleótido modificado en el método puede ser que el oligonucleótido que lo contiene puede ser digerido por una ARNasa específica de la cadena simple.

A fin de lograr los objetivos de la invención, a modo de ejemplo, las bases de ácido nucleico en los mononucleótidos de una o varias posiciones de un polinucleótido u oligonucleótido pueden comprender guanina, adenina, uracilo, timina o citidina o, alternativamente, una o varias de las bases de ácido nucleico pueden comprender una base modificada como, pero sin limitación, xantina, alilamino-uracilo, alilaminotimidina, hipoxantina, 2-aminoadenina, 5-propiniluracilo, 5-propinilcitosina, 4-tiouracilo, 6-tioguanina, aza y deazauracilos, timidinas, citosinas, adeninas o guaninas. Además, pueden comprender una base de ácido nucleico que se deriva con un resto de biotina, un resto de digoxigenina, un resto fluorescente o quimioluminiscente, un resto neutralizante o algún otro resto. La invención no se limita a las bases de ácido nucleico listadas; esta lista se brinda para mostrar un ejemplo del amplio rango de bases que se pueden usar para un fin particular en un método.

Respecto de los ácidos nucleicos o polinucleótidos de la invención, uno o varios de los restos de azúcar pueden comprender 2'-desoxirribosa o, alternativamente, uno o varios de los restos de azúcar pueden ser algún otro resto de azúcar tales como, pero sin limitación, ribosa o 2'-fluoro-2'-desoxirribosa o 2'-O-metil-ribosa, que proporcionan resistencia a ciertas nucleasas o 2'-amino-2'-desoxirribosa o 2'-azido-2'-desoxirribosa, que pueden ser etiquetadas haciéndolas reaccionar con tinturas o productos químicos visibles, fluorescentes, fluorescentes infrarrojos u otras tinturas o productos químicos detectables que tienen un resto electrofílico, fotorreactivo, de alquilo u otro resto químico reactivo.

Los ligamientos de internucleósido de ácidos nucleicos o polinucleótidos de la invención pueden ser ligamientos de fosfodiéster o, alternativamente, uno o varios de los ligamientos de internucleósido pueden comprender ligamientos modificados tales como, pero sin limitación, fosforotioato, fosforoditioato, fosforoselenato o fosforodiselenato que son resistentes a ciertas nucleasas.

Cuando se refiere a un oligonucleótido o una porción de un oligonucleótido que exhibe una "secuencia aleatoria", eso implica que el oligonucleótido o su porción están sintetizados (por ejemplo, usando un sintetizador de oligonucleótido) usando cantidades iguales de todas las cuatro bases canónicas de nucleótido (A, G, C y T o U) para cada posición de nucleótido dentro de la porción de secuencia aleatoria. Este método resulta en la síntesis de una mezcla de oligonucleótidos que comprende  $(4^n + 1)$  de diferentes oligonucleótidos, donde "n" es igual a la cantidad de posiciones de nucleótido dentro de la porción de secuencia aleatoria. Así, en estas formas de realización, el oligonucleótido comprende una mezcla de muchos oligonucleótidos diferentes, que representan todas las posibles secuencias de la porción de secuencia aleatoria. Cuando se refiere a un oligonucleótido o una porción de un oligonucleótido que exhibe una "secuencia semialeatoria", esto implica que el oligonucleótido semialeatorio o la porción se sintetiza (por ejemplo, usando un sintetizador de oligonucleótido) en donde algunas posiciones de nucleótido se sintetizan usando iguales cantidades de todas las cuatro de las bases canónicas de nucleótido (A, G, C y T o U) (es decir, aquellas posiciones son "aleatorias" tal como se describió con anterioridad), pero una o varias otras posiciones dentro de la porción semialeatoria se sintetizan usando sólo una, dos o tres, más que las cuatro de las bases canónicas de nucleótidos (es decir, A, C, G y T o U). En algunas formas de realización, un oligonucleótido contiene uno o varios pares de bases con una o varias bases de ácido nucleico distintas de las reglas de formación de pares de bases estándar donde A forma un par con T o U y G forma un par con C y un "nucleótido degenerado" es un nucleótido que contiene una base degenerada. Una "porción" o "región", usadas de modo indistinto en la presente, de un polinucleótido o un oligonucleótido (incluyendo un cebador) es una secuencia contigua de 2 o más bases. En otras formas de realización, una región o porción es de al menos aproximadamente cualquiera de 1, 2, 3, 5, 10, 15, 20, 25, 50, 75 o incluso más nucleótidos contiguos.

Un “cebador” es un oligonucleótido (“oligo”), en general con un grupo 3'-OH libre, que se puede extender por una polimerasa de ácido nucleico. Para una polimerasa dependiente de la plantilla, en general al menos la porción 3' del cebador oligo es complementaria a una porción de un ácido nucleico de plantilla, a la que se “une” el oligo (o “complejos”, “se fusiona”, o “hibrida”), por ligación de hidrógeno y otras fuerzas moleculares, a la plantilla para dar un complejo de cebador/plantilla para inicio de la síntesis por una ADN polimerasa y que se extiende (es decir, “el cebador extendido”) por adición de bases covalentemente ligadas con su extremo 3' que son complementarias a la plantilla en el proceso de síntesis de ADN. El resultado es un producto de extensión del cebador. Las ADN polimerasas dependientes de la plantilla (incluyendo transcriptasas inversas) requieren en general la complejización de un cebador de oligonucleótido en una plantilla monocatenaria para iniciar la síntesis del ADN (“cebado”), pero las ARN polimerasas en general no requieren un cebador para la síntesis de ARN que es complementario a una plantilla de ADN (transcripción).

Una “ADNasa específica de la cadena simple” implica una ADNasa que digiere específicamente el ADN monocatenario, pero no digiere ARN o ARN monocatenario o ADN que se fusiona o se complejiza con ARN o ADN complementario, ya sea que dicho ARN o ADN complementario sea parte de otra molécula de ácido nucleico (por ejemplo, por formación de pares de bases intermoleculares) o una porción de la misma molécula de ácido nucleico (por ejemplo, por formación de bases intramoleculares). La ADNasa específica de la cadena simple puede ser una endonucleasa o una exonucleasa, siempre que sea activa en la digestión específica de ADN monocatenario en monómeros u oligodesoxirribonucleótidos cortos. En algunas formas preferidas de realización, los oligodesoxirribonucleótidos, incluyendo cebadores, se remueven de la mezcla de reacción después de la etapa del método en la que se usan por digestión con una ADNasa específica de cadena simple. La exonucleasa I, exonucleasa VII y Rec J2278 exonucleasa son ADNasas específicas de cadena simple de ejemplo.

Una “ARN polimerasa de tipo T7” (RANP) significa aquí T7 ARN polimerasa (por ejemplo, ver Studier, FW et al., pp. 60-89 en *Methods in Enzymology*, Vol. 185, ed. por Goeddel, DV, Academic Press, 1990) o una RANP derivada de un “bacteriófago de tipo T7”, que implica un bacteriófago que tiene una organización genética similar a la del bacteriófago T7. La organización genética de todos los fagos de tipo T7 que se examinaron se hallaron esencialmente iguales a la de T7. Ejemplos de bacteriófagos de tipo T7 incluyen, pero no se limitan a fagos de *Escherichia coli* T3, phi I, phi II, W31, H, Y, AI, 122, cro, C21, C22 y C23; fago de *Pseudomonas putida* gh-1; fago de *Salmonella typhimurium* SP6; fagos de *Serratia marcescens* IV; fago de *Citrobacter* Villi; y fago de *Klebsiella* No. 11 (Hausmann, *Current Topics in Microbiology and Immunology* 75:77-109, 1976; Korsten et al., *J. Gen. Virol.* 43:57-73, 1975; Dunn, et al., *Nature New Biology* 230:94-96, 1971; Towle, et al., *J. Biol. Chem.* 250:1723-1733, 1975; Butler y Chamberlin, *J. Biol. Chem.* 257:5772-5778, 1982), así como formas mutantes de tales RNAP (por ejemplo, Sousa et al., patente U.S. No. 5.849.546; Padilla, R y Sousa, R, *Nucleic Acids Res.*, 15: e138, 2002; Sousa, R y Mukherjee, S, *Prog Nucleic Acid Res Mol Biol.*, 73: 1-41, 2003; Guillerez, J, et al., solicitud de patente U.S. No. 20040091854). En los ejemplos preferidos, el promotor usado es una secuencia promotora de tipo salvaje o mutante que es reconocida por una ARN polimerasa de tipo T7. En algunas formas de realización, el promotor puede ser monocatenario, como un pseudopromotor (por ejemplo, Ohmichi et al., *Proc. Natl. Acad. Sci. USA* 99:54-59, 2002) o un promotor N4 vRNAP, en cuyo caso se usa la proteína truncada comprende el dominio de aminoácido 1.106 transcripcionalmente activo (correspondientes a los aminoácidos 998-2103) de N4 vRNAP (designado “mini-vRNAP”; EPICENTRE Biotechnologies, Madison, WI, Estados Unidos) (Kazmierczak, K.M., et al., *EMBO J.*, 21: 5815-5823, 2002).

Tal como se usa en la presente, “ADN objetivo” se refiere a cualquier dsADN de interés que se somete a transposición, por ejemplo, para generar una biblioteca de fragmentos de ADN rotulados (por ejemplo, fragmentos de ssADN o de dsADN lineales con rótulo en 5' y 3' o fragmentos de ssADN circulares rotulados).

“ADN objetivo” se puede derivar de cualquier fuente in vivo o in vitro, incluyendo una o varias células, tejidos, órganos u organismos, ya sea vivos o muertos o de cualquier fuente biológica o ambiental (por ejemplo, agua, aire, suelo). Por ejemplo, en algunas formas de realización, el ADN objetivo comprende o consiste en dsADN eucariota y/o procarionta que se origina o que se deriva de seres humanos, animales, plantas, hongos (por ejemplo, mohos o levaduras), bacterias, virus, viroides, micoplasma u otros microorganismos. En algunas formas de realización, el ADN objetivo comprende o consiste en ADN genómico, subADN genómico, ADN cromosómico (por ejemplo, de un cromosoma aislado o una porción de un cromosoma, por ejemplo, de uno o varios genes o loci de un cromosoma), ADN mitocondrial, ADN de cloroplasto, plásmido u otro ADN derivado episomal (o ADN recombinante contenido dentro) o cADN bicatenario hecho por transcripción inversa de ARN usando una ADN polimerasa dependiente de ARN o transcriptasa inversa para generar cADN de primera cadena y luego extendiendo un cebador fusionado con el cADN de primera cadena para generar dsADN. En algunas formas de realización, el ADN objetivo comprende múltiples moléculas de dsADN o se prepara de moléculas de ácidos nucleicos (por ejemplo, múltiples moléculas de dsADN o preparadas de ADN genómico o cADN preparado de ARN o de una fuente biológica (por ejemplo, célula, tejido, órgano, organismo) o ambiental (por ejemplo, agua, aire, suelo, saliva, esputo, orina, heces). En algunas formas de realización, el ADN objetivo proviene de una fuente in vitro. Por ejemplo, en algunas formas de realización, el ADN objetivo comprende o consiste en dsADN que se prepara in vitro de ADN monocatenario (ssADN) o de ARN monocatenario o bicatenario (por ejemplo, usando métodos que son bien conocidos en la técnica, como la extensión del cebador usando una ADN polimerasa dependiente de ADN y/o dependiente de ARN (transcriptasa inversa). En algunas formas de realización, el ADN objetivo comprende o consiste en dsADN que se prepara de todas o una porción de una o varias moléculas de ADN o ARN bicatenario o monocatenario usando cualquier método conocido en la técnica, incluyendo métodos para: amplificación de ADN o ARN (por ejemplo, PCR o PCR por transcriptasa inversa (RT-PCR), métodos

de amplificación mediada por transcripción, con amplificación de todas o una porción de una o varias moléculas de ácido nucleico); clonación molecular de todas o una porción de una o varias moléculas de ácidos nucleicos en un plásmido, fósido, BAC u otro vector que se replica posteriormente en una célula huésped apropiada; o captura de una o varias moléculas de ácidos nucleicos por hibridación, como por hibridación en sondas de ADN en una matriz o micromatriz (por ejemplo, por “captura de secuencias”; por ejemplo, usando kits y/o matrices de ROCHE NIMBLEGEN, AGILENT o FEBIT).

En algunas formas de realización, “ADN objetivo” significa dsADN que se prepara o modifica (por ejemplo, usando diversas técnicas bioquímicas o biológicas moleculares) antes de ser usado para generar una biblioteca de fragmentos de ADN rotulados (por ejemplo, ssADN lineal con rótulo en 5' y 3' o dirrotulado o fragmentos de dsADN o fragmentos de ssADN circulares rotulados). Por ejemplo, los presentes inventores observaron que la representación de los datos de secuencia de próxima generación de los extremos de ADN objetivo que comprende moléculas de dsADN con un tamaño inferior a 10 Kb era baja en comparación con la representación de los datos de secuencia de la mitad de ese ADN objetivo. Sin quedar ligado por la teoría, una posible explicación para esta observación es que la probabilidad de hallar fragmentos de ADN con dos composiciones de extremo del transposón insertadas en orientaciones opuestas en los extremos de una molécula de dsADN lineal es menor que la probabilidad de hallar fragmentos de ADN con dos composiciones de extremo del transposón insertados en orientaciones opuestas en la mitad de la molécula de dsADN lineal. Así, en algunas formas de realización, a fin de generar bibliotecas de fragmentos de ADN dirrotulados o fragmentos de ADN circulares rotulados que representan mejor las secuencias terminales, el método también comprende proporcionar ADN objetivo para usar en el método que comprende dsADN (por ejemplo, ADN bicatenario genómico o cADN preparado a partir de ARN, como mARN) que ya tiene un rótulo en el extremo 5' y/o 3'. Por ejemplo, en algunas formas de realización, el ADN objetivo comprende cADN bicatenario que se prepara de ARN: sintetizando cADN de primera cadena extendiendo un cebador de síntesis de cADN de primera cadena que tiene una porción 3' y una porción 5', en donde la porción 3' es complementaria a la porción de extremo 3' del ARN y la porción 5' comprende un primer rótulo, luego uniendo un segundo rótulo con el extremo 3' del cADN de primera cadena usando un oligonucleótido de rótulo terminal y una ADN polimerasa tal como se describe en otra parte de la presente y luego usando una ADN polimerasa para sintetizar cADN bicatenario extendiendo un cebador de síntesis de cADN de segunda cadena que se fusiona con el segundo rótulo. De modo alternativo, en algunas otras formas de realización preferidas, a fin de generar bibliotecas de fragmentos de ADN dirrotulados que representan mejor las secuencias terminales, el ADN objetivo usado en el método para generar fragmentos de ADN dirrotulados o fragmentos de ADN circulares rotulados comprende dsADN circular que se prepara por ligamiento intramolecular de dsADN lineal (por ejemplo, que se prepara por ligamiento intramolecular de ADN bicatenario genómico o de cADN bicatenario preparado a partir de ARN, como mARN). Así, en algunas formas de realización, el método también comprende: ligación de dsADN lineal usando una ligasa (por ejemplo, T4 ADN ligasa) para generar dsADN circular para usar como ADN objetivo en el método. En algunas formas de realización del método que comprende generar dsADN circular para usar como ADN objetivo ligando dsADN lineal, el dsADN lineal se trata con T4 ADN polimerasa y T4 polinucleótido quinasa (por ejemplo, usando END-It™ ADN End Repair Kit (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos) anterior a la etapa de ligamiento a fin de preparar terminales romos y fosforilar los extremos 5'.

Tal como se usa en la presente, un “fragmento de ADN” implica una porción o parte o segmento de un ADN objetivo que se cliva o se libera o se rompe de una molécula de ADN más larga de modo que ya no está unida con la molécula principal. Un fragmento de ADN puede ser bicatenario (un “fragmento de dsADN”) o monocatenario (un “fragmento de ssADN”) y el proceso de generación de fragmentos de ADN a partir del ADN objetivo se menciona como “fragmentación” del ADN objetivo. En algunas formas de realización de preferencia, el método se usa para generar una “biblioteca de fragmentos de ADN” que comprende una colección o población de fragmentos de ADN rotulados.

Una “plantilla” es una molécula de ácido nucleico que es copiada por una polimerasa de ácido nucleico, como una ADN polimerasa. Ya sea que la molécula de ácido nucleico comprenda dos cadenas (es decir, es “bicatenaria”) o sólo una cadena (es decir, es “monocatenaria”), la cadena de dicha molécula de ácido nucleico que sirve para especificar la secuencia de nucleótidos exhibida por un ácido nucleico que se sintetiza es la “plantilla” o “la cadena de plantilla”. El ácido nucleico sintetizado por la polimerasa de ácido nucleico es complementario a la plantilla. Tanto el ARN como el ADN se sintetizan siempre en la dirección 5' a 3', comenzando en el extremo 3' de la cadena de plantilla y las dos cadenas de un dúplex de ácido nucleico siempre se alinean de modo tal que los extremos 5' de las dos cadenas están en extremos opuestos del dúplex (y, por necesidad, así son los extremos 3'). Un cebador se requiere para las plantillas de ARN y ADN para iniciar la síntesis por una ADN polimerasa, pero no se requiere un cebador para iniciar la síntesis por una ARN polimerasa dependiente de ADN, que usualmente se denomina de modo simple una “ARN polimerasa”.

“Transferasa terminal”, también mencionada como “desoxirribonucleotidiltransferase terminal” o “TdT”, es una ADN polimerasa que cataliza la adición independiente de la plantilla (o “cola”) de trifosfatos de desoxirribonucleósido (dNTPs) o un trifosfato de didesoxirribonucleósido simple en los términos de 3'-hidroxilo de ADN. Una transferasa terminal común usado en la técnica, que es asequible en comercios, se produce en una cepa de E. coli que expresa el gen recombinante de timo de ternero. En algunas formas de realización, la invención también comprende la etapa de incubación de fragmentos de ADN rotulados, después de desnaturalización, con TdT y dNTP en condiciones y durante un tiempo suficiente, en donde los fragmentos de ADN rotulados en 5' y 3' que tienen un segundo rótulo que comprende una cola de ADN homopolimérico se sintetizan. En algunas formas de realización, la cola de ADN homopolimérico se usa luego como un sitio cebador para la síntesis de cADN bicatenario. En algunas formas de realización, el cebador usado para sintetizar la segunda cadena de ADN tiene una porción 3' que es complementaria

al segundo rótulo que comprende la cola homopolimérica y una porción 5' que exhibe una secuencia deseada arbitraria que no es complementaria al primer rótulo, el ADN objetivo o el segundo rótulo que comprende la cola homopolimérica. Por ejemplo, en algunas formas de realización, la porción 5' del cebador exhibe una secuencia promotora antisentido para un promotor de ARN polimerasa y el método también comprende la incubación del cADN bicatenario resultante con la ARN polimerasa en condiciones y durante un tiempo suficiente en que se sintetiza ARN.

En algunas formas de realización, los oligonucleótidos del extremo del transposón usado en el método de la presente divulgación exhiben sólo las secuencias de extremo del transposón necesarias en una reacción de transposición. Sin embargo, en algunas formas de realización, al menos un oligonucleótido del extremo del transposón exhibe adicionalmente una o varias otras secuencias de nucleótidos 5' de la secuencia de extremo del transposón. Así, en algunas formas de realización, el método o kit usa una cadena transferida que tiene una porción 3' y una porción 5', en donde la porción 3' la secuencia del extremo del transposón transferida y la porción 5' exhiben una o varias secuencias adicionales que no participan en la formación de un complejo funcional con la transposasa. No hay límite para usar secuencias adicionales APRA una o varias secuencias adicionales en la porción 5' de la cadena transferida, cuyas secuencias se pueden usar para llevar a cabo cualquier finalidad deseada. Por ejemplo, en algunas formas de realización de la invención, la porción 5' de la cadena transferida exhibe una o varias secuencias de rótulo adicionales (por ejemplo, una secuencia de rótulo que permite la captura por fusión con una secuencia específica sobre una superficie, como una perla o una sonda en un microchip o matriz; por ejemplo, para la captura en una perla para secuenciación de próxima generación; por ejemplo, una secuencia de rótulo 454A o 454B para la captura en la perla para secuenciación usando un secuenciador Roche 454 Next-Gen) o una o varias secuencias para identificación, detección (por ejemplo, detección fluorescente) o clasificación de los productos del método. En algunas otras formas de realización, la porción 5' de la cadena transferida exhibe uno o varios nucleótidos adicionales o secuencias o un grupo químico o resto que comprende o consiste en una unión por afinidad (por ejemplo, una secuencia de rótulo que permite la captura por fusión con una secuencia específica sobre una superficie como una perla o una sonda en un microchip o matriz. En algunas formas de realización de preferencia, el tamaño de las una o varias secuencias adicionales en la porción 5' de la cadena transferida se minimizan a fin de minimizar la probabilidad o la frecuencia de inserción de la cadena transferida en sí misma durante la reacción in vitro de transposasa. Por ejemplo, en algunas formas de realización, el tamaño de la porción 5' de la cadena transferida es inferior a aproximadamente 150 nucleótidos, inferior a aproximadamente 100 nucleótidos, inferior a aproximadamente 75 nucleótidos, inferior a aproximadamente 50 nucleótidos, inferior a aproximadamente 25 nucleótidos o inferior a aproximadamente 15 nucleótidos.

En algunas formas de realización, el extremo 5' de la cadena transferida tiene grupo 5'-monofosfato. En algunas formas de realización, tanto la cadena transferida como la cadena no transferida tienen un grupo 5'-monofosfato. En algunas formas de realización de preferencia, sólo el extremo 5' de la cadena no transferida tiene un grupo 5'-monofosfato. En algunas otras formas de realización, no hay un grupo 5'-monofosfato en el extremo 5' de la cadena transferida.

El término "transposasa" con respecto a la presente invención pretende implicar una enzima capaz de formar un complejo funcional con un extremo de transposón o una secuencia de extremo de transposón necesaria en una reacción de transposición. Una transposasa de la invención también incluye integrasas de retrotransposones y retrovirus.

Una "reacción de transposición" es una reacción en la que uno o varios extremos del transposón se insertan en un ADN objetivo en sitios aleatorios o sitios casi aleatorios. Los componentes esenciales en una reacción de transposición son una transposasa y oligonucleótidos de ADN que exhiben las secuencias de nucleótido del extremo de transposón, incluyendo la secuencia del extremo del transposón transferida y su complemento, la secuencia del extremo del transposón no transferida, así como otros componentes necesarios para formar un complejo de transposición funcional. El método de esta invención se ejemplifica empleando un complejo de transposición formado por una Tn5 transposasa hiperactiva y un extremo de transposón de tipo Tn5 (Goryshin, I. y Reznikoff, W.S., *J. Biol. Chem.*, 273: 7367, 1998) o una MuA transposasa y un extremo de transposón Mu que comprende secuencias terminales R1 y R2 (Mizuuchi, K., *Cell*, 35: 785, 1983; Savilahti, H, et al., *EMBO J.*, 14: 4893, 1995). Sin embargo, cualquier sistema de transposición que es capaz de insertar un extremo de transposón en una manera aleatoria o casi aleatoria con suficiente eficacia en el rótulo 5' y fragmento a ADN objetivo para su uso pretendido se puede usar en la presente invención. Los ejemplos de sistemas de transposición conocidos en la técnica que podrían ser evaluados para los presentes métodos incluyen, pero sin limitación, *Staphylococcus aureus* Tn552 (Colegio OR et al., *J Bacteriol.*, 183: 2384-8, 2001; Kirby C et al., *Mol Microbiol.*, 43: 173-86, 2002), Tyl (Devine SE y Boeke JD., *Nucleic Acids Res.*, 22: 3765-72, 1994 y la solicitud de patente internacional No. WO 95/23875), Transposon Tn7 (Craig, NL, *Science*. 271: 1512, 1996; Craig, NL, Review en: *Curr Top Microbiol Immunol.*, 204: 27-48, 1996), Tn10 e IS10 (Kleckner N, et al., *Curr Top Microbiol Immunol.*, 204: 49-82, 1996), Mariner transposasa (Lampe DJ, et al., *EMBO J.*, 15: 5470-9, 1996), Tel (Plasterk RH, *Curr Top Microbiol Immunol*, 204: 125-43, 1996), P Element (Gloor, GB, *Methods Mol Biol.*, 260: 97-114, 2004), Tn3 (Ichikawa H y Ohtsubo E., *J Biol Chem*. 265: 18829- 32, 1990), secuencias de inserción bacteriana (Ohtsubo, F y Sekine, Y, *Curr. Top. Microbiol. Immunol*. 204: 1-26, 1996), retrovirus (Brown PO, et al., *Proc Natl Acad Sci USA*, 86: 2525-9, 1989) y retrotransposón de levadura (Boeke JD y Corces VG, *Annu Rev Microbiol*. 43: 403-34, 1989).

El método para insertar un extremo de transposón en una secuencia objetivo se puede llevar a cabo in vitro usando

un sistema de transposón apropiado para el que un sistema de transposición in vitro apropiado está a disposición o que se puede desarrollar en base al conocimiento en la técnica. En general, un sistema de transposición in vitro apropiado para usar en los métodos de la presente invención requiere, como mínimo, una enzima transposasa de suficiente pureza, suficiente concentración y suficiente actividad de transposición in vitro y un extremo del transposón con el que la transposasa forma un complejo funcional con la respectiva transposasa que es capaz de catalizar la reacción de transposición. Las secuencias del extremo de transposón de transposasa apropiadas que se pueden usar en la invención incluyen, pero sin limitación, secuencias del extremo del transposón de tipo salvaje, derivado o mutante que forman un complejo con una transposasa seleccionada de entre una forma de tipo salvaje, derivada o mutante de la transposasa. Las transposasas de ejemplo que fueron usadas exitosamente por los solicitantes en los métodos de la presente invención incluyen formas de tipo salvaje o mutante de Tn5 transposasa y MuA transposasa (a pesar de que la EZ-Tn5 transposasa era significativamente más eficaz que una cantidad de proteína equivalente de MuA transposasa en la generación de fragmentos de ADN rotulados en 5' en los métodos de la presente invención), pero se puede usar cualquier otra transposasa para las que las composiciones y las condiciones para una transposición in vitro eficaz de los extremos del transposón definidos se conocen o posteriormente se desarrollan en los presentes métodos. Las secuencias de extremo del transposón reconocidas por las formas de tipo salvaje o mutantes de Tn5 transposasa o MuA transposasa son preferidas y aquellas secuencias de extremo del transposón que resultan con las máximas eficacias de transposición cuando se complejizan con la transposasa, junto con las correspondientes enzimas transposasas ópticamente activas que se complejizan con ellas, son las más preferidas para las formas de realización de la presente invención. Con preferencia, se selecciona un transposón en donde la secuencia del extremo de transposasa requerida por la transposasa para la transposición no es demasiado grande y las secuencias del extremo del transposón son del mínimo tamaño posible que funcionan bien para la finalidad pretendida y que son de suficiente tamaño, de modo que la misma secuencia está presente sólo raramente o, con preferencia, si no está presente en absoluto, en el ADN objetivo o muestra de ADN. A modo de ejemplo, las secuencias del extremo del transposón de las secuencias de extremo del transposón EZ-Tn5<sup>TM</sup> derivado de Tn5 comprenden sólo 19 nucleótidos, mientras que algunas otras transposasas requieren de secuencias terminales mucho más grandes para transposición (por ejemplo, MuA transposasa requería secuencias de extremo del transposón de aproximadamente 51 nucleótidos).

Los sistemas apropiados de transposición in vitro que se pueden usar para insertar un extremo de transposón en un ácido nucleico objetivo incluyen, pero sin limitación, aquellas que utilizan la Tn5 Transposasa EZTn5<sup>TM</sup> hiperactiva asequible de EPICENTRE Technologies, Madison, WI o la MuA Transposasa HyperMu<sup>TM</sup> hiperactiva de EPICENTRE u otra MuA transposasa, como la asequible de Finnzymes Oy, Espoo, Finlandia. Los oligonucleótidos de extremo del transposón que exhiben las secuencias de los respectivos extremos del transposón se pueden sintetizar usando un sintetizador de oligonucleótido o se pueden comprar de una fuente comercial en base a la información disponible de los respectivos vendedores o usando información bien conocida en la técnica. Por ejemplo, las secuencias de nucleótidos del extremo mosaico de transposón hiperactivo para EZ-Tn5<sup>TM</sup> transposasa se presentan en el Ejemplo 1 y está a disposición información adicional relacionada con la EZ-Tn5<sup>TM</sup> transposasa en la literatura publicada y en línea en [www.EpiBio.com](http://www.EpiBio.com) de EPICENTRE Biotechnologies, Madison, WI, Estados Unidos.

En algunas formas de realización, la inserción de un extremo del transposón en ADN objetivo de acuerdo con la presente invención también se puede llevar a cabo in vivo. Si la transposición se lleva a cabo in vivo, la transposición en el ADN objetivo se logra preferentemente por electroporación de un complejo sináptico de una transposasa y una composición de extremo del transposón apropiada en la célula huésped tal como se describe en la patente U.S. No. 6.159.736 Este método de transposición se ejemplifica empleando un complejo de transposición formado por una Tn5 transposasa hiperactiva y una composición de extremo del transposón Tn5 apropiada usando métodos similares a los descritos por Goryshin, I. y Reznikoff, W. S. (J. Biol. Chem., 273: 7367, 1998) o un complejo de transposición formado por MuA Transposasa HyperMu<sup>TM</sup> hiperactiva (EPICENTRE, Madison, WI) y una composición de extremo del transposón MuA apropiada que exhibe las secuencias terminales R1 y R2 reconocidas por la transposasa. Los complejos sinápticos apropiados o complejos Transposome<sup>TM</sup> (EPICENTRE) entre una composición de extremo del transposón y una transposasa se pueden preparar tal como se describe en la patente U.S. No. 6.159.736 y patentes relacionadas de Goryshin y Reznikoff o tal como se describe en la literatura de productos para complejos EZ-Tn5<sup>TM</sup> Transposome<sup>TM</sup> de tipo Tn5 o para complejos HyperMu<sup>TM</sup> MuA Transposome<sup>TM</sup> de EPICENTRE Technologies, Madison, WI, excepto porque se usan oligonucleótidos que exhiben sólo un extremo del transposón en vez de un polinucleótido u oligonucleótido que tiene dos extremos del transposón, usualmente en o cerca de cada extremo de los respectivos polinucleótido u oligonucleótido.

La invención también se refiere a o comprende kits y composiciones individuales para cada uno de los métodos de la invención. Un kit es una combinación de composiciones individuales de utilidad para llevar a cabo un método de la invención, en donde las composiciones se optimizan para usar juntas en el método. Una composición comprende un componente individual o una mezcla de componentes para al menos una etapa de un método de la invención. La invención se refiere a cualquier kit que se puede ensamblar a partir de una combinación de dos composiciones cualesquiera de la invención y cualquier composición nueva que se usa en un kit o método de la invención. De modo alternativo, se puede ensamblar un kit a partir de un componente o composición simple en un formato de uso conveniente, por ejemplo, prealiquotado en una porción de uso individual y puede incluir opcionalmente un grupo de instrucciones para uso del componente o composición.

Descripción de la invención

## Introducción

La presente invención se refiere a métodos y composiciones para tratar ácido nucleico y en particular, métodos y composiciones para fragmentar y rotular ADN usando composiciones de transposón. Los métodos y composiciones de la presente invención son de utilidad para generar bibliotecas de fragmentos de ssADN lineales dirrotulados o fragmentos de ssADN circulares rotulados (y sus productos de amplificación) de ADN objetivo que comprende cualquier dsADN de interés (incluyendo cADN bicatenario preparado de ARN) de cualquier fuente para análisis genómico, subgenómico, transcriptómico o metagenómico o análisis de expresión de ARN (por ejemplo, para usar en la preparación de un objetivo etiquetado para análisis de micromatriz; por ejemplo, para análisis de la variación del número de copias, para la detección y el análisis de polimorfismos de nucleótidos simples y para hallar genes de muestras ambientales tales como fuentes de suelo o agua). Los métodos son de utilidad en una variedad de procesos, incluyendo, pero sin limitación, procesos para la amplificación del genoma completo de uno o varios organismos, incluyendo uno o varios organismos microbianos o ambientales para los cuales las condiciones de cultivo o crecimiento son desconocidas (por ejemplo, amplificación de genoma completo o WGA), PCR en tiempo real, PCR en emulsión, hibridación genómica comparativa (CGH), secuenciación genómica comparativa (CGS) y para preparar sondas específicas de ADN (por ejemplo, sondas específicas del cromosoma, por ejemplo, pinturas de cromosoma o, por ejemplo, sondas específicas de genes o locus) para aplicaciones tales como hibridación fluorescente in situ (FISH). En algunas formas de realización, los métodos también se usan para generar plantillas para secuenciación de ADN masivamente paralela (denomina "secuenciación de próxima generación"). Cada uno de estos procesos o aplicaciones se usan tanto para fines investigativos como de diagnóstico molecular.

La presente invención proporciona métodos y composiciones para generar una biblioteca de fragmentos de ADN rotulados de ADN objetivo que comprenden ADN bicatenario (dsADN) contenido en cualquier muestra de interés. Los métodos son más fáciles, rápidos, requieren menos manos a la vez, se pueden realizar con muestras más pequeñas y menores cantidades de ácidos nucleicos de muestra, son más eficientes en la rotulación de los extremos de los fragmentos y generan fragmentos de ADN dirrotulados que con cualitativa y/o cuantitativamente representativos de los ácidos nucleicos de muestra de los que se generan. Los métodos se pueden llevar a cabo con facilidad a mano sin un instrumento, pero también se adaptan con facilidad para la automatización robótica en un ambiente de alto rendimiento.

## Formas de realización de los métodos

Todas las formas de realización de los métodos de la presente invención divulgados en la presente utilizan una reacción de transposición in vitro para romper simultáneamente un ADN objetivo en fragmentos y unir un rótulo con el extremo 5' de cada fragmento. Como todos los métodos están relacionados, a menos que se establezca específicamente otra cosa con respecto a una forma de realización particular, también se puede usar un método que está presente en la presente con respecto a una forma de realización con otra forma de realización descrita en la presente. Todas las formas de realización de los métodos divulgados en la presente que usan una reacción de transposición in vitro se pueden realizar efectuando la reacción con transposasa separada y composiciones de extremo del transposón o una composición de transposoma individual que comprende un complejo estable formado entre la transposasa y la composición de extremo del transposón. En consecuencia, se entenderá que cualquier método que describa el uso de una transposasa y una composición de extremo del transposón también podría usar una composición de transposoma hecha a partir de una transposasa y la composición de extremo del transposón y cualquier método que describa el uso de una composición de transposoma también podría usar la transposasa separada y una composición de extremo del transposón de la cual se compone la composición de transposoma. Esto se ilustra por medio de las dos siguientes descripciones de un método general de la invención.

Una forma de realización de la invención es un método para generar una biblioteca de fragmentos de ADN rotulados de ADN objetivo que comprende cualquier dsADN de interés (por ejemplo, para usar como secuenciación de próxima generación o plantillas de amplificación), que comprende: incubación del ADN objetivo en una reacción de transposición in vitro con al menos una transposasa y una composición de extremo del transposón con la que la transposasa forma un complejo de transposición, la composición de extremo del transposón que comprende (i) una cadena transferida que exhibe una secuencia del extremo del transposón transferida y, opcionalmente, una secuencia adicional 5' de la secuencia del extremo del transposón transferida y (ii) una cadena no transferida que exhibe una secuencia que es complementaria a la secuencia del extremo del transposón transferida, en condiciones y durante un tiempo suficiente en donde se producen múltiples inserciones en el ADN objetivo, cada una de las cuales da como resultado la unión de un primer rótulo que comprende o que consiste en la cadena transferida con el extremo 5' de un nucleótido en el ADN objetivo, fragmentando así el ADN objetivo y generando una población de fragmentos de ADN fusionados con rótulo en 5', cada uno de los cuales tiene el primer rótulo en el extremo 5'; y luego unión de los extremos 3' de los fragmentos de ADN con rótulo en 5' con el primer rótulo o con un segundo rótulo, generando así una biblioteca de fragmentos de ADN rotulados (por ejemplo, que comprende ya sea fragmentos de ssADN circulares rotulados o fragmentos de ADN rotulados en 5' y 3' (o fragmentos de ADN dirrotulados)).

En una forma de realización preferida, tal como se describió inmediatamente antes, el método se lleva a cabo usando transposasa separada y composiciones de extremo del transposón, mientras que, en algunas otras formas de realización preferidas, el método se lleva a cabo usando una composición de transposoma que comprende el complejo formado entre la transposasa y la composición de extremo del transposón.

Así, una forma de realización preferida de la invención es un método para generar una biblioteca de fragmentos de ADN rotulados de ADN objetivo en una reacción de transposición *in vitro* que comprende cualquier dsADN de interés (por ejemplo, para usar como secuenciación de próxima generación o plantillas de amplificación), que comprende: incubación del ADN objetivo con una o varias composiciones de transposoma, cada una de las cuales comprende un complejo entre una transposasa y una composición de extremo del transposón con la que la transposasa forma un complejo de transposición, donde la composición de extremo del transposón comprende (i) una cadena transferida que exhibe una secuencia del extremo del transposón transferida y, opcionalmente, una secuencia adicional 5' de la secuencia del extremo del transposón transferida y (ii) una cadena no transferida que exhibe una secuencia que es complementaria a la secuencia del extremo del transposón transferida, en condiciones y durante un tiempo suficiente donde se producen múltiples inserciones en el ADN objetivo, cada una de las cuales da como resultado la unión de un primer rótulo que comprende o que consiste en la cadena transferida con el extremo 5' de un nucleótido en el ADN objetivo, fragmentando así el ADN objetivo y generando una población de fragmentos de ADN fusionados con rótulo en 5', cada uno de los cuales tiene el primer rótulo en el extremo 5'; y luego uniendo los extremos 3' de los fragmentos de ADN con rótulo en 5' con el primer rótulo o con un segundo rótulo, generando así una biblioteca de fragmentos de ADN rotulados (por ejemplo, que comprenden ya sea fragmentos de ssADN circulares rotulados o fragmentos de ADN rotulados en 5' y 3' (o "fragmentos de ADN dirrotulados")).

En algunas formas de realización de cualquiera de los métodos de la invención, la cantidad de la transposasa y la composición de extremo del transposón o de la composición de transposoma usada en la reacción de transposición *in vitro* está entre aproximadamente 1 picomol y aproximadamente 25 picomoles por 50 nanogramos de ADN objetivo por reacción de 50 microlitros. En algunas formas de realización de preferencia de cualquiera de los métodos de la invención, la cantidad de la transposasa y la composición de extremo del transposón o de la composición de transposoma usada en la reacción de transposición *in vitro* es de entre aproximadamente 5 picomoles y aproximadamente 50 picomoles por 50 nanogramos de ADN objetivo por 50 microlitros de reacción. En algunas formas de realización de preferencia de cualquiera de los métodos de la invención, en donde la transposasa es la Tn5 transposasa hiperactiva y la composición de extremo del transposón comprende dicha Tn5 transposasa hiperactiva y una composición de extremo del transposón que comprende el extremo del transposón MEDS, las cantidades de dicha transposasa y la composición de extremo del transposón o dicha composición de transposoma usada en la reacción de transposición *in vitro* es de entre aproximadamente 5 picomoles y aproximadamente 25 picomoles por 50 nanogramos de ADN objetivo por 50 microlitros de reacción. En algunas formas de realización de preferencia de cualquiera de los métodos de la invención, en donde la transposasa es una Tn5 transposasa hiperactiva o MuA transposasa, las concentraciones finales de la transposasa y la composición de extremo del transposón o de la composición de transposoma usada en la reacción de transposición *in vitro* es de al menos 250 nM; en algunas otras formas de realización, las concentraciones finales de Tn5 transposasa hiperactiva o MuA transposasa y de su respectiva composición de extremo del transposón o composición de transposoma es de al menos 500 nM.

En algunas formas de realización de cualquiera de los métodos de la invención, el tiempo de reacción para la reacción de transposición *in vitro* es de dos horas o menos, una hora o menos, 30 minutos o menos o 15 minutos o menos. En algunas formas de realización de preferencia de cualquiera de los métodos de la invención, el tiempo de reacción para la reacción de transposición *in vitro* es de 5 minutos o menos. En algunas formas de realización de preferencia de cualquiera de los métodos de la invención, en donde la composición de transposoma comprende la Tn5 transposasa hiperactiva y una composición de extremo del transposón que comprende el extremo del transposón MEDS, el tiempo de reacción para la reacción de transposición *in vitro* es de 5 minutos o menos.

En algunas formas de realización, el método también comprende la etapa de amplificar no selectivamente los fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados o fragmentos de ssADN circulares rotulados usando una ADN polimerasa termoestable y al menos un cebador que es complementario del primer rótulo o el segundo rótulo. En algunas formas de realización de preferencia del método donde sólo un transposoma se usa en la reacción de transposición *in vitro*, la etapa de amplificación de los fragmentos de ADN rotulados comprende la amplificación de los fragmentos de ADN dirrotulados o los fragmentos de ssADN circular rotulados usando un cebador simple que exhibe la secuencia de al menos una porción de la cadena transferida. En algunas formas de realización, la etapa de amplificación de los fragmentos de ADN rotulados usando un cebador simple comprende una reacción de PCR o reacción de replicación del círculo de rodadura. En algunas formas de realización, la porción 5' de un cebador usado para la amplificación comprende o consiste en un dominio de rótulo de secuenciación.

En una forma de realización preferida de cualquiera de los métodos de la invención, la biblioteca de fragmentos de ADN se usa para proporcionar plantillas para secuenciación de ADN o amplificación de ácidos nucleicos.

La invención comprende varias formas de realización para generar una biblioteca de fragmentos de ADN rotulados que comprende ya sea fragmentos de ADN dirrotulados o fragmentos de ssADN circulares rotulados, tal como se trata más abajo.

Uso de una ADN polimerasa con actividad de desplazamiento de cadena o de 5' nucleasa que genera fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados.

Una forma de realización preferida del método comprende: incubación del ADN objetivo en la reacción de transposición

in vitro con al menos un transposoma en condiciones y durante un tiempo suficiente para generar una población de fragmentos de ADN fusionados con rótulo en 5'; y luego incubación de la población de fragmentos de ADN fusionados con rótulo en 5' con una ADN polimerasa que tiene actividad de desplazamiento de cadena o de 5' nucleasa en condiciones sin termociclado y en donde los fragmentos de ADN fusionados con rótulo en 5' no se desnaturalizan, en donde la ADN polimerasa extiende el extremo 3' de cada cadena de los fragmentos de ADN fusionados con rótulo en 5' usando la cadena complementaria como una plantilla y desplaza o digiere la cadena no transferida, generando así la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de dsADN dirrotulados.

Una forma de realización preferida del método comprende: incubación del ADN objetivo en la reacción de transposición in vitro con al menos un transposoma para generar una población de fragmentos de ADN fusionados con rótulo en 5'; incubación de la población de fragmentos de ADN fusionados con rótulo en 5' con la ADN polimerasa que tiene actividad de desplazamiento de cadena o de 5' nucleasa para generar fragmentos de dsADN dirrotulados; y desnaturalización de los fragmentos de dsADN dirrotulados para generar la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ssADN dirrotulados (por ejemplo, por calentamiento hasta 95°C y enfriamiento rápido). En una versión preferida de esta forma de realización del método, la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ssADN dirrotulados se genera a partir del ADN objetivo en un tubo simple sin realizar ninguna etapa de purificación interviniente.

En algunas formas de realización del método que comprende la generación de una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados usando una ADN polimerasa que tiene actividad de desplazamiento de cadena o de 5' nucleasa, el método también comprende la etapa de amplificación de los fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados usando una ADN polimerasa termoestable y al menos un cebador que es complementario al segundo rótulo. En algunas formas de realización de preferencia de este método, la etapa de amplificación de la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados comprende la amplificación de la biblioteca de fragmentos de ADN rotulados por PCR usando sólo un oligodesoxirribonucleótido que exhibe la secuencia de al menos una porción de la cadena transferida como un cebador de PCR y los fragmentos de ADN dirrotulados como plantillas. Así, esta forma de realización es un método para la amplificación por PCR de un cebador de una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados generados a partir del ADN objetivo. Si el ADN objetivo comprende ADN genómico total de un organismo, esta forma de realización es un método para la amplificación de genoma completo no selectiva.

En algunas formas de realización de preferencia, en donde una composición simple de extremo del transposón se usa en la reacción de transposición in vitro del método que comprende la generación de la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados usando una ADN polimerasa que tiene actividad de desplazamiento de cadena o de 5' nucleasa y luego amplificación de los fragmentos de ADN dirrotulados generados por PCR, se usan dos diferentes cebadores de PCR, cada uno de cuyos cebadores de PCR exhibe la secuencia de al menos una porción del extremo del transposón transferido que comprende la composición de extremo del transposón. En algunas formas de realización preferidas, cada cebador de PCR comprende una porción 3' y una porción 5', en donde la porción 3' exhibe la respectiva secuencia del extremo del transposón transferida y la porción 5' exhibe la secuencia de un respectivo dominio de rótulo para un fin particular (por ejemplo, un dominio de rótulo de secuenciación o un dominio de rótulo de amplificación y opcionalmente un dominio de rótulo de direccionamiento para la secuenciación o amplificación de próxima generación).

En algunas formas de realización de preferencia de cualquiera de los métodos que comprende la generación de la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados usando una ADN polimerasa que tiene actividad de desplazamiento de cadena o de 5' nucleasa, el al menos un transposoma en la reacción de transposición in vitro comprende o consiste en dos diferentes transposomas. En algunas formas de realización de preferencia, en donde se usan dos diferentes transposomas, cada uno de los dos transposomas comprende la misma transposasa pero una distinta composición de extremo del transposón. En algunas formas de realización de preferencia, en donde se usan dos diferentes transposomas, los dos diferentes transposomas comprenden cada uno la misma transposasa y las composiciones de extremo del transposón comprenden distintas cadenas transferidas. En algunas formas de realización de preferencia, en donde se usan dos diferentes transposomas, cada uno de los dos transposomas comprende distintas enzimas transposasa y distintas composiciones de extremo del transposón, cada uno de los cuales forma un complejo funcional con la respectiva transposasa. En algunas formas de realización de preferencia del método, en donde se usan dos diferentes composiciones de extremo del transposón en la reacción de transposición in vitro y en donde la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ssADN dirrotulados se genera usando una ADN polimerasa que tiene actividad de desplazamiento de cadena o de 5' nucleasa, el primer rótulo exhibe la secuencia de la cadena transferida de una composición de extremo del transposón y el segundo rótulo exhibe la secuencia de la cadena no transferida de la otra composición de extremo del transposón.

En algunas formas de realización de preferencia del método que comprende la generación de la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados usando una ADN polimerasa que tiene actividad de desplazamiento de cadena o de 5' nucleasa, en donde se usan dos diferentes composiciones de extremo del transposón en la reacción de transposición in vitro y el método también comprende la etapa de amplificación de los fragmentos de ADN dirrotulados generados por PCR, se usan dos diferentes cebadores de PCR, uno de cuyos cebadores de PCR exhibe la secuencia de al menos una porción de la cadena transferida que compone una

- composición de extremo del transposón y el otro de los cebadores de PCR exhibe la secuencia de al menos una porción de la cadena transferida que compone la otra composición de extremo del transposón. En algunas formas de realización de preferencia, en donde la cadena transferida que compone cada respectiva composición de extremo del transposón comprende una porción 3' y una porción 5', en donde la porción 3' exhibe la respectiva secuencia del extremo del transposón transferida y la porción 5' de cada cadena transferida respectiva exhibe una diferente secuencia que comprende un dominio de rótulo para un fin particular (por ejemplo, un dominio de rótulo de secuenciación o un dominio de rótulo de amplificación y opcionalmente un dominio de rótulo de direccionamiento para secuenciación o amplificación de próxima generación), donde cada cebador de PCR exhibe la secuencia del dominio de rótulo del respectivo oligonucleótido de transposón transferido.
- 5
- 10 Uso de transferasa terminal para generar fragmentos de ADN rotulados que comprenden fragmentos de ADN dirrotulados.
- Otra forma de realización del método comprende: incubación del ADN objetivo en la reacción de transposición in vitro con el al menos un transposoma para generar los fragmentos de dsADN con rótulo en 5'; desnaturalización de los fragmentos de dsADN con rótulo en 5' para generar fragmentos de ssADN con rótulo en 5'; e incubación de los fragmentos de ssADN con rótulo en 5' con una ADN polimerasa que consiste en una transferasa terminal y al menos un sustrato de dNTP para la transferasa terminal en condiciones y durante un tiempo suficiente en donde la transferasa terminal une el segundo rótulo que consiste en poli(dNMP) con el extremo 3' de los fragmentos de ADN rotulados en 5', generando así una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados (por ejemplo, la FIG. 3). En algunas formas de realización de este método, el extremo 3' del oligonucleótido del extremo del transposón no transferido que compone la composición de extremo del transposón del transposoma se bloquea (por ejemplo, usando un oligonucleótido del extremo del transposón no transferido que tiene un didesoxi nucleótido o un 3'-O-metil-nucleótido como el 3'-nucleótido terminal), que, cuando está bloqueado, el 3' nucleótido evita la adición de poli(dNMP) por la transferasa terminal, evitando así la rotulación de fondo del oligonucleótido del extremo del transposón no transferido.
- 15
- 20
- 25 Aún otra forma de realización del método comprende: incubación del ADN objetivo en la reacción de transposición in vitro con el al menos un transposoma para generar los fragmentos de ADN con rótulo en 5'; incubación de los fragmentos de ADN con rótulo en 5', sin una etapa previa de desnaturalización, con una ADN polimerasa que consiste en una transferasa terminal y al menos un sustrato dNTP para la transferasa terminal en condiciones y durante un tiempo suficiente en donde la transferasa terminal une el segundo rótulo que consiste en el poli(dNMP) con el extremo 3' de los fragmentos de ADN con rótulo en 5', generando así una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados. En algunas formas de realización de este método, el extremo 3' del oligonucleótido del extremo del transposón no transferido que compone la composición de extremo del transposón del transposoma está bloqueado (por ejemplo, usando un oligonucleótido del extremo del transposón no transferido que tiene un didesoxi nucleótido o un 3'-O-metil-nucleótido como el nucleótido 3'-terminal).
- 30
- 35 Uso de una ADN polimerasa y un oligonucleótido de rótulo terminal para generar fragmentos de ADN rotulados que comprenden fragmentos de ADN dirrotulados.
- Aún otra forma de realización del método comprende: incubación del ADN objetivo en la reacción de transposición in vitro con el al menos un transposoma para generar los fragmentos de dsADN con rótulo en 5'; desnaturalización de los fragmentos de dsADN con rótulo en 5' para generar fragmentos de ssADN con rótulo en 5' (por ejemplo, calentando hasta 95°C y enfriamiento rápidamente); y unión del segundo rótulo con los fragmentos de ssADN con rótulo en 5' usando una ADN polimerasa y un oligonucleótido de rótulo terminal (por ejemplo, FIG. 4), generando así una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados. En algunas formas de realización preferidas, la etapa de unión del segundo rótulo con el extremo 3' de los fragmentos de ADN rotulados en 5' usando una ADN polimerasa y un oligonucleótido de rótulo terminal comprende:
- 40
- 45 (1) proporcionar un oligonucleótido de rótulo terminal que comprende o que consiste en una porción 5' y porción 3', en donde la porción 5' exhibe una secuencia que es complementaria a la secuencia del segundo rótulo que se desea unir con el término 3' de los fragmentos de ssADN con rótulo en 5' y la porción 3' exhibe una secuencia aleatoria que comprende o que consiste en tres a ocho (por ejemplo, 3, 4, 5, 6, 7 u 8 nucleótidos aleatorios, de los cuales el nucleótido 3'-terminal está bloqueado, de modo que no es capaz de ser extendido por la ADN polimerasa;
- 50 (2) poner en contacto los fragmentos de ssADN con rótulo en 5' con el oligonucleótido de rótulo terminal en condiciones y durante un tiempo suficiente en donde el oligonucleótido de rótulo terminal se fusiona con los fragmentos de ssADN con rótulo en 5'; y
- (3) poner en contacto los fragmentos de ssADN con rótulo en 5' a los que el oligonucleótido de rótulo terminal se fusiona con la ADN polimerasa en una mezcla de reacción y en condiciones de polimerización de ADN y durante un tiempo suficiente en donde el término 3' de los fragmentos de ssADN con rótulo en 5' se extienden usando el oligonucleótido de rótulo terminal como una plantilla, con lo que el segundo rótulo se une con sus términos 3' y se generan fragmentos de ssADN con rótulo en 5' y 3'. En algunas formas de realización, se usa una secuencia semialeatoria en lugar de la secuencia aleatoria en el oligonucleótido de rótulo terminal.
- 55

En algunas variantes de esta forma de realización, el oligonucleótido de rótulo terminal comprende o consiste en desoxirribonucleótidos. En algunas variantes de esta forma de realización, el oligonucleótido de rótulo terminal comprende o consiste en ribonucleótidos, en cuyas formas de realización la ADN polimerasa es una ADN polimerasa dependiente del ARN. En algunas formas de realización de preferencia, la porción 3' del oligonucleótido de rótulo terminal consiste en siete nucleótidos aleatorios. En algunas formas de realización de preferencia del método en donde un oligonucleótido de rótulo terminal se usa para unir el segundo rótulo con los fragmentos de ssADN con rótulo en 5', el segundo rótulo no es complementario al primer rótulo.

Uso de una ligasa dependiente de la plantilla (u homóloga) y un oligonucleótido de rotulación de ligamiento para generar fragmentos de ADN rotulados que comprenden fragmentos de ADN dirrotulados.

Una forma de realización preferida del método comprende: incubación del ADN objetivo en la reacción de transposición in vitro con el al menos un transposoma en condiciones y durante un tiempo suficiente para generar una población de fragmentos de ADN con rótulo en 5' fusionados; y luego incubación de la población de fragmentos de dsADN con rótulo en 5' fusionados con una ADN ligasa dependiente de la plantilla (u homóloga) y un oligodeoxinucleótido con rótulo de ligamiento que comprende o que consiste en una porción 3' y una porción 5', en donde la porción 3' exhibe un segundo rótulo que exhibe cualquier secuencia que se desea unir con el extremo 3' de la población de fragmentos de ADN fusionados con rótulo en 5' (por ejemplo, una secuencia arbitraria) y la porción 5' tiene un grupo 5'-monofosfato y exhibe una secuencia aleatoria, en condiciones y durante un tiempo suficiente, en donde el segundo rótulo se une con los fragmentos de ADN fusionados con rótulo en 5', generando así una biblioteca de fragmentos de ADN que comprende fragmentos de ADN dirrotulados fusionados. En algunas formas de realización de preferencia, el método también comprende la etapa de desnaturalización de la biblioteca de fragmentos de ADN que comprende fragmentos de ADN dirrotulados fusionados (por ejemplo, por calentamiento hasta 95°C y rápido enfriamiento), generando así una biblioteca de fragmentos de ADN que comprende fragmentos de ssADN dirrotulados.

En algunas formas de realización de preferencia, el oligonucleótido de rotulación de ligamiento comprende una porción 5' que exhibe una secuencia aleatoria que consiste en aproximadamente tres a aproximadamente ocho nucleótidos.

En algunas formas de realización de preferencia, el oligonucleótido de rotulación de ligamiento comprende una porción 5' que exhibe una secuencia aleatoria que consiste en cuatro nucleótidos. En algunas formas de realización de preferencia, la ligasa dependiente de la plantilla es ADN ligasa de E. coil. En una versión preferida de esta forma de realización del método, la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ssADN dirrotulados es generada a partir del ADN objetivo en un tubo simple sin llevar a cabo ninguna etapa de purificación interviniente.

Uso de una composición de extremo de transposón en horquilla y una ligasa dependiente de la plantilla para generar una biblioteca de fragmentos de ADN rotulados que comprenden fragmentos de ADN circulares rotulados, fragmentos de dsADN de cola de milano o fragmentos de ADN dirrotulados.

En una forma de realización preferida del método para generar una biblioteca de fragmentos de ADN rotulados de ADN objetivo en una reacción de transposición in vitro, el método comprende: incubación del ADN objetivo con una o varias composiciones de transposoma, donde cada una comprende un complejo entre una transposasa y una composición de extremo del transposón en horquilla con la que la transposasa forma un complejo de transposición, donde la composición de extremo del transposón en horquilla comprende o consiste en un oligonucleótido con contenido de 5'-fosfato que exhibe una secuencia del extremo del transposón no transferida en su extremo 5', una secuencia del extremo del transposón transferida en su extremo 3' y una secuencia de rótulo arbitraria interviniente entre la secuencia del extremo del transposón no transferida y la secuencia del extremo del transposón transferida que es suficientemente larga para permitir la formación de tallo-bucle intramolecular; en condiciones y durante un tiempo suficiente en donde inserción de la composición de extremo del transposón en horquilla en el ADN objetivo genera una población de fragmentos de ADN fusionados con rótulo en 5'; luego incubación de la población de fragmentos de ADN fusionados con rótulo en 5' con una o varias secuencias aleatorias de oligonucleótidos con contenido de 5'-fosfato que, solos o en combinación, tienen la misma longitud que las brechas monocatenarias en los fragmentos de ADN fusionados con rótulo en 5' que dan como resultado a continuación la reacción de transposición in vitro, en condiciones y durante un tiempo suficiente en donde las brechas monocatenarias en la población de fragmentos de ADN fusionados con rótulo en 5' se rellenan por fusión de los oligonucleótidos de secuencia aleatoria con el ADN objetivo en las brechas monocatenarias; y luego, incubación de la población de fragmentos de ADN fusionados con rótulo en 5' con brechas monocatenarias rellenas con una ligasa dependiente de la plantilla en condiciones y durante un tiempo suficiente en donde los oligonucleótidos fusionados de secuencia aleatoria se ligan entre sí o con los extremos 5' de fragmentos de ADN con rótulo en 5' adyacentes, generando así la biblioteca de fragmentos de ADN circulares rotulados.

En otra forma de realización preferida del método para generar una biblioteca de fragmentos de ADN rotulados de ADN objetivo en una reacción de transposición in vitro, el método comprende: incubación del ADN objetivo con una o varias composiciones de transposoma, donde cada una comprende un complejo entre una transposasa y una composición de extremo del transposón en horquilla con la que la transposasa forma un complejo de transposición, donde la composición de extremo del transposón en horquilla comprende o consiste en un oligonucleótido con contenido de 5'-fosfato que exhibe una secuencia del extremo del transposón no transferida en su extremo 5', una secuencia del extremo del transposón transferida en su extremo 3' y una secuencia de rótulo arbitraria interviniente

entre la secuencia del extremo del transposón no transferida y la secuencia del extremo del transposón transferida que es suficientemente larga para permitir la formación de tallo-bucle intramolecular; en condiciones y durante un tiempo suficiente en donde inserción de la composición de extremo del transposón en horquilla en el ADN objetivo genera una población de fragmentos de ADN fusionados con rótulo en 5'; luego incubación de la población de fragmentos de ADN fusionados con rótulo en 5' con una ADN polimerasa que carece de actividades de 5' a 3' exonucleasa y 5' nucleasa dependiente de la estructura y desplazamiento de cadena, en condiciones y durante un tiempo suficiente en donde las brechas monocatenarias que están presentes en la población de fragmentos de ADN fusionados con rótulo en 5' después de la reacción de transposición in vitro se rellenan por extensión de los extremos 3' de cada fragmento de ADN fusionado con rótulo en 5' por la ADN polimerasa; y luego, incubación de la población de fragmentos de ADN fusionados con rótulo en 5' con brechas monocatenarias rellenas en con una ligasa dependiente de la plantilla en condiciones y durante un tiempo suficiente en donde los extremos 3' de los productos de extensión de ADN polimerasa fusionada se ligan con los extremos 5' de fragmentos de ADN adyacentemente fusionados con rótulo en 5', generando así la biblioteca de fragmentos de ADN circulares rotulados.

En algunas formas de realización preferidas de este método, tanto la ADN polimerasa como la ligasa dependiente de la plantilla se proporcionan en una mezcla de reacción simple y tanto la extensión de la ADN polimerasa como el ligamiento dependiente de la plantilla se llevan a cabo en la mezcla de reacción simple.

En algunas formas de realización de cualquiera de estos métodos para generar una biblioteca de fragmentos de ADN circulares rotulados, el método comprende además, después de la etapa de incubación con la ligasa dependiente de la plantilla para generar la biblioteca de fragmentos de ADN circulares rotulados, una o varias etapas para eliminar ssADN y dsADN lineal no ligado (por ejemplo, que comprende los oligonucleótidos de secuencia aleatoria, el ADN objetivo lineal y/o las composiciones de extremo del transposón en horquilla que no se unen con el ADN objetivo). En una forma de realización preferida para remover el ssADN y dsADN lineal no ligado, el método comprende además: tratamiento de la mezcla de reacción que contiene los fragmentos de ADN circulares rotulados con T5 exonucleasa.

En algunas formas de realización de preferencia de cualquiera de estos métodos para generar una biblioteca de fragmentos de ADN circulares rotulados, el método comprende además: clivaje de los fragmentos de ADN circulares rotulados en cada una de las estructuras de bucle derivadas de las composiciones de extremo del transposón en horquilla para generar fragmentos de dsADN de cola de milano, donde cada cadena tiene una porción del rótulo en su extremo 5' y una porción del rótulo en su extremo 3'. En algunas formas de realización, la etapa de clivaje de los fragmentos de ADN circulares rotulados en cada una de las estructuras de bucle comprende: poner en contacto los fragmentos de ADN circulares rotulados con una composición de enzima de clivaje en condiciones y durante un tiempo suficiente en donde los fragmentos de ADN circulares rotulados se clavan en los sitios de clivaje para generar los fragmentos de dsADN de cola de milano. En algunas formas de realización, la etapa de clivaje de los fragmentos de ADN circulares rotulados en cada una de las estructuras de bucle comprende: fusión con los fragmentos de ADN circulares rotulados un oligodesoxirribonucleótido que se fusiona con un sitio de restricción dentro del rótulo y luego incubación con la endonucleasa de restricción que cliva en el sitio de restricción bicatenario en condiciones y durante un tiempo suficiente para generar la biblioteca de los fragmentos de dsADN de cola de milano. En algunas formas de realización de preferencia, la etapa de clivaje de los fragmentos de ADN circulares rotulados en cada una de las estructuras de bucle comprende: poner en contacto los fragmentos de ADN circulares rotulados con una ADN glicosilasa y una AP endonucleasa, en donde la ADN glicosilasa remueve la base de ácido nucleico de un nucleótido no canónico (por ejemplo, un dUMP o 8-oxo-dGMP) que está presente dentro del rótulo y la AP endonucleasa cliva los fragmentos de ssADN circulares rotulados en el sitio abásico resultante; en alguna forma de realización, la ADN glicosilasa se selecciona de entre uracil-N-glicosilasa y proteína FPG y la AP endonucleasa se selecciona de entre endonucleasa III o endonucleasa IV de *E. coli*.

En algunas formas de realización de preferencia de cualquiera de los métodos para generar una biblioteca de fragmentos de dsADN de cola de milano, el método comprende además la etapa de: desnaturalización de la biblioteca de fragmentos de dsADN de cola de milano para generar una biblioteca de fragmentos de ssADN lineales dirrotulados.

Uso de una ligasa independiente de la plantilla para generar fragmentos de ADN rotulados que comprenden fragmentos de ssADN circulares rotulados o fragmentos de ssADN lineales dirrotulados.

Una forma de realización preferida del método comprende: incubación del ADN objetivo en la reacción de transposición in vitro con el al menos un transposoma, en donde el extremo 5' de la cadena transferida que comprende el transposoma tiene un grupo 5'-fosfato, en condiciones y durante un tiempo suficiente para generar una población de fragmentos de ADN fusionados con rótulo en 5'; luego desnaturalizar los fragmentos de dsADN fusionados con rótulo en 5' para obtener fragmentos de ssADN con rótulo en 5' (por ejemplo, por calentamiento hasta 95°C y rápido enfriamiento); y luego incubar los fragmentos de ssADN con rótulo en 5' en una reacción de ligamiento con una ligasa independiente de la plantilla (o no homóloga) en condiciones y durante un tiempo suficiente en donde los fragmentos de ssADN con rótulo en 5' se ligan intramolecularmente para generar una biblioteca de fragmentos de ssADN circulares rotulados, cada uno de los cuales exhibe la secuencia de una porción del ADN objetivo y la secuencia del rótulo.

En una versión preferida de esta forma de realización del método, la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ssADN circulares rotulados es generada a partir del ADN objetivo en un tubo simple sin realizar ninguna etapa de purificación interviniente. En una forma de realización preferida, la ligasa independiente de

la plantilla se selecciona de entre ARN ligasa termoestable de bacteriófago TS2126 y una ARN ligasa arqueal (por ejemplo, ARN ligasa 1 de *Methanobacterium thermoautotrophicum*). En algunas formas de realización de preferencia, se proporciona la ligasa dependiente de la plantilla en una forma adenilada y la etapa de incubación de los fragmentos de ssADN con rótulo en 5' con la ligasa independiente de la plantilla se lleva a cabo sin añadir ATP o NAD a la reacción de ligamiento.

En algunas formas de realización de preferencia, el método también comprende: clivaje de los fragmentos de ssADN circulares rotulados en un sitio dentro del rótulo, generando así una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ssADN lineales dirrotulados. En algunas formas de realización, la etapa de clivaje comprende la fusión de un oligodesoxirribonucleótido que es complementario a un sitio de restricción monocatenario dentro del rótulo de los fragmentos de ssADN circulares rotulados y luego clivaje de los fragmentos de ssADN circulares rotulados en el sitio de restricción usando la endonucleasa de restricción que reconoce el sitio de restricción. En algunas otras formas de realización, la etapa de clivaje comprende poner en contacto los fragmentos de ssADN circulares rotulados con una ADN glicosilasa y una endonucleasa, en donde la ADN glicosilasa remueve la base de ácido nucleico de un nucleótido no canónico (por ejemplo, dUMP u 8-oxo-dGMP) que está presente dentro del rótulo y la endonucleasa cliva los fragmentos de ssADN circulares rotulados en el sitio abásico resultante; en algunas formas de realización, la ADN glicosilasa se selecciona de entre uracil-N-glicosilasa y proteína FPG y la AP endonucleasa se selecciona de entre endonucleasa III o endonucleasa IV *E. coli*.

En algunas formas de realización, el método también comprende la etapa de amplificación de la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ssADN circulares rotulados o fragmentos de ssADN lineales dirrotulados, generando así una biblioteca amplificada de fragmentos de ADN rotulados. En algunas formas de realización de preferencia, la etapa de amplificación de la biblioteca de fragmentos de ADN rotulados comprende la realización de una reacción de polimerasa en cadena (PCR), generando así una biblioteca amplificada de fragmentos de ADN rotulados que comprende fragmentos ampliados de ADN dirrotulados. En algunas formas de realización de preferencia, la reacción de PCR se lleva a cabo usando un primer cebador de PCR y un segundo cebador de PCR, donde cada uno tiene una porción 3' y una porción 5', en donde la porción 3' del primer cebador de PCR es complementario a una secuencia exhibida por el rótulo en los fragmentos de ADN rotulados y la porción 3' del segundo cebador de PCR es complementario a una secuencia que es complementaria al rótulo y en donde cada porción 5' comprende un dominio de rótulo de secuenciación que comprende o consiste en un rótulo de secuenciación apropiado que permite el uso de los fragmentos de ADN dirrotulados amplificados generados como plantillas para la secuenciación de próxima generación usando una plataforma de secuenciación de próxima generación particular (por ejemplo, los rótulos de secuenciación de Roche 454A y 454B, los rótulos de secuenciación ILLUMINA™ SOLEXA™, los rótulos de secuenciación SOLID™ de Applied Biosystems, los rótulos de secuenciación SMRT™ de Pacific Biosciences, los rótulos de secuenciación de Pollonator Polony o los rótulos de secuenciación de Complete Genomics).

Métodos para generar bibliotecas de fragmentos de ADN con representación mejorada de secuencias en los extremos del ADN objetivo.

Los inventores observaron determinados datos de secuencia que indicaban que la representación de datos de secuencia de próxima generación de los extremos de ADN objetivo que comprende moléculas de dsADN con un tamaño inferior a 10 Kb era baja en comparación con la representación de los datos de secuencia del medio de ese ADN objetivo. Sin quedar ligados por la teoría, una posible explicación para esta observación es que la probabilidad de hallar fragmentos de ADN con dos composiciones de extremo del transposón insertadas en orientaciones opuestas en los extremos de una molécula de dsADN lineal es menor que la probabilidad de hallar fragmentos de ADN con dos composiciones de extremo del transposón insertadas en orientaciones opuestas en el medio de la molécula de dsADN lineal. A fin de resolver este problema, los inventores desarrollaron métodos adicionales para generar una biblioteca de fragmentos de ADN en donde hay una mejor representación de fragmentos de ADN que exhiben las secuencias en los extremos de las moléculas de dsADN que componen el ADN objetivo.

Una forma de realización preferida es un método para generar una biblioteca de fragmentos de ADN en donde hay una mejor representación de fragmentos de ADN que exhiben las secuencias en los extremos de las moléculas de dsADN que componen el ADN objetivo, que comprende: incubación del ADN objetivo con al menos una composición de transposoma que comprende al menos una transposasa y al menos una composición de extremo del transposón con la que forma un complejo de transposición, donde la composición de extremo del transposón comprende una cadena transferida y una cadena no transferida, en una reacción de transposición *in vitro* en condiciones y durante un tiempo suficiente en donde la cadena transferida se une con el ADN objetivo, generando fragmentos de dsADN con rótulo en 5' que comprende fragmentos de ssADN fusionados con rótulo en 5', cada uno de los cuales tiene un primer rótulo que comprende o que consiste en la cadena transferida en el extremo 5'; desnaturalización de los fragmentos de dsADN con rótulo en 5' para liberar los fragmentos de ssADN con rótulo en 5'; y luego circularización de los fragmentos de ssADN con rótulo en 5' por ligamiento intramolecular con una ligasa independiente de la plantilla que liga el ssADN (por ejemplo, ARN ligasa de bacteriófago TS2126; por ejemplo, ligasa de ssADN CIRCLIGASE™ termoestable, EPICENTRE, Madison, WI, Estados Unidos), generando así la biblioteca de fragmentos de ssADN circulares rotulados. En algunas formas de realización, la al menos una transposasa y la al menos una composición de extremo del transposón se añaden a la reacción como componentes separados más que como el componente individual que comprende la composición de transposoma. En algunas formas de realización, los fragmentos de ssADN circulares rotulados se usan como plantillas de secuenciación de próxima generación o, después de etiquetar, como

objetivo para fusión con sondas en una matriz o micromatriz o para otras aplicaciones descritas en otra parte en la presente. En algunas otras formas de realización, el método también comprende la etapa de linealización de los fragmentos de ssADN circulares rotulados dentro del primer rótulo, generando así fragmentos de ssADN lineales dirrotulados. En algunas de estas formas de realización que comprenden la linealización de los fragmentos de ssADN circulares rotulados, el primer rótulo comprende múltiples dominios de rótulo, en donde la etapa de linealización del primer rótulo da como resultado una porción del primer rótulo en el extremo 5' y otra porción del primer rótulo en el extremo 3'. Por ejemplo, en algunas formas de realización, la cadena transferida de la composición de extremo del transposón transferida exhibe el primer rótulo que comprende múltiples dominios de rótulo (por ejemplo, los dominios de rótulo de secuenciación Roche 454A y Roche 454B), de los que al menos un dominio de rótulo se une con el extremo 3' de los fragmentos de ssADN dirrotulados generados a partir de la etapa de linealización de los fragmentos de ssADN circulares rotulados. Por ejemplo, en algunas formas de realización, los fragmentos de ADN con rótulo en 5' se generan usando una composición de extremo del transposón que comprende una cadena transferida que contiene uno o varios nucleótidos que permiten el clivaje en los sitios de dichos nucleótidos y la etapa de linealización de los fragmentos de ssADN circulares rotulados dentro del rótulo comprende el clivaje de los fragmentos de ssADN circulares rotulados en dichos uno o varios nucleótidos. Por ejemplo, en algunas formas de realización, la cadena transferida contiene uno o varios nucleótidos de desoxiuridina o uno o varios nucleótidos de 8-oxoguanina (por ejemplo, sintetizados usando un sintetizador de oligonucleótido) y la etapa de linealización de los fragmentos de ssADN circulares rotulados dentro del rótulo comprende el clivaje de los fragmentos de ssADN circulares rotulados por incubación de los fragmentos de ssADN circulares rotulados con uracil-ADN glicosilasa o formamidopirimidina-ADN glicosilasa, respectivamente y una endonucleasa que cliva el ADN en un sitio abásico (por ejemplo, endonucleasa IV). Por ejemplo, en algunas otras formas de realización, los fragmentos de ssADN circulares rotulados se linealizan dentro del rótulo por fusión de un oligonucleótido complementario al rótulo y la linealización usando una endonucleasa de restricción que reconoce un sitio de restricción dentro del rótulo bicatenario. En algunas de cualquiera de las formas de realización que comprende la linealización de los fragmentos de ssADN circulares rotulados, el método también comprende la purificación de los fragmentos de ssADN dirrotulados (por ejemplo, usando una columna de limpieza Qiagen PCR); en algunas de estas formas de realización, los fragmentos de ssADN dirrotulados se usan como plantillas de secuenciación de próxima generación o, después del etiquetado, como objetivo para fusión con sondas en una matriz o micromatriz o para otras aplicaciones descritas en otra parte en la presente.

Amplificación de fragmentos de ADN rotulados y otras formas de realización.

En algunas formas de realización de cualquiera de los métodos descritos para generar la biblioteca de fragmentos de ADN rotulados, el método también comprende: amplificación de la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados, fragmentos de ssADN circulares rotulados o fragmentos de ADN de cola de milano.

En algunas formas de realización de cualquiera de los métodos, el método también comprende la etapa de: amplificar la biblioteca de fragmentos de ssADN lineales dirrotulados, fragmentos de ADN circulares rotulados o fragmentos de dsADN de cola de milano usando una reacción de polimerasa en cadena (PCR). Así, en algunas formas de realización, el método también comprende (a) proporcionar (i) primeros y segundos cebadores de PCR, en donde al menos el extremo 3' del primer cebador de PCR es complementario a al menos una porción de la secuencia de rótulo de los fragmentos de ADN circulares rotulados o a al menos una porción de la secuencia de rótulo que se une con el extremo 3' de los fragmentos de dsADN de cola de milano o a al menos una porción de la secuencia de rótulo que se une con el extremo 3' de los fragmentos de ssADN lineales y en donde al menos el extremo 3' del segundo cebador de PCR es complementario a al menos una porción del complemento de la secuencia de rótulo de los fragmentos de ADN circulares rotulados (es decir, en donde al menos el extremo 3' del segundo cebador de PCR exhibe una secuencia que es idéntica a al menos una porción de la secuencia de rótulo) o en donde al menos el extremo 3' del segundo cebador de PCR es complementario a al menos una porción del complemento de la secuencia de rótulo que se une con el extremo 5' de los fragmentos de dsADN de cola de milano o los fragmentos de ssADN lineales dirrotulados (es decir, en donde al menos el extremo 3' del segundo cebador de PCR exhibe una secuencia que es idéntica a al menos una porción de la secuencia de rótulo que se une con el extremo 5' de los fragmentos de dsADN de cola de milano o los fragmentos de ssADN lineales dirrotulados) y (ii) una ADN polimerasa termoestable que se puede usar para PCR; y (b) incubación de los fragmentos de ADN circulares rotulados o los fragmentos de dsADN de cola de milano o los fragmentos de ssADN lineales dirrotulados con los respectivos primer y segundo cebador de PCR y la ADN polimerasa termoestable en condiciones de amplificación por PCR y durante un tiempo suficiente en donde se generan los fragmentos de dsADN lineales dirrotulados amplificados.

En algunas formas de realización en donde el método comprende la amplificación de los fragmentos de ADN circulares rotulados usando PCR, el primer cebador de PCR es complementario a una secuencia de rótulo en al menos una porción de la estructura de bucle de la composición de extremo del transposón en horquilla que se inserta en el ADN objetivo de los fragmentos de ADN circulares rotulados y/o el segundo cebador de PCR exhibe una secuencia que es idéntica a al menos una porción de la estructura de bucle de la composición de extremo del transposón en horquilla que se inserta en el ADN objetivo de los fragmentos de ADN circulares rotulados. En algunas formas de realización, el primer cebador de PCR es complementario a al menos una porción de la secuencia del extremo del transposón transferida o la secuencia del extremo del transposón no transferida y el segundo cebador de PCR es idéntico a al menos una porción de la secuencia del extremo del transposón transferida o la secuencia del extremo del transposón no transferida.

En algunas formas de realización, la porción 5' del primer cebador de PCR o la porción 5' del segundo cebador de PCR o las porciones 5' de ambos cebadores de PCR primero y segundo comprenden o consisten en los primeros y segundos rútuos de secuenciación, respectivamente, para la generación de las plantillas para la secuenciación de próxima generación para una plataforma de secuenciación particular (por ejemplo, rútuos de secuenciación para: una plataforma de secuenciación ROCHE 454A o 454B; para una plataforma de secuenciación ILLUMINA SOLEXA; para una plataforma de secuenciación APPLIED BIOSYSTEMS SOLID™; para una plataforma de secuenciación SMRT™ de PACIFIC BIOSCIENCES; para una plataforma de secuenciación POLLONATOR POLONY; para una plataforma de secuenciación HELICOS; para una plataforma de secuenciación COMPLETE GENOMICS; para una plataforma de secuenciación INTELLIGENT BIOSYSTEMS; o para cualquier otra una plataforma de secuenciación). En algunas formas de realización, la porción 5' del primer cebador de PCR o la porción 5' del segundo cebador de PCR comprende o consiste adicionalmente en un dominio de rútuos de direccionamiento u otro dominio de rútuos para fines particulares. En otras formas de realización, el rútuos de los fragmentos de ADN circulares rotulados comprende un rútuos de secuenciación para secuenciación de próxima generación usando una plataforma particular.

En formas de realización del método en donde una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados se genera usando una ADN polimerasa que tiene actividad de 5' nucleasa o de desplazamiento de cadena, la etapa de amplificación de la biblioteca comprende usando sólo un cebador de oligodesoxirribonucleótido individual que es complementario al segundo rútuos para amplificar la biblioteca de fragmentos de ADN rotulados por PCR. En algunas formas de realización, el cebador individual usado para PCR exhibe al menos una porción de la secuencia del extremo del transposón transferida. En algunas otras formas de realización, el cebador individual usado para PCR exhibe al menos una porción de la secuencia de la porción 5' del oligonucleótido del extremo del transposón transferido. En algunas otras formas de realización preferidas, la etapa de amplificación de la biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados usando un cebador de oligonucleótido simple comprende: proporcionar un cebador de oligonucleótido simple que es complementario al segundo rútuos en el extremo 3' de los fragmentos de ADN rotulados y una ADN polimerasa termoestable que es apropiada para PCR; e incubación de la biblioteca de fragmentos de ADN rotulados con el cebador de oligonucleótido y la ADN polimerasa termoestable en condiciones de amplificación por PCR durante un tiempo suficiente en donde la biblioteca de fragmentos de ADN rotulados se amplifica por PCR, generando una biblioteca de fragmentos de ADN rotulados amplificados.

En algunas otras formas de realización en donde una biblioteca de fragmentos de ADN rotulados que comprende fragmentos de ADN dirrotulados no se genera usando una ADN polimerasa que tiene una actividad de 5' nucleasa o de desplazamiento de cadena, la etapa de amplificación de la biblioteca comprende llevar a cabo una reacción de polimerasa en cadena (PCR), donde el método también comprende: (1) proporcionar (a) cebadores de PCR primero y segundo, en donde al menos el extremo 3' del primer cebador de PCR es complementario a al menos una porción del primer rútuos en el extremo 3' de los fragmentos de ADN dirrotulados o los fragmentos de ADN de cola de milano o a al menos una porción del rútuos en los fragmentos de ssADN circulares rotulados y al menos el extremo 3' del segundo cebador de PCR es complementario a al menos una porción del complemento del segundo rútuos de los fragmentos de ADN dirrotulados o los fragmentos de ADN de cola de milano o a al menos una porción del complemento del rútuos en los fragmentos de ssADN circulares rotulados y (b) una ADN polimerasa termoestable que es apropiada para PCR; y (2) incubación de la biblioteca de fragmentos de ADN rotulados con los cebadores de PCR y la ADN polimerasa termoestable en condiciones de amplificación por PCR y durante un tiempo suficiente en donde la biblioteca de fragmentos de ADN rotulados se amplifica para generar una biblioteca de fragmentos de ADN rotulados amplificados. En algunas formas de realización, el primer o el segundo cebador de PCR comprenden una porción 5' y una porción 3', en donde la porción 5' no es complementaria a la secuencia en el respectivo rútuos o su complemento en los fragmentos de ADN rotulados y la porción 3' es complementaria a la secuencia del respectivo rútuos o su complemento. En algunas formas de realización, la porción 5' de los primeros y segundos cebadores de PCR comprenden o consisten en el primer y el segundo rútuos de secuenciación apropiado que permiten su uso para generar plantillas para secuenciación de próxima generación (por ejemplo, los rútuos de secuenciación de Roche 454A y 454B o los primeros y segundos rútuos de secuenciación apropiados para otra plataforma de secuenciación; por ejemplo, sin limitación, la plataforma sólida de Illumina Solexa o Applied Biosystems).

Una amplia variedad de enzimas y kits son asequibles para realizar la reacción de amplificación por PCR. Por ejemplo, en algunas formas de realización, la amplificación por PCR se lleva a cabo usando ya sea el sistema de PCR FAILSAFF™ o el sistema de PCR extralargo MASTERAMP™ de EPICENTRE Biotechnologies, Madison, WI, tal como se describe por el fabricante. Estos sistemas permiten una rápida optimización de las condiciones de reacción de PCR usando una serie de 2X premezclas de PCR proporcionadas con cada sistema para identificar la premezcla óptima para una plantilla particular y el par de cebadores. Sin embargo, la divulgación no está limitada al uso de aquellos productos o condiciones para la reacción de amplificación y se puede usar cualquier ADN polimerasa termoestable apropiada y la mezcla de reacción que permita la amplificación de la secuencia entre el cebador que se fusiona con la secuencia objetivo y el cebador que se fusiona con el transposón.

La divulgación tampoco está limitada al uso de PCR para amplificar la biblioteca de fragmentos de ADN rotulados. Cualquier método de amplificación apropiado (por ejemplo, amplificación de círculo de rodadura, amplificación del dribocebador (por ejemplo, patente U.S. No. 7.413.857), ICAN, UCAN, ribospia, rotulación terminal (solicitud de patente U.S. No. 20050153333), una amplificación de aARN de tipo Eberwine o amplificación de desplazamiento de cadena) que amplifica la misma secuencia y genera una composición apropiada y una cantidad de producto de

amplificación para el fin pretendido se puede usar en formas de realización de la presente divulgación. Por ejemplo, algunos métodos de desplazamiento de cadena que se pueden usar se describen en la publicación de patente PCT Nros. WO 02/16639; WO 00/56877; y AU 00/29742; de Takara Shuzo Company, Kioto, Japón; patentes U.S. Nros. 5.523.204, 5.536.649, 5.624.825, 5.631.147, 5.648.211, 5.733.752, 5.744.311, 5.756.702 y 5.916.779 de Becton Dickinson and Company; patente U. S. Nros. 6.238.868, 833 6.309y 6.326.173 de Nanogen/Becton Dickinson Partnership; patente U. S. Nros. 5.849.547, 5.874.260 y 6.218.151 de Bio Merieux; patentes U.S. Nros. 5.786.183, 6.087.133 y 6.214.587 de Gen-Probe, Inc.; patente U.S. No. 6.063.604 de Wick et al.; patente U. S. No. 6.251.639 de Kurn; patente U. S. No. 6.410.278 y la publicación PCT No. WO 00/28082 de Eiken Kagaku Kabushiki Kaishi, Tokio, Japón; patentes U. S. Nros. 5.591.609, 5.614.389, 5.773.733, 5.834.202 y 6.448.017 de Auerbach; y patentes U.S. Nros. 6.124.120 y 6.280.949 de Lizardi.

En formas de realización preferidas de la invención, no es necesario seleccionar el tamaño de la biblioteca de fragmentos de ADN con rótulo en 5' generada en la reacción de transposición in vitro o la biblioteca de fragmentos final de ADN rotulados. En el evento, es necesaria la selección del tamaño o la purificación para ciertas aplicaciones, los fragmentos de ADN con rótulo en 5' se pueden seleccionar en tamaño por electroforesis en gel de agarosa (por ejemplo, usando un gel de agarosa no desnaturizante de baja temperatura de fusión de una agarosa de porcentaje apropiado para el rango de tamaño deseado de fragmentos de ADN) y purificar (por ejemplo, para eliminar los oligonucleótidos del extremo del transposón no insertados, otros productos de reacción y gel de agarosa; por ejemplo, por digestión de la porción de gel de agarosa con contenido de rango de tamaño deseado de fragmentos de ADN con rótulo en 5' con enzima de digestión de gel de agarosa GELase™, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos, seguido de precipitación de alcohol y otras etapas de limpieza de acuerdo con direcciones con el producto GELase o usando cualquier otro método de purificación conocido en la técnica). En algunas formas de realización, se usa una etapa de purificación que comprende precipitación de polietilenglicol (PEG) para precipitar la biblioteca de fragmentos de ADN rotulados sin precipitación de las sustancias contaminantes (por ejemplo, sin limitación, oligonucleótidos de rotulación de ligación no ligados u otros componentes de reacción). En algunas formas de realización, se usa una columna de rotación o cualquier otro método conocido en la técnica.

En algunas formas de realización, los fragmentos de ADN circulares rotulados se usan como plantillas para la secuenciación de ADN.

En algunas formas de realización, los fragmentos de ADN rotulados se usan como plantillas para la secuenciación de ADN.

En algunas formas de realización, la biblioteca de fragmentos de ADN rotulados se usa como plantilla para una reacción de amplificación (por ejemplo, una reacción de amplificación por PCR usando cebadores de PCR que son complementarios al primer y al segundo rótulo de fragmentos de ADN rotulados que comprenden fragmentos de ADN dirrotulados o fragmentos de ADN de cola de milano o que son complementarios al rótulo de fragmentos de ADN rotulados que comprenden fragmentos de ssADN circulares rotulados). En algunas formas de realización de preferencia, la biblioteca de fragmentos de ADN rotulados amplificados comprende la mayoría o aproximadamente todas las secuencias exhibidas por el ADN objetivo. En algunas formas de realización en donde el ADN objetivo comprende ADN genómico de un organismo, la reacción de amplificación es una reacción de amplificación de genoma completo.

En algunas formas de realización del método que comprende la amplificación de los fragmentos de ADN rotulados, los amplificados se etiquetan por incorporación de un nucleótido etiquetado durante una o varias etapas del método de amplificación (por ejemplo, el método de reacción de amplificación por PCR). En algunas formas de realización, la biblioteca de fragmentos de ADN rotulados amplificados que contienen la etiqueta se usa para detectar o capturar o detectar y capturar los fragmentos de ADN rotulados amplificados que contienen la etiqueta para una aplicación particular.

Algunas formas de realización de cualquiera de los métodos de la invención para generar una biblioteca de fragmentos de ADN rotulados (por ejemplo fragmentos de ADN dirrotulados) comprenden la generación de una biblioteca de fragmentos de ADN rotulados "etiquetados" que contienen uno o múltiples restos (por ejemplo, una o múltiples moléculas de unión por afinidad) que permitan la captura de los fragmentos de ADN rotulados etiquetados en una superficie o uno o múltiples restos detectables que permitan la detección de los fragmentos de ADN rotulados etiquetados (por ejemplo, que se fusionan con un ADN complementario, como un ADN complementario en un cromosoma). Es decir, algunas formas de realización de cualquiera de los métodos descritos que comprenden además la amplificación de la biblioteca de fragmentos de ADN rotulados comprenden la generación de una biblioteca de fragmentos de ADN rotulados amplificados "etiquetados" que comprende uno o múltiples restos (por ejemplo, una o múltiples moléculas de unión por afinidad) que permitan la captura en una superficie o uno o múltiples restos detectable que permitan la detección de los fragmentos de ADN rotulados etiquetados (por ejemplo, que se fusionan con un ADN complementario, como ADN complementario en un cromosoma). En algunas formas de realización, la biblioteca de fragmentos de ADN rotulados etiquetados o fragmentos de ADN rotulados amplificados etiquetados se genera usando al menos un oligonucleótido etiquetado (por ejemplo, un oligonucleótido del extremo del transposón transferido etiquetado, un oligonucleótido de rotulación de ligamiento etiquetado o al menos un cebador de amplificación etiquetado, como al menos un (o más de un) cebador de PCR). En algunas otras formas de realización, una biblioteca de fragmentos de ADN rotulados amplificados etiquetados se genera incluyendo un dNTP etiquetado que se incorpora

en los productos de amplificación durante la reacción de amplificación. El dNTP etiquetado puede tener cualquier etiqueta conocida en la técnica que se puede usar para generar fragmentos de ADN rotulados amplificados etiquetados, ya sea por etiquetado directo o por etiquetado indirecto. Por "etiquetado directo", se entiende que el resto de captura o la etiqueta detectable se une directamente con los fragmentos de ADN rotulados amplificados sin otro resto entre la captura o el resto detectable y el fragmento de ADN rotulado o el fragmento de ADN rotulado amplificado. Por "etiquetado indirecto", se entiende que hay al menos otro resto entre la captura o resto detectable y el fragmento de ADN rotulado o fragmento de ADN rotulado etiquetado. Un ejemplo de etiquetado directo es incorporar un nucleótido etiquetado con tintura en los fragmentos de ADN rotulados, mientras que un ejemplo de etiquetado indirecto es incorporar un nucleótido etiquetado con biotina en los fragmentos de ADN rotulados y luego etiquetar los fragmentos de ADN rotulados con un resto detectable con tintura al incubar con estreptavidina etiquetada con tintura en condiciones en las que la estreptavidina etiquetada con tintura se une con los nucleótidos etiquetados con biotina. La invención comprende el uso de cualquier método apropiado para generar la biblioteca de fragmentos de ADN rotulados etiquetados o fragmentos de ADN rotulados amplificados etiquetados, en donde la etiqueta se usa posteriormente para captura o detección.

En algunas otras formas de realización, los fragmentos de ADN rotulados en una biblioteca preparados usando un método de la invención luego se etiquetan, directa o indirectamente, poniendo en contacto la biblioteca de fragmentos de ADN rotulados con una molécula de tintura reactiva (por ejemplo, cualquiera de las tinturas fluorescentes reactivas que contienen un N-hidroxisuccinimidilo o éster de "NHS" de Molecular Probes, Eugene, OR) o con una molécula de unión por afinidad reactiva (por ejemplo, un reactivo de biotilación reactivo como un compuesto de biotina-NHS, de Pierce Chemical Company, Rockford, IL). Por ejemplo, en algunas formas de realización, la biblioteca de fragmentos de ADN rotulados amplificados etiquetados se genera incorporando un dNTP que contiene un grupo aminoalilo durante la reacción de amplificación y luego la biblioteca de fragmentos de ADN rotulados amplificados que contienen el grupo aminoalilo se pone en contacto con el éster de NHS de tintura fluorescente etiquetado o el éster de biotina-NHS para generar fragmentos de ADN rotulados amplificados etiquetados con tintura fluorescente o fragmentos de ADN rotulados amplificados etiquetados con biotina, respectivamente. Los expertos en la técnica conocerán o sabrán cómo hallar muchos métodos y reactivos específicos adicionales, incluyendo kits, por ejemplo, de Molecular Probes, para etiquetar la biblioteca de fragmentos de ADN rotulados amplificados para un fin particular (por ejemplo, para permitir la captura en una superficie o detección). Por ejemplo, los ejemplos incluyen uno o varios nucleótidos modificados que tienen un grupo aminoalilo, un grupo propinilo, un grupo biotina, una tintura fluorescente u otra tintura detectable o cualquier otra molécula detectable o combinación de moléculas conocidas en la técnica, incluyendo puntos cuánticos, una enzima (por ejemplo, una fosfatasa, una peroxidasa o una pirofosfatasa) o una proteína detectable (por ejemplo, ficobiliproteína, ficoeritrina). En algunas otras formas de realización, una biblioteca de fragmentos de ADN rotulados amplificados etiquetados se genera por incorporación de uno o varios dNTPs modificados que se etiquetan con una molécula de unión por afinidad o un resto detectable durante la reacción de amplificación, por ejemplo, durante una reacción de amplificación por PCR, por ejemplo, por incorporación de uno o varios dNTPs modificados que tiene un grupo aminoalilo, un grupo biotina, una tintura fluorescente u otra tintura detectable u otro resto que permita ser detectado, ya sea directa o indirectamente después del etiquetado con cualquier otra molécula detectable o combinación de moléculas conocidas en la técnica, incluyendo puntos cuánticos o una enzima o proteína detectable (por ejemplo, ficobiliproteína, ficoeritrina) que se une con una molécula de unión por afinidad (por ejemplo, como estreptavidina, un anticuerpo).

En algunas formas de realización, los fragmentos de ADN rotulados (por ejemplo, fragmentos de ADN dirrotulados se usan para la preparación de fragmentos de ADN etiquetados para hibridación de sondas unidas a una superficie (por ejemplo, como ADN objetivo etiquetado para hibridación en sondas de ADN en una matriz o micromatriz). En algunas formas de realización, los fragmentos de ADN rotulados (por ejemplo, que comprenden fragmentos de ADN dirrotulados) se usan para hibridación en cromosomas o partes de cromosomas en células fijas o secciones de tejidos (por ejemplo, para hibridación fluorescente in situ o FISH).

En algunas formas de realización, el método comprende la generación de ADN fragmentos rotulados etiquetados o fragmentos de ADN rotulados amplificados etiquetados (por ejemplo, de ADN dirrotulados etiquetados o fragmentos de ADN dirrotulados amplificados etiquetados) para usar en la hibridación de cromosomas (por ejemplo, en donde los fragmentos de ADN rotulados etiquetados se preparan a partir de ADN objetivo que comprende ADN de uno o varios cromosomas específicos para usar como "pinturas cromosómicas" (por ejemplo, para hibridación en uno o varios cromosomas en células fijas o secciones de tejidos, por ejemplo, usando hibridación fluorescente in situ o FISH para aplicaciones tales como tipeo de cromosomas o para investigación, diagnósticos médicos, identificación del sexo de un organismo u otras aplicaciones biológicas celulares). En algunas formas de realización, el método comprende la generación de fragmentos de ADN rotulados etiquetados o fragmentos de ADN rotulados amplificados etiquetados de ADN objetivo que comprende partes de cromosomas (por ejemplo, en donde los fragmentos de ADN rotulados se preparan a partir de ADN que codifica uno o varios genes o loci específicos de uno o varios cromosomas (por ejemplo, para hibridación en uno o varios cromosomas en células fijas o secciones de tejido, por ejemplo, usando hibridación fluorescente in situ o FISH o para usar como sondas específicas de genes o sondas específicas de loci en ensayos in vitro para aplicaciones tales como ensayos específicos de análisis o ensayos de diagnóstico para aplicaciones de investigación médica, industrial, ambiental o molecular o de biología celular).

En algunas formas de realización, la hibridación de fragmentos de ADN rotulados etiquetados en sondas sobre una superficie (por ejemplo, una matriz o micromatriz, una varilla graduada, un punto cuántico, una perla o un microcanal

en un dispositivo microfluidico) se usa para detectar, cuantificar, determinar las cantidades relativas o caracterizar una o varias moléculas de ADN o sus porciones que es una fuente natural o que proviene de una fuente natural (por ejemplo, ADN genómico de una célula; por ejemplo, ADN humano para evaluación de variación de número de copias o "CNV" o ADN de una célula bacteria patógena, fúngica, de micoplasma, viral o de nematodo que es un patógeno) o de una fuente in vitro (por ejemplo, cADN bicatenario hecho por transcripción inversa de ARN, como mARN o ARN no codificante o ARN viral que se aísla de una fuente natural o que se amplifica de una fuente natural usando un método de amplificación de ácido nucleico, como un método de amplificación de ADN o ARN).

En algunas otras formas de realización, en las que el método comprende amplificar los fragmentos de ADN rotulados, el método comprende generar fragmentos de ADN rotulados amplificados etiquetados mediante la incorporación de uno o más dNTPs que tienen una molécula de ligación por afinidad o una parte detectable durante la reacción de amplificación, por ejemplo durante una reacción de por PCR (por ejemplo, mediante la incorporación de uno o más dNTPs identificados que tiene un grupo aminoalilo, un grupo biotina, un tinte fluorescente u otro tinte detectable, u otra parte que le permite ser detectado, sea directamente, sea indirectamente después de etiquetado con cualquier otra molécula detectable o combinación de moléculas conocida en la especialidad, lo que incluye puntos cuánticos, o una enzima o proteína detectable (por ejemplo, ficobilproteína, ficoeritrina) que está ligada a una molécula de ligación por afinidad (por ejemplo, tal como estreptavidina, un anticuerpo).

En algunas otras formas de realización, los fragmentos de ADN rotulados o los fragmentos de ADN rotulados amplificados preparados mediante un método de la invención se etiquetan mediante la incorporación de uno o más dNTPs modificados que tienen una molécula de ligación por afinidad o una parte detectable durante la reacción de amplificación (por ejemplo, durante la respectiva transcripción, reacción de RCR o PCR, por ejemplo, por incorporación de uno o más dNTPs modificados que tienen un grupo aminoalilo, un grupo biotina, un grupo digoxigenina, un tinte fluorescente u otro tinte detectable, u otra parte que le permite ser detectado, sea directamente, sea indirectamente después de etiquetado con cualquier otra molécula detectable o combinación de moléculas conocida en la especialidad, lo que incluye puntos cuánticos, o una enzima o proteína detectable (por ejemplo, ficobilproteína, ficoeritrina) que está ligada a una molécula de ligación por afinidad (por ejemplo, tal como estreptavidina, un anticuerpo). En algunas formas de realización, los respectivos productos se utilizan para la preparación de fragmentos de ácido nucleico etiquetados para su hibridación a sondas fijadas a una superficie (por ejemplo, en forma de ácido nucleico objetivo apuntado para su hibridación con sondas de ADN sobre una matriz o micromatriz). En algunas formas de realización, los respectivos productos etiquetados se utilizan para su hibridación a cromosomas o partes de cromosomas en células fijas o secciones de tejidos (por ejemplo, para la hibridación fluorescente in situ o FISH). En algunas formas de realización, la hibridación de productos etiquetados con sondas sobre una superficie se utiliza para detectar, cuantificar, determinar cantidades relativas, o caracterizar una o más porciones de un ADN objetivo tomado de una fuente natural (por ejemplo ADN genómico procedente de una célula; por ejemplo, para la evaluación de la variación de la cantidad de copias o CNV (*copy-number variation*) o procedente de una fuente in vitro (por ejemplo, ADN bicatenario hecho mediante transcripción inversa de ARN, tal como mARN o ARN no codificante (ncARN), que ha sido aislado de una fuente natural o que ha sido amplificado a partir de una fuente natural mediante la utilización de un método de amplificación de ARN).

En algunas formas de realización de métodos que comprenden generar una biblioteca de fragmentos de ADN circulares rotulados, el oligonucleótido extremo de transposón transferido, además de presentar la secuencia del extremo de transposón transferido en su porción 3', también presenta una secuencia de una cadena de un promotor polimerasa bicatenario en su porción 5'. En algunas formas de realización de métodos que comprenden generar una biblioteca de fragmentos de ADN de doble rótulo mediante la utilización de un oligonucleótido rotulador de ligación y de una liga dependiente de plantilla, el oligonucleótido rotulador de ligación presenta una secuencia de una cadena de un promotor polimerasa bicatenaria en su porción 3'. En algunas formas de realización de métodos en los que el nucleótido extremo de transposón transferido o el oligonucleótido rotulador de ligación no presenta una secuencia promotora de polimerasa de ARN, el método comprende además la amplificación por PCR de los fragmentos de ADN dirrotulados mediante la utilización de por lo menos un cebador de PCR que es un "cebador promotor". El cebador promotor tiene una "cola-5' o una porción de "5'-cola" que no tiene una fusión con los fragmentos de ADN dirrotulados y que presenta la secuencia de una cadena de un promotor polimerasa bicatenaria, y una porción 3' que tiene una fusión con el rótulo primero o segundo de los fragmentos 5' 3'-rotulados o de sus complementos.

En algunas formas de realización preferidas en las que el nucleótido extremo de transposón transferido, el oligonucleótido rotulador de ligación, o un cebador de PCR presenta una secuencia promotora de polimerasa de ARN, el ARN polimerasa promotor es un ARN polimerasa de tipo T7 promotor y el método comprende además la etapa de transcribir los fragmentos de ADN 5'- y 3'-rotulados in vitro mediante la utilización de una ARN polimerasa de tipo T7 que reconoce el promotor. Lo más preferible es que la ARN polimerasa y el promotor sean elegidos entre T7 RNAP, T3 RNAP y SP6 RNAP y los correspondientes promotores cognatos. Sin embargo, las etapas de transcripción de un método de la divulgación que pueda utilizar cualquier RNAP para el que se conozca una secuencia promotora adecuada que permita la transcripción con una elevada especificidad puede obtenerse. Los kits y enzimas para la transcripción in vitro pueden obtenerse en el comercio de muchos vendedores y las mezclas de reacción y condiciones adecuadas para llevar a cabo las etapas de la presente divulgación que comprenden la transcripción in vitro pueden utilizar dichos productos de acuerdo con los fabricantes. Por ejemplo, la transcripción in vitro mediante la utilización de T7 RNAP puede llevarse a cabo mediante la utilización del kit AMPLISCRIBE™ T7-Flash™ Transcription Kit o el kit AMPLISCRIBE™ T7 High Yield Transcription Kit de EPICENTRE Biotechnologies, Madison, WI como se describe

en la literatura del producto. De manera similar, si se utiliza T3 RNAP o SP6 RNAP en un método de la invención para la transcripción in vitro, puede utilizarse un kit AMPLISCRIBE™ T3-Flash™ High Yield Transcription Kit o el kit AMPLISCRIBE™ SP6 High Yield Transcription Kit (EPICENTRE Biotechnologies, Madison, WI), respectivamente, como se describe.

5 En algunas formas de realización, el nucleótido extremo de transposón transferido, el oligonucleótido rotulador de ligación, o un cebador de PCR presenta, además de la secuencia promotora de polimerasa de ARN, secuencias adicionales para la traducción, como por ejemplo y sin limitación, un sitio de ligación de ribosoma y un codón de iniciación de traducción (que también lleva la designación de “señal de iniciación de la traducción”), y el método adicionalmente comprende traducir el ARN transcripto. En algunas de estas formas de realización, el método  
10 comprende además la etapa de la traducción in Vitro de los transcritos de ARN resultantes. Los sistemas y kits para la traducción in vitro de los transcritos de ARN también están disponibles en el comercio de muchas fuentes y pueden utilizarse para la presente divulgación. A título de ejemplo y no de limitación, el reticulocito lisado de conejo, el extracto de germen de trigo, y los sistemas de extracto de E. coli S30 vistos por PROMEGA Corporation, Madison, WI pueden utilizarse para la presente invención. Más aún, los kits para la transcripción acoplada in vitro y para la traducción  
15 in Vitro también están disponibles en el comercio y pueden utilizarse, tales como los sistemas TNT® Quick Coupled Transcription/ Translation Systems de Promega.

En algunas formas de realización preferidas del método, la biblioteca de fragmentos de ADN dirrotulados generados a partir de ADN objetivo que comprende una muestra de ADN procedente de un genoma entero de una célula u organismo se amplifican por PCR (es decir, el método comprende o consiste en un método para la amplificación del  
20 genoma completo). En algunas formas de realización, el método para la amplificación del genoma completo se utiliza para amplificar el genoma entero a partir de una célula individual. En algunas formas de realización preferidas del método de la amplificación del genoma completo en la presente, la biblioteca de fragmentos de ADN rotulados se genera a partir de una muestra de ADN de un genoma entero de una célula u organismo que son ampliados por PCR mediante la utilización del cebador oligonucleótido individual (o cebador de PCR) que es complementario con respecto  
25 al segundo rótulo.

En algunas formas de realización, los fragmentos de ADN rotulados generados mediante la utilización de un método de la invención se generan a partir de ADN objetivo que comprende o consiste en genomas y/o ADNc bicatenaria preparado a partir de ARN de todos los organismos (por ejemplo, múltiples organismos) que se hallan presentes en una muestra ambiental (por ejemplo, para aplicaciones metagenómicas o metatranscriptómicas, lo que incluye  
30 aplicaciones industriales, médicas o de investigación).

En algunas otras formas de realización del método, la biblioteca de fragmentos de ADN rotulados se genera a partir de un ADN objetivo que comprende ADN que comprende o consisten un único cromosoma o en una porción de un cromosoma. En algunas de estas formas de realización, el método comprende la ampliación por PCR de la biblioteca de fragmentos de ADN rotulados generado a partir del ADN objetivo que comprende o consiste en ADN de un  
35 cromosoma individual o de una porción de un cromosoma, lo que incluye una porción de un cromosoma que comprende uno o más genes o lugares de genes bajo condiciones en donde los productos amplificados por PCR se etiquetan con una parte detectable (por ejemplo, un tinte fluorescente, infrarrojo fluorescente, quimioluminiscente, visible, u otro tinte detectable; por ejemplo, mediante la utilización de un dNT etiquetado con tinte en el PCR. En algunas formas de realización, los productos amplificados por PCR que se etiquetan con el tinte detectable se utilizan  
40 para teñir células fijas in situ (por ejemplo, los productos de amplificación por PCR se utilizan como pinturas de cromosomas). Por lo tanto, en algunas formas de realización preferidas, el método comprende o consiste en un método para preparar pinturas para cromosomas o pinturas de subcromosomas o marcadores de cromosomas.

En algunas formas de realización, los fragmentos de ADN rotulados o los fragmentos amplificados de ADN rotulados generados mediante la utilización del método se utilizan como el ADN objetivo para una segunda ronda de fragmentación y rotulación mediante la utilización de un método de la invención. En algunas formas de realización, el mismo transposoma se utiliza tanto en la primera como en la segunda ronda del método. En algunas formas de  
45 realización, una segunda transposasa diferente y extremos de transposón diferentes se utilizan para la segunda ronda.

En algunas formas de realización, los fragmentos de ADN rotulados o los fragmentos amplificados de ADN rotulados generados mediante la utilización del método son clonados en un vector (por ejemplo, en un vector COPYCONTROL™  
50 fósido, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos). En algunas formas de realización en donde el método comprende además clonar los fragmentos de ADN rotulados o los fragmentos amplificados de ADN rotulados y en donde los fragmentos de ADN rotulados o los fragmentos amplificados de ADN rotulados (por ejemplo, los fragmentos amplificados por PCR de ADN rotulados presentan un promotor ARN polimerasa, el método comprende además transcribir por lo menos una cadena de los fragmentos clonados de ADN rotulados o los fragmentos  
55 amplificados de ADN rotulados. En algunas formas de realización, los fragmentos clonados de ADN rotulados o los fragmentos amplificados de ADN rotulados son transcritos in vitro mediante la utilización de una ARN polimerasa que reconoce el promotor ARN polimerasa. En algunas formas de realización, los fragmentos clonados de ADN rotulados o los fragmentos amplificados de ADN rotulados son transcritos in vivo en una célula huésped que es capaz de una expresión inducible de la ARN polimerasa que reconoce el promotor ARN polimerasa y que seguidamente transcribe plantillas de ADN que contienen el promotor al que se liga la ARN polimerasa (por ejemplo, el sistema pET se utiliza  
60 ampliamente para la expresión de proteínas in vivo a partir de una ARN polimerasa de tipo T7 inducida). En algunas

formas de realización preferidas, la ARN polimerasa para la expresión in vitro o in vivo es una ARN polimerasa de tipo T7 y la transcripción se inicia a partir de un respectivo promotor ARN de tipo T7 cognato. En algunas formas de realización preferidas, la ARN polimerasa de tipo T7 se selecciona entre T7 ARN polimerasa, T3 ARN polimerasa, y SP6 ARN polimerasa.

5 En algunas formas de realización de cualquiera de los métodos, sea el nucleótido extremo de transposón transferido, el oligonucleótido rotulador de ligación, o un cebador de PCR, contiene, o está unida a una molécula de afinidad (por ejemplo, biotina o digoxigenina), y el método adicionalmente comprende las etapas siguientes: proporcionar una superficie sólida que está recubierta de manera covalente o no covalente con una sustancia de ligación de afinidad que es capaz de ligarse específicamente y de formar un par de ligación específico con la molécula de afinidad (por ejemplo, estreptavidina o avidina para ligar biotina, o un anticuerpo para ligar digoxigenina); y, sea antes sea después de la etapa en la que interviene, poner en contacto los productos generados mediante la utilización del nucleótido extremo de transposón transferido, el oligonucleótido rotulador de ligación, o el cebador de PCR que se une químicamente a la molécula de afinidad bajo condiciones y durante un tiempo suficiente en el que se liga a la sustancia de ligación de afinidad que está unida a la superficie sólida.

15 La invención no se limita a una superficie sólida en particular, que puede ser porosa o no porosa, y de cualquier composición, tamaño o forma que sea adecuada para el método y aplicación particulares. A título de ejemplo, pero no de limitación, la superficie sólida puede seleccionarse del grupo consistente en perlas magnéticas, perlas recubiertas, portaobjetos, las cavidades de una placa de microtitulación, y palillos de inmersión consistentes en vidrio, material plástico, (por ejemplo, látex o poliestireno), sílice, Teflon, u otro material adecuado. La finalidad de la superficie sólida que está recubierta con la sustancia de ligación de afinidad es las de permitir la manipulación (por ejemplo, la captura y lavado para remover las moléculas en una mezcla reactiva), aislación, y captura del nucleótido extremo de transposón transferido, el oligonucleótido rotulador de ligación, o el cebador de PCR que está químicamente unido a la molécula de afinidad, o para permitir la manipulación, aislación, y captura de los fragmentos 5'-de ADN rotulados, de los fragmentos 5'- y 3'- de ADN rotulados, o de los productos de PCR generados a partir de los mismos. A efectos de impedir la ligación no específica, en algunas formas de realización, el soporte sólido es tratado con un gran exceso de una sustancia seleccionado del grupo consistente en: tRNA libre de SADN; proteína (por ejemplo, BSA), polisacárido (por ejemplo, glicógeno, sulfato de dextrano o heparina). La invención tampoco se limita a una sustancia de afinidad específica ni a una sustancia de ligación de afinidad, siempre y cuando sean capaces de ligarse específicamente y de formar un par de ligación específico.

20 Por lo tanto, en algunas formas de realización, los fragmentos de ADN rotulados o los fragmentos de ADN amplificados rotulados son capturados, aislados, purificados o utilizados en otro método mediante su ligación a la superficie sólida, en donde el método comprende los pasos de: poner en contacto los fragmentos de ADN rotulados o los fragmentos amplificados de ADN rotulados que contienen la molécula de afinidad con la superficie sólida en la presencia de reactivos y bajo condiciones que facilitan su ligación a la sustancia de ligación por afinidad que está fijada a la superficie sólida, en donde los fragmentos de ADN rotulados o los fragmentos amplificados de ADN rotulados son ligados a la superficie.

En algunas formas de realización preferidas, la molécula de afinidad es biotina y la sustancia de ligación de afinidad es avidina o estreptavidina, o en donde la molécula de afinidad es digoxigenina y la sustancia de ligación de afinidad es un anticuerpo que liga específicamente digoxigenina.

40 Como se utiliza en la presente, los términos "transposasa" y "ADN polimerasa" y "ligasa" se refieren a moléculas de proteína o a agregados de moléculas de proteína que son responsables de la catalización de reacciones químicas y biológicas específicas. En términos generales, un método, composición, o kit de la invención no se limita al uso de una transposasa particular o de enzima ADN polimerasa procedente de una fuente en particular. En cambio, un método, composición o kit de la presente invención comprende cualquier transposasa o enzima ADN polimerasa de cualquier fuente que tenga una actividad enzimática equivalente a la de las enzimas particulares divulgadas en la presente con respecto al método, composición, o kit, particulares. Más aún, los métodos de la presente invención también incluyen formas de realización en las que solamente una enzima particular que es provista y utilizada en una etapa del método es reemplazada por una combinación de dos o más enzimas que, cuando se las utiliza en combinación, sea utilizadas por separado de una manera escalonada o utilizados conjuntamente con la misma mezcla de reacción, permite obtener resultados que son idénticos a los resultados obtenidos mediante la utilización de la única enzima particular. Los métodos, tampones y condiciones de reacción presentados en la presente, inclusive en los ejemplos, son preferidos en la presente para las formas de realización de los métodos, composiciones y kits de la presente invención. Sin embargo, otros tampones de almacenamiento de enzimas, tampones de reacción, y condiciones de reacción para ser utilizados para algunas de las enzimas de la invención son conocidos en la especialidad, que también pueden ser adecuados para su uso en la presente invención, y se hallan incluidos en la presente.

Formas de realización de composiciones y kits

La invención también se refiere a o comprende kits y composiciones para un método de la descripción. Un kit es una combinación de composiciones individuales útiles para llevar a cabo un método de la invención, en donde las composiciones están optimizadas para su uso conjunto en el método. Una composición comprende un componente individual para por lo menos una etapa de un método de la invención. La invención se refiere a cualquier kit que pueda

ser ensamblado a partir de una combinación de cualesquiera dos composiciones o kits novedosos de la invención, o de cualquier novedosa composición que se utilice en un kit. En algunas formas de realización, el kit o la composición comprende o consiste en un subconjunto de cualquier kit o composición descrito en la presente, en cualquier combinación adecuada y por cualquier razón, de manera de proporcionar al usuario una flexibilidad para adaptar el método a una finalidad o aplicación particular, o para permitir al usuario emplear otras composiciones junto con el kit o composición que comprende o consiste en el subconjunto.

#### Formas de realización de composiciones

Una forma de realización de una composición de la invención es una composición de transposoma que comprende: (i) una cadena transferida que tiene una porción-3' que presenta la secuencia extrema de transposón transferido y una porción-5' que presenta la secuencia de un dominio de rótulo, e: (ii) un cadena no transferida que contiene 5'-fosfato que presenta solamente la secuencia extrema de transposón no transferido, en donde la transposasa forma un complejo con la composición de extremo del transposón que es activa en una reacción de transposición in vitro. En algunas formas de realización, el dominio de rótulo es un dominio de rótulo destinado a ser utilizado en el secuenciado o amplificación de generación siguiente. En algunas formas de realización, el dominio de rótulo se selecciona de entre un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de detección, un dominio de rótulo de dirección, un dominio de rótulo de amplificación, y un dominio de promotor de transcripción.

Otra composición de la divulgación es una composición de extremo del transposón transferida en donde la cadena transferida comprende una porción 3' y una porción 5', en donde la porción 3' presenta una secuencia extrema de transposón transferida y la porción 5' comprende un dominio promotor de la transcripción que presenta una secuencia promotora de ARN polimerasa.

Otra composición de la divulgación es una composición de extremo del transposón en horquilla que comprende o consiste en un oligonucleótido que contiene 5'-fosfato que presenta una secuencia extrema de transposón no transferida en su extremo 5', una secuencia extrema de transposón transferida en su extremo 3', y una secuencia rótulo arbitraria interviniente entre la secuencia extrema de transposón no transferida y la secuencia extrema de transposón transferida que es suficientemente larga para permitir la formación de un vástago-bucle intramolecular. En algunas formas de realización preferidas, la composición de extremo del transposón en horquilla presenta las secuencias extremas de transposón Tn 5 transposasa hiperactiva. En algunas otras formas de realización, la composición de extremo del transposón en horquilla está adenilada en su extremo -5' en lugar de tener un grupo 5'-fosfato.

La invención también comprende composiciones para llevar los métodos a la práctica. Por ejemplo, una composición de la invención es un oligonucleótido que comprende una 3' -porción y una porción 5', en donde la porción 3' presenta una secuencia extrema de transposón transferida y la porción 5' presenta un dominio de sitio de restricción (por ejemplo para una endonucleasa de restricción de corte raro tal como NotI o AscI, o para una endonucleasa de restricción de tipo II tal como FokI). Por ejemplo, otra composición de la invención es un oligonucleótido que comprende una porción 3' y una porción 5', en donde la porción 3' presenta una secuencia extrema de transposón transferida y la porción 5' presenta una secuencia promotora de polimerasa de ARN (por ejemplo, para el fago T7, T3, SP6 o N4 ARN polimerasa). En algunas formas de realización preferidas, la secuencia promotora de polimerasa de ARN es una secuencia de promotor de sentido para cualquiera de estas ARN polimerasas. En algunas otras formas de realización, la secuencia promotora de polimerasa de ARN es una secuencia de promotor antisentido para cualquiera de estas ARN polimerasas. Otra composición de la invención es un oligonucleótido que comprende una porción 3' y una porción 5', en donde la porción 3' presenta una secuencia extrema de transposón transferida y la porción 5' presenta un dominio de rótulo seleccionado de entre un dominio de rótulo de secuenciación, un dominio de rótulo de amplificación, un dominio de rótulo de captura, un dominio de rótulo de dirección, un dominio de rótulo de detección, y un dominio de rótulo de sitio de detección.

En algunas formas de realización preferidas, la secuencia extrema de transposón transferida es la composición de extremo del transposón transferida MEDS o pMEDS para EZ-Tn5<sup>TM</sup> transposasa (EPICENTRE). En algunas formas de realización preferidas, el dominio de rótulo de secuencias presenta un rótulo de secuenciación que es adecuado para una plataforma de secuenciación ROCHE 454, una plataforma de secuenciación ILLUMINA<sup>TM</sup> SOLEXA<sup>TM</sup>, una plataforma de secuenciación SOLID<sup>TM</sup> de LIFE TECHNOLOGIES/APPLIED BIOSYSTEMS, una plataforma de secuenciación SEQPL, una plataforma de secuenciación SMRT<sup>TM</sup> de PACIFIC BIOSYSTEMS, una plataforma de secuenciación POLLONATOR Polony, una plataforma de secuenciación COMPLETE GENOMICS, una plataforma de secuenciación INTELLIGENT BIOSYSTEMS o una plataforma de secuenciación HELICOS.

#### Formas de realización de kits

Una forma de realización de la divulgación es un kit para generar una biblioteca de fragmentos de ADN rotulados en 5' a ser utilizada en la preparación de plantillas para la amplificación de generación siguiente o de ácidos nucleicos, en donde el kit comprende: una composición de transposoma que comprende una transposasa y una composición de extremo del transposón que comprende: (i) una cadena transferida que tiene una porción 3' que presenta la secuencia extrema de transposón transferida y una porción 5' que presenta la secuencia para un dominio de rótulo destinado a

ser utilizado en una reacción de secuenciación o amplificación de generación siguiente; y: (ii) una cadena no transferida que contiene 5'-fosfato que presenta solamente la secuencia extrema de transposón no transferida, en donde la transposasa forma un complejo con la composición de extremo del transposón que es activo en una reacción de transposición in vitro; y un tampón de reacción que contiene dimetilformamida en una cantidad que tiene por resultado

5

que se halla presente en la reacción de transposición in vitro con una concentración final del 10%.

En algunas formas de realización, el kit contiene adicionalmente por lo menos otro componente enzima seleccionada entre: una ADN polimerasa que tiene una actividad de 5' nucleasa o de desplazamiento de cadena; una ADN polimerasa que carece de actividad de 5' nucleasa, una ADN ligasa dependiente de plantilla, y una ligasa independiente de plantilla. En algunas formas de realización, el por lo menos otro componente enzima se selecciona

10

entre: una mezcla FAILSAFE™ ADN polimerasa; Taq ADN polimerasa, Tfi ADN polimerasa, T4 ADN polimerasa, ADN ligasa de E. coli, ARN ligasa de bacteriófago TS2126 termoestable, ARN ligasa de Mth Rn 1 termoestable, y ssADN ligasa CIRCLIGASE™ termoestable.

En algunas formas de realización preferidas en las que la por lo menos una enzima en el kit es una ligasa dependiente de plantilla (por ejemplo, ADN ligasa de E. coli), una elevada proporción de las moléculas de ligasa están adeniladas y no se provee ATP en el kit. En algunas formas de realización en donde la por lo menos una enzima en el kit es una ligasa dependiente de plantilla (por ejemplo, ADN ligasa de E. coli), el kit comprende adicionalmente un oligonucleótido rotulador de ligación que comprende una porción 3' y una porción 5', en donde la porción 3' presenta una secuencia de un dominio de rótulo y la porción 5' presenta una secuencia aleatoria consistente en de aproximadamente tres a aproximadamente ocho nucleótidos. En algunas formas de realización preferidas, el oligonucleótido rotulador de ligación comprende una porción 5' que presenta una secuencia aleatoria consistente en cuatro nucleótidos. En algunas otras formas de realización en donde la por lo menos una enzima en el kit es una ligasa dependiente de plantilla, el kit comprende adicionalmente una composición de extremo del transposón en horquilla. En algunas formas de realización en las que composición de extremo del transposón de horquilla tiene un extremo 5' que está adenilado, menos del 50% de las moléculas que componen la ligasa de ácido nucleico dependiente de plantilla provista en el kit está adenilada y no se provee ATP ni NAD en el kit.

15

20

25

En algunas formas de realización preferidas en las que la por lo menos una enzima en una ligasa independiente de plantilla, seleccionada entre ARN ligasa de bacteriófago TS2126 termoestable, ARN ligasa de Mth Rn 1 termoestable, y ssADN ligasa de CIRCLIGASE™ termoestable, la ligasa independiente de plantilla se provee en una forma elevadamente adenilada y no se provee ATP en el kit.

30

En una forma de realización preferida del kit, el transposoma comprende un transposoma de tipo salvaje o hiperactivo o un transposoma Tn5 hiperactivo que se provee con una concentración de manera tal que la concentración final del transposoma en la reacción de transposición in vitro es de por lo menos 250 nM. En algunas otras formas de realización, la concentración final del transposoma Tn5 de tipo salvaje o hiperactivo o del transposoma MuA es de por lo menos 500 nM.

35

Una forma de realización preferida es un kit para generar fragmentos de ssADN circulares rotulados mediante la utilización de EZ-Tn5™ transposasa y de ADN ligasa de E. coli, comprendiendo el kit: (1) una forma de tipo salvaje o mutante de Tn5 transposasa (por ejemplo, EZ-Tn5™ transposasa); (2) una composición de extremo del transposón que consiste en una cadena transferida que presenta la secuencia extrema de transposón transferida y una cadena no transferida que presenta la secuencia extrema de transposón no transferida para EZ-Tn5 transposasa; (3) tampón de reacción EZ-Tn5 transposasa; y (4) una ligasa de ácido nucleico independiente de plantilla que pueda catalizar la ligación intramolecular de ssADN en la ausencia de una plantilla de ligación (por ejemplo, seleccionado entre la ARN ligasa del TS2126R termoestable (Patente de los Estados Unidos No. 7,303,901); ssADN ligasa de CIRCLIGASE™ termoestable (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos); y ARN ligasa de Mth.1). En una forma de realización preferida, la transposasa en el kit es una forma de tipo salvaje o mutante de Tn5 transposasa (por ejemplo, EZ-Tn5™ transposasa) con una concentración mayor que o igual a aproximadamente 5 unidades por microlitro; aproximadamente 10-20 unidades por microlitro; aproximadamente 20-40 unidades por microlitro; aproximadamente 40-60 unidades por microlitro; de aproximadamente 60-80 unidades por microlitro; o de aproximadamente 80-100 unidades por microlitro. En una forma de realización preferida del kit que comprende EZ Tn5™ transposasa y la ligasa independiente de plantilla, la composición de extremo del transposón EZ-Tn5 pMEDS comprende tanto una cadena EZ-Tn5 pMETS transferida que tiene un grupo 5'-monofosfato y una cadena EZ-Tn5 pMENTS no transferida que tiene un grupo 5'-monofosfato.

40

45

50

En una forma de realización preferida, la transposasa en el kit es una forma de tipo salvaje o mutante de Tn5 transposasa (por ejemplo, EZ-Tn5™ transposasa) con una concentración superior o igual aproximadamente 5 unidades por microlitro; aproximadamente 10-20 unidades por microlitro; aproximadamente 20-40 unidades por microlitro; aproximadamente 40-60 unidades por microlitro; aproximadamente 60-80 unidades por microlitro, o aproximadamente 80-100 unidades por microlitro. En una forma de realización preferida del kit que comprende EZ-Tn5™ transposasa y la ligasa de ácido nucleico independiente de plantilla, la composición de extremo del transposón EZ-Tn5 pMEDS comprende tanto una cadena EZ-Tn5 pMETS transferida que tiene un grupo 5'-monofosfato y una cadena EZ-Tn5 pMENTS no transferida que tiene un grupo 5'-monofosfato.

55

60

Los métodos, composiciones y kits divulgados son útiles para generar fragmentos de ADN circulares rotulados o

5 fragmentos de dsADN o fragmentos de ssADN lineales dirrotulados (y productos de amplificación de los mismos) a partir de ADN objetivo de cualquier fuente para análisis genómico, subgenómico o metagenómico (por ejemplo, destinado a ser utilizado en la preparación de objetivo etiquetado para análisis de micromatriz; por ejemplo, para el análisis de la variación de números de copias, para la detección y análisis de polimorfismos de nucleótidos individuales), y para hallar genes a partir de muestras ambientales tales como fuentes de suelo o agua). Los métodos son útiles en una variedad de procesos, lo que incluye procesos para la amplificación del genoma entero de uno o más organismos, lo que incluye uno o más organismos microbianos o ambientales cuyas condiciones de cultivo son desconocidas (por ejemplo, amplificación de genoma entero (WGA, genoma completo amplificación), PCR en tiempo real, PCR de emulsión, CGH (hibridación genómica comparativa), CGS (secuenciación genómica comparativa), y para la preparación de sondas específicas para ADN (por ejemplo, sondas específicas para cromosomas, por ejemplo, pinturas para cromosomas) para aplicaciones tales como FISH (hibridación fluorescente in situ). En algunas formas de realización, también se utilizan los métodos para generar plantillas para el secuenciado de ADN masivo en paralelo (denominado "secuenciado de generación siguiente"). Cada uno de estos procesos o aplicaciones encuentra usos tanto para fines de investigación o de diagnóstico molecular.

15 Descripción detallada de las formas de realización de la invención

En los capítulos siguientes se presenta la descripción detallada de formas de realización de la invención dadas a título de ejemplo:

I. Fragmentación y dirrotulación de ADN mediante la utilización de una transposasa y de una ADN polimerasa

II. Fragmentación, rotulación y amplificación de ADN objetivo mediante cebador simple

20 III. Fragmentación y rotulación de ADN mediante la utilización de una transposasa y de una ligasa

IV. generación de fragmentos de ssADN circulares rotulados a partir de Ds-ADN objetivo mediante la utilización de una transposasa y de una ligasa

V. Fragmentación y rotulación de Ds-AN mediante transposición in vitro de composición de extremo del transposón en horquilla

25 I. Fragmentación y dirrotulación de ADN mediante la utilización de una transposasa y de una ADN polimerasa

La presente invención comprende métodos y composiciones y se refiere a kits para utilizar transposasa para generar fragmentos rotulados en 5' a partir de ADN objetivo que comprenden o consisten en una o más moléculas bicatenarias (dsADN), y seguidamente unir un segundo rótulo que presenta una secuencia de ADN diferente que el primer rótulo a los extremos -3' de dichos fragmentos de ADN rotulados en 5'. El primer rótulo presenta la secuencia de la cadena transferida del extremo de transposón reconocido por la transposasa, y opcionalmente, también presenta una o más secuencias adicionales que están alejadas en 5' de la secuencia de dicho extremo de transposón transferido. El segundo rótulo se une a los extremos 3' de los fragmentos rotulados 5' in vivo mediante la utilización de una ADN polimerasa sin llevar a cabo una reacción de amplificación por PCR.

Uno de los métodos de la invención comprende: incubar una transposasa y un extremo de transposón con el que forma un complejo de transposición en una reacción de transposición in vitro con ADN objetivo bajo condiciones y durante el tiempo suficiente en donde el extremo de transposón transferido se une al ADN objetivo, generándose fragmentos 5' de ADN rotulados que tienen un primer rótulo en sus extremos 5'; incubar los fragmentos 5' de ADN rotulados con una ADN polimerasa bajo condiciones de polimerización de ADN, las cuales condiciones no comprenden el termociclado, durante un tiempo suficiente en donde un segundo rótulo que presenta una secuencia que es diferente del primer rótulo se une a los extremos 3' de los fragmentos 5' de ADN rotulados, generándose fragmentos de 5'- y 3' de ADN rotulados. En algunas formas de realización, el método comprende además la etapa de amplificar los fragmentos 5'- y 3'- de ADN rotulados mediante la utilización de una ADN polimerasa y de por lo menos un cebador que es complementario con respecto al segundo rótulo. El ADN objetivo puede comprender o consistir en ADN dicatenario procedente de cualquier fuente in vivo o in Vitro, tal como ADN genómico, ADN subgenómico, plásmido u otro ADN de derivación episomal o de ADN recombinante en la presente, o de ADN de doble cadena eDNA hecha mediante transcripción inversa de ARN. El ADN genómico puede comprender o consistir en uno o más genomas de una fuente biológica o ambiental. Por lo tanto, los métodos, composiciones y kits son útiles para generar fragmentos de ADN rotulados en 5' y 3' y, opcionalmente, para amplificar fragmentos de ADN rotulados en 5' y 3' generados a partir de ADN objetivo desde cualquier fuente destinada a ser utilizada en métodos y aplicaciones conocidos en la técnica para análisis genómico, subgenómico, o metagenómico (por ejemplo, para el análisis de la variación de la cantidad de copias, polimorfismos de nucleótidos simples, u otros métodos de análisis genómico. Los métodos son útiles en una variedad de procesos, lo que incluye, pero no se limitan, el análisis metagenómico de genomas de uno o más organismos, lo que incluye uno o más organismos microbianos o ambientales cuyas condiciones de cultivo o desarrollo son desconocidas, PCR de tiempo real, PCR de emulsión, hibridación genómica comparativa (CGH), secuenciación genómica comparativa (CGS). Los métodos son particularmente útiles para generar plantillas para algunos tipos de secuenciación masivamente paralelo de ADN (llamados "secuenciación de próxima generación").

Uso de una polimerasa de ADN que desplaza una cadena con actividad de 5' nucleasa para generar fragmentos de

ADN que tienen rótulos 5' y 3' que presentan las secuencias de diferentes extremos de transposón.

En algunas formas de realización preferidas, la presente invención provee un método para generar una biblioteca de fragmentos de ADN que comprenden fragmentos de ADN rotulados en 5' y 3' a partir de ADN objetivo que comprende o consiste en una o más moléculas bicatenarias (dsADN), que comprende:

5 Proporcionar:

1. ADN objetivo que comprende o consisten una o más moléculas bicatenarias (dsADN) (por ejemplo, ADN o ADNc bicatenario genómico eucariota y/o procariota preparado mediante transcripción inversa de ARN),

10 2. una transposasa (por ejemplo, una transposasa de tipo salvaje o mutante; por ejemplo, transposasa Tn5 de tipo salvaje, por ejemplo, transposasa EZ-Tn5<sup>TM</sup>, o por ejemplo, transposasa HYPERMU<sup>TM</sup> MuA, EPICENTRE 45 Biotechnologies, Madison, WI, Estados Unidos), y

15 3. un extremo de transposón que es capaz de formar un complejo funcional con la transposasa en una reacción de transposición (por ejemplo, el extremo de transposón de extremo exterior 19-bp (OE, outer end), extremo de transposón de extremo interior (IE, inner end), o extremo de transposón "extremo mosaico" ME, mosaic end) reconocido por una Tn5 transposasa de tipo salvaje o mutante, por ejemplo, por la transposasa EZ-Tn5<sup>TM</sup>, o el extremo de transposón R1 y R2, por ejemplo, por la HYPERMU<sup>TM</sup> MuA transposasa), comprendiendo dicho extremo de transposón un ADN bicatenario consistente en una cadena transferida y una cadena no transferida, que, combinados, presentan las secuencias del extremo de transposón bicatenario, en donde la cadena transferida presenta la secuencia de un primer rótulo; y

20 4. una ADN polimerasa que desplaza la cadena o digiere ADN que es fusionado a una cadena de plantilla corriente abajo del extremo 3' de la molécula de ADN que se está extendiendo por dicha ADN polimerasa (es decir, la ADN tiene una actividad de desplazamiento de cadena y/o de 5' nucleasa; por ejemplo, Taq ADN polimerasa, Tfi ADN polimerasa, mezcla de ADN de FAILSAFE<sup>TM</sup>; phi29 ADN polimerasa, E. coli ADN polimerasa I y DISPLACEACE<sup>TM</sup> ADN polimerasa, todas ellas provistas por EPICENTRE Biotechnologies, Madison, WI, Estados Unidos);

25 Incubar el ADN objetivo con la transposasa y el extremo de transposón bajo condiciones y durante un tiempo suficiente en donde la inserción del extremo de transposón catalizado por la transposasa en ambas cadenas del ADN objetivo genera fragmentos de ADN rotulados en 5', cada uno de los cuales tiene rótulo de cebador en su extremo 5' (por ejemplo, FIG. 2); e

30 Incubar fragmentos de ADN rotulados en 5' generados en la reacción de transposición in vitro con la ADN polimerasa bajo condiciones y durante un tiempo suficiente en donde la ADN polimerasa extiende los extremos 3' de los fragmentos de ADN rotulados en 5', con lo que se une un segundo rótulo que presenta por lo menos una porción de la secuencia extrema de transposón no transferida los fragmentos de ADN rotulados en 5' y se generan fragmentos de ADN rotulados en 5' y 3'.

35 En algunas formas de realización de la divulgación, la cadena transferida presenta solamente la secuencia extrema de transposón transferido y, por lo tanto, el primer rótulo que se halla presente en los fragmentos de ADN rotulados en 5' presenta solamente la secuencia extrema de transposón transferido. En algunas formas de realización de la invención, la cadena transferida comprende o consiste en una porción 3' y una porción 5', en donde la porción 3' presenta la secuencia del extremo de transposón transferido y la porción 5' presenta cualquier otra secuencia deseada, en las cuales formas de realización el primer rótulo comprende o consiste tanto en la porción 3' como en la 5'-porción. En formas de realización en las que la cadena transferida comprende o consiste en una porción 3' y una porción 5', la cadena no transferida puede, sin que ello sea necesario, presentar una secuencia que es complementaria a la porción 5' de la cadena transferida.

45 En algunas formas de realización en donde la cadena transferida comprende o consiste en una porción 3' y una porción 5', la porción 5' presenta un rótulo de secuenciación (por ejemplo, el rótulo de secuenciación Roche 454A; como se diagrama, por ejemplo, en la FIG. 10) y la porción 3' presenta la secuencia: de la cadena transferida del extremo de transposón. Esto genera fragmentos de ADN con rotulo 5' con un primer rótulo que comprende o consiste en un rótulo de secuenciación (por ejemplo, el rótulo de secuenciación Roche 454A). Seguidamente, la ADN polimerasa se utiliza para unir un segundo rótulo que comprende o consiste en el otro rótulo de secuenciación (por ejemplo, el rótulo de secuenciación Roche 454B) a los fragmentos de ADN rotulados en 5', con lo que se genera una biblioteca de fragmentos de ADN que comprenden fragmentos de ADN rotulados en 5' y 3' con ambos rótulos de secuenciación (por ejemplo, el 454A y 454B; como se muestra esquemáticamente en la FIG. 10). Los fragmentos de ADN 3'-y rotulados en 5' en el intervalo de tamaños diseñados se utilizan como plantillas para el secuenciación de próxima generación mediante la utilización del sistema Roche 454 Genome Sequencer FLX System. En otras formas de realización, los fragmentos de ADN 5'- y 3'-f rotulados in la biblioteca se generan con rótulos primero y segundo que son adecuados como rótulos de secuenciación para el secuenciado de generación siguiente mediante la utilización de cualquiera otras plataformas de secuenciación (por ejemplo, mediante la utilización de la plataforma de secuenciación ROCHE 454, la plataforma de secuenciación ILLUMINA<sup>TM</sup> SOLEXA<sup>TM</sup>, la plataforma de secuenciación SOLID<sup>TM</sup> de LIFE TECHNOLOGIES/APPLIED BIOSYSTEMS, la plataforma de secuenciación SMRT<sup>TM</sup> de PACIFIC BIOSCIENCES, la plataforma de secuenciación POLLONATOR Polony, la plataforma de secuenciación COMPLETE

GENOMICS, la plataforma de secuenciación INTELLIGENT BIOSYSTEMS, o la plataforma de secuenciación HELICOS ).

En algunas formas de realización, se utilizan múltiples extremos de transposón bicatenario para una transposasa en particular o para múltiples extremos de transposón diferentes reconocidos por diferentes enzimas transposasa. En algunas formas de realización preferidas en las que se utiliza una ADN polimerasa que desplaza cadenas o una ADN polimerasa que tiene una actividad de 5' nucleasa, se insertan dos extremos de transposón diferentes cercanos entre sí en cadenas opuestas del ADN objetivo, y seguidamente la ADN polimerasa extiende los fragmentos de extremo -3 y extremo -5 ADN rotulados mediante la utilización de la cadena opuesta como una plantilla, con lo que se genera se genera una biblioteca de fragmentos de ADN rotulados en 5' y 3' con rótulos que presentan secuencias de extremos de transposón sobre el extremo 3' que son diferentes que sobre el extremo 5' (por ejemplo, en donde la cadena transferida del primer extremo de transposón y la cadena transferida del segundo extremo de transposón son diferentes y están unidos a cadenas opuestas del ADN objetivo (por ejemplo, FIG. 7).

Por lo tanto, en algunas formas de realización, el método comprende proporcionar adicionalmente:

5. una segunda transposasa que reconoce un extremo de transposón diferente del extremo de transposón reconocido por la primera transposasa (que en la presente recibe la denominación del "primer extremo de transposón" reconocido por la "primera transposasa"); y

6. un segundo extremo de transposón que es capaz de formar un complejo funcional con la segunda transposasa en una reacción de transposición, comprendiendo dicho extremo de transposón una cadena transferida y una cadena no transferida, que, combinados entre sí, presentan las secuencias del extremo de transposón bicatenario, en donde la cadena transferida presenta una secuencia que se complementa con la secuencia presentada por el rótulo; e

Incubar el ADN objetivo con la primera transposasa y el primer extremo de transposón y la segunda transposasa y el segundo extremo de transposón bajo condiciones y durante un tiempo suficiente en donde las inserciones primera y segunda catalizadas por transposasa de los extremos de transposón primero y segundo en el ADN objetivo genera fragmentos de ADN, cada uno de los cuales presenta la secuencia de la cadena transferida del primer extremo de transposón o el segundo extremo de transposón sobre su extremo 5'; e

Incubar los fragmentos de ADN rotulados en 5' con la ADN polimerasa bajo condiciones de polimerización de ADN, las cuales condiciones no comprenden la desnaturalización de dsADN ni el termociclado, y durante un tiempo suficiente en donde la ADN polimerasa extiende el extremo 3' de los fragmentos de ADN 5' - rotulados, con lo que se genera la unión del segundo rótulo a los fragmentos de ADN rotulados en 5' y se genera una biblioteca de fragmentos de ADN rotulados (por ejemplo, fragmentos de ADN 5' y 3' rotulados) sin llevar a cabo una reacción de amplificación.

En algunas formas de realización, el método comprende al mismo tiempo incubar el ADN objetivo tanto con la primera transposasa como con los oligonucleótidos del primer extremo de transposón y la segunda transposasa y los oligonucleótidos del segundo extremo de transposón en la misma mezcla de reacción. En algunas otras formas de realización, el método se lleva a cabo consecutivamente empezando por incubar el ADN objetivo con la primera transposasa y los oligonucleótidos del primer extremo de transposón oligonucleótidos y seguidamente se incuban los productos de dicha reacción con la segunda transposasa y los oligonucleótidos del segundo extremo de transposón. En algunas de las formas de reacción en donde el método se lleva a cabo consecutivamente, los productos de la reacción del ADN objetivo con la primera transposasa y los oligonucleótidos de primer extremo de transposón son purificados antes de incubar dichos productos con la segunda transposasa y los oligonucleótidos del segundo extremo de transposón.

En algunas formas de realización del método en donde la cadena transferida comprende o consiste en una porción 3' y una porción 5', la porción 5' presenta la secuencia de un rótulo de secuenciación (por ejemplo, un primer rótulo de secuenciación Roche 454, por ejemplo, el Roche 454A) y la porción 3' presenta la secuencia de la cadena transferida del extremo de transposón. La expresión "rótulo de secuenciación", tal como se la utiliza en la presente, se refiere a un rótulo que se une al extremo 5' o al extremo 3' de un fragmento de ADN monocatenario generado a partir de la molécula de ADN objetivo, el cual rótulo tiene la finalidad de facilitar el secuenciado de dicho fragmento de ADN. Por ejemplo, en algunas formas de realización, el rótulo de secuenciación provee un sitio para capturar dicha cadena de fragmento de ADN sobre una superficie y/o cebar la síntesis de ADN sobre dicho fragmento de ADN y/o el complemento de dicho fragmento de ADN (por ejemplo, como se utilizan los rótulos de secuenciación Roche 454A y 454B para el Genome Sequencer FLX System Roche 454). Por lo tanto, cuando la porción 5' de la cadena transferida presenta la secuencia de un rótulo de secuenciación, los 5'-fragmentos de ADN rotulados tienen un primer rótulo que comprende o consiste en el rótulo de secuenciación (por ejemplo, el rótulo de secuenciación Roche 454). Seguidamente, la ADN polimerasa se une a un segundo rótulo que comprende o consiste en un segundo rótulo de secuenciación (por ejemplo, el rótulo de secuenciación Roche 454B) a los fragmentos de ADN rotulados en 5', con lo que se genera una biblioteca de fragmentos de ADN rotulados en 5' y 3' con rótulos de secuenciación en cada extremo (por ejemplo, los rótulos de secuenciación Roche 454A y Roche 454B). Los fragmentos de ADN rotulados en 5' y 3' en la biblioteca con el intervalo deseado de tamaños se utilizan como plantillas para el secuenciado de generación siguiente mediante la utilización del Genome Sequencer FLX System Roche 454. En otras formas de realización, los fragmentos de ADN rotulados en 5' y 3' se generan con rótulos primero y segundo que son rótulos de secuenciación

para el secuenciado de generación siguiente mediante la utilización de otras plataformas de secuenciación (por ejemplo, mediante la utilización de la plataforma de secuenciación ROCHE 454, la plataforma de secuenciación ILLUMINA™ SOLEXA™, la plataforma de secuenciación SOLID™ de LIFE TECHNOLOGIES/APPLIED BIOSYSTEMS, la plataforma de secuenciación SMRT™ de PACIFIC BIOSCIENCES, la plataforma de secuenciación POLLONATOR Polony, la plataforma de secuenciación COMPLETE GENOMICS, la plataforma de secuenciación INTELLIGENT BIOSYSTEMS o la plataforma de secuenciación HELICOS).

En algunas formas de realización, cada uno de los diferentes extremos de transposón bicatenario comprende una cadena diferente transferida que tiene una 5' porción y una 3' porción, en donde la porción 5' de cada cadena diferente transferida presenta una diferente secuencia de rótulo deseada y la porción 3' presenta la respectiva secuencia extrema de transposón transferido. En algunas formas de realización, por ejemplo, como se muestra en un ejemplo presentado en la FIG. 8, los fragmentos de ADN rotulados en 5' y 3' en la biblioteca tienen ambos un primer rótulo en su extremo 5' y un segundo rótulo en su extremo 3'. Los diferentes extremos de transposón mostrados en la FIG. 8 han sido insertados en diferentes lugares durante los eventos de transposición in vitro separados catalizados por la misma transposasa. Sin embargo, en algunas otras formas de realización, se utilizan diferentes transposasas que forman complejos de transposición con diferentes extremos de transposón. Los diferentes rótulos en las 5'-porciones de la cadena transferida de cada extremo de transposón pueden mostrar cualquier secuencia deseada para cualquier finalidad que se desee. A título de ejemplo, en algunas formas de realización, el primer rótulo y el segundo rótulo de los fragmentos de ADN rotulados en 5' y 3' presentan las secuencias de los rótulos de secuenciación Roche 454A y 454B y, después de aislar los fragmentos en el intervalo de tamaños deseado, se utilizan como plantillas para la generación siguiente mediante la utilización del Genome Sequencer FLX System Roche 454. De manera similar, en otras formas de realización, los fragmentos de ADN 5'- y 3'- rotulados, después de haberse aislado aquellos que se hallan en el intervalo de tamaños deseado, se utilizan como plantillas para la generación siguiente mediante la utilización de otra plataforma de secuenciación (por ejemplo, mediante la utilización de la plataforma de secuenciación ROCHE 454, la plataforma de secuenciación ILLUMINA™ SOLEXA™, las plataforma de secuenciación SOLID™ de LIFE TECHNOLOGIES / APPLIED BIOSYSTEMS, la plataforma de secuenciación SMRT™ de PACIFIC BIOSCIENCES, la plataforma de secuenciación POLLONATOR Polony, la plataforma de secuenciación COMPLETE GENOMICS, la plataforma de secuenciación INTELLIGENT BIOSYSTEMS, la plataforma de secuenciación HELICOS. En algunas formas de realización preferidas, los fragmentos de ADN rotulados en 5' y 3' son generados mediante la utilización de este método a partir de ADN objetivo que comprende un genoma entero de una célula u organismo.

En algunas formas de realización, la cadena transferida del primer extremo de transposón o del segundo extremo de transposón se etiqueta con una molécula de ligación por afinidad (por ejemplo, biotina) o con una molécula detectable (por ejemplo, un tinte fluorescente) que permite la captura (por ejemplo, mediante la utilización de una superficie a la que se une estreptavidina para capturar las moléculas biotiniladas) o la detección de fragmentos de ADN rotulados en 5' y 3' que tienen un rótulo con la molécula de ligación por afinidad o la molécula detectable en el extremo 5'.

Adición de un dominio de rótulo a fragmentos de ADN 5'- y 3'- rotulados. En algunas formas de realización, se utiliza una ADN polimerasa y un oligonucleótido que comprende una plantilla para un dominio de rótulo para añadir un dominio de rótulo a los extremos 3' de fragmentos de ADN rotulados en 5' y 3' en la biblioteca de fragmentos de ADN rotulados (por ejemplo, FIG. 19). En algunas formas de realización, la ADN polimerasa utilizada para añadir el dominio de rótulo es una ADN polimerasa termoestable y el oligonucleótido es un cebador de PCR, y el dominio de rótulo se une al segundo rótulo mediante la realización de PCR.

### III. Fragmentación, rotulado y amplificación de ADN objetivo con cebador simple

Un método preferido comprende: incubar un complejo de transposoma que consiste en una transposasa y un extremo de transposón con el que forma un complejo de transposición con ADN objetivo en una reacción de transposición in vitro bajo condiciones y durante un tiempo suficiente en donde el extremo de transposón transferido se inserta en múltiples sitios en ambas cadenas del ADN objetivo; incubar los productos de la reacción de transposición in vitro con una ADN polimerasa que tiene una actividad de desplazamiento de cadena y/o de 5'- nucleasa bajo condiciones y durante un tiempo suficiente en donde el extremo 3' de cada cadena de ADN objetivo que tiene el extremo de transposón transferido unido a su extremo 5' se extiende mediante la utilización de la cadena opuesta del ADN objetivo como una plantilla, en donde cada una de dichas extensiones catalizadas por ADN polimerasa desplaza o digiere el extremo de transposón no transferido que es fusionado al siguiente extremo de transposón transferido adyacente que se une a la cadena opuesta del ADN objetivo, con lo que se genera una biblioteca de fragmentos de ADN dirrotulados, cada uno de los cuales comprende una porción diferente del ADN objetivo con una cadena transferida en su extremo 5' y una cadena no transferida en su extremo 3', en donde la población de la totalidad de los fragmentos de ADN dirrotulados es sustancialmente representativa de la secuencia del ADN objetivo a partir del cual fueron generados; e incubar la biblioteca de fragmentos de ADN dirrotulados con una ADN polimerasa termoestable y un único cebador que presenta por lo menos una porción de la secuencia extrema de transposón transferido bajo condiciones de termociclado de PCR, con lo que se genera una biblioteca amplificada de fragmentos de ADN dirrotulados. En algunas formas de realización preferidas, el método comprende generar la biblioteca amplificada de fragmentos de ADN dirrotulados en la presencia de uno o más dNTPs etiquetados que se utilizan como sustratos por la ADN polimerasa termoestable, con lo que se genera una biblioteca amplificada de fragmentos de ADN dirrotulados etiquetados.

Por lo tanto, una forma de realización es un método in vitro para la amplificación mediante cebador simple de

fragmentos de ADN generados a partir de un ADN objetivo, comprendiendo el método; llevar a cabo una reacción de transposición en la presencia de un ADN objetivo y en la presencia de un extremo de transposón que consiste en una cadena transferida que presenta la secuencia extrema de transposón transferida y una cadena no transferida que presenta la secuencia extrema de transposón no transferida, resultando dicha reacción de transposición en múltiples inserciones que comprenden unir el extremo de transposón transferido a cada cadena del ADN objetivo, con lo que se generan fragmentos de ADN rotulados en 5' que son fusionados entre sí, cada uno de los cuales tiene un rótulo cebador en su extremo 5' que presenta la secuencia de extremo de transposón transferida; extender los fragmentos de extremo 3' de ADN rotulados en 5' con la ADN polimerasa que tiene una actividad de desplazamiento de cadena y/o de 5' nucleasa mediante la utilización de las cadenas opuestas a las que los fragmentos de ADN rotulados en 5' están fusionados como plantillas; y llevar a cabo una reacción de amplificación por PCR mediante la utilización de un cebador simple que presenta por lo menos una porción de la secuencia extrema de transposón transferida, una ADN polimerasa termoestable, y por lo menos un dNTP etiquetado que se utiliza como un substrato por la ADN polimerasa termoestable, con lo que se genera la biblioteca amplificada de fragmentos de ADN dirrotulados que son representativos del ADN objetivo.

En algunas formas de realización preferidas, el ADN objetivo se selecciona entre ADN genómico eucariota y/o procaríota o DNAC bicatenaria preparada mediante transcripción inversa de ARN.

En algunas formas de realización preferidas, el transposoma es un complejo de una forma de tipo salvaje o hiperactiva de una transposasa seleccionada entre Tn5 transposasa, MuA transposasa, Sleeping Beauty transposasa, Mariner transposasa, Tn7 transposasa, Tn10 transposasa, Tyl transposasa, y Tn552 transposasa y un extremo de transposón con el que la transposasa forma un complejo que es activo en una reacción de transposición.

En algunas formas de realización preferidas, se utiliza una enzima individual o una mezcla de enzimas tanto como la ADN polimerasa que tiene una actividad de desplazamiento de cadenas y de 5' nucleasa y la ADN polimerasa termoestable, la cual ADN polimerasa o mezcla se selecciona entre formas de tipo salvaje o recombinantes de TAQ ADN polimerasa, Tfi ADN polimerasa, Tth ADN polimerasa, y mezcla de ADN FAILSAFE<sup>FM</sup> ADN polimerasa.

En algunas formas de realización preferidas, el por lo menos un dNTP etiquetado comprende una etiqueta, por ejemplo, una cianina, (por ejemplo, Cy5.5, Cy5, Cy3, Cy2), FITC, Alexa Fluors (por ejemplo, 647, 594), Texas Red, JOE, 5-FAM, 6-FAM, VIC, HEX, 6-ROX, Rhodamine, Lissamine, Cyan 500, etc. (Véase, por ejemplo, Handbook of Molecular Probes, R. Haugland, Molecular Probe, Eugene, OR.,).

En algunas formas de realización preferidas, la biblioteca o una biblioteca ampliada de fragmentos de ADN rotulados en 5' y 3' generados mediante la utilización de este método son de ADN que comprende, o que consiste en, de un genoma completo de una célula u organismo. En algunas formas de realización, la biblioteca o una biblioteca amplificada de fragmentos de ADN dirrotulados generados mediante la utilización de este método se basa en ADN objetivo que comprende, o que consiste en, genomas y/o ADNc bicatenario de todos los organismos (por ejemplo, múltiples organismos) que se hallan presentes en una muestra ambiental (por ejemplo, para investigaciones o aplicaciones metagenómicas o metatranscriptómicas).

En algunas formas de realización preferidas del método, la cadena transferida presenta solamente la secuencia extrema de transposón transferida y, por ello, el primer rótulo que se halla presente en los fragmentos de ADN rotulados en 5' presenta solamente la secuencia del extremo de transposón transferida. En algunas otras formas de realización, la cadena transferida comprende o consiste en una porción 3' y una porción 5', en donde la porción 3' presenta la secuencia extrema de transposón transferida y la porción 5' presenta cualquier otro nucleótido o secuencia de nucleótidos deseados, en las cuales formas de realización el primer rótulo comprende o consiste en tanto la porción 3' como la porción 5'. En formas de realización en las que la cadena transferida comprende o consiste en una porción 3' y una porción 5', la cadena no transferida puede, si bien no es necesario, presentar una secuencia que se complementa con la porción 5' de la cadena transferida. En algunas formas de realización preferidas en las que la cadena transferida comprende o consiste en una porción 3' y una porción 5', la porción 5' presenta por lo menos un nucleótido que comprende un dominio de captura (por ejemplo, un nucleótido que comprende una parte biotina, que puede ser capturada por una parte estreptavidina que está adherida a una superficie; o a otra molécula de ligación por afinidad).

## II. Fragmentación y rotulación de ADN mediante la utilización de una transposasa y de una ligasa

La presente invención comprende métodos y composiciones y se refiere a kits para la utilización de una transposasa para generar fragmentos rotulados en 5' a partir de ADN objetivo que comprende o que consiste en una o más moléculas bicatenarias (dsADN), y seguidamente se une un segundo rótulo que presenta una secuencia de ADN diferente del primer rótulo a los extremo 3's de dichos fragmentos de ADN rotulados en 5'. El primer rótulo presenta la secuencia de la cadena transferida del extremo de transposón reconocida por la transposasa, y opcionalmente, también presenta una o más otras secuencias que son 5'- de la secuencia de dicho extremo de transposón transferido. El segundo rótulo se une a los extremos 3's de los fragmentos de ADN rotulados en 5' in vitro mediante la utilización de una ligasa de ácido nucleico.

Un método de la invención comprende: incubar una transposasa y un extremo de transposón con el que forma un

complejo de transposición en una reacción de transposición in vitro con h ADN objetivo bajo condiciones y durante un tiempo suficiente en donde el extremo de transposón transferido se une al ADN objetivo, generar fragmentos de ADN rotulados en 5' que tienen un primer rótulo en sus extremos 5'; incubar los fragmentos de ADN rotulados en 5' con una ligasa de ácido nucleico y un oligonucleótido rotulador de ligación que comprende o consiste en un segundo rótulo bajo condiciones y durante un tiempo suficiente en donde el oligonucleótido rotulador de ligación se une a los extremos 3' de los fragmentos de ADN rotulados en 5', generar una biblioteca de fragmentos de ADN 5'- y 3' rotulados. En algunas formas de realización, el método comprende además la etapa de amplificar la biblioteca de fragmentos de ADN rotulados en 5' y 3' mediante la utilización de una ADN polimerasa y o por lo menos un cebador que se complementa con el segundo rótulo. En algunas formas de realización, la etapa de amplificar la biblioteca de fragmentos de ADN rotulados en 5' y 3' mediante la utilización de una ADN polimerasa comprende la amplificación por PCR mediante la utilización de una ADN polimerasa termoestable, un primer cebador de PCR que es complementario con respecto al segundo rótulo, y un segundo cebador de PCR que presenta una secuencia que es idéntica a por lo menos una porción de la secuencia presentada por el primer rótulo.

Una forma de realización preferida de la presente invención es un método para generar una biblioteca de fragmentos de ADN rotulados que comprenden fragmentos de ADN rotulados en 5' y 3' de ADN objetivo que comprenden o que consisten en una o más moléculas bicatenarias (dsADN), que comprende:

Proporcionar:

1. ADN objetivo que comprende o que consiste en, una o más moléculas bicatenarias (dsADN) (por ejemplo, ADN genómico eucariota y/o procariota o ADNc bicatenario preparado mediante transcripción inversa de ARN),

2. una transposasa (por ejemplo, una transposasa de tipo salvaje o mutante; por ejemplo, Tn5 transposasa de tipo salvaje o mutante, por ejemplo, EZ-Tn5<sup>TM</sup> transposasa, o, por ejemplo, HYPERMU<sup>TM</sup> MuA transposasa, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos), y

3. un extremo de transposón que es capaz de formar un complejo funcional con la transposasa en una reacción de transposición (por ejemplo, el extremo de transposón de extremo exterior (OE, outer end) de 19 bp, el extremo de transposón extremo interior (IE, inner end), o el extremo de transposón ME ("mosaic end", extremo mosaico) reconocido por una Tn5 transposasa de tipo salvaje o mutante, por ejemplo por EZ-Tn5<sup>TM</sup> transposasa, o el extremo de transposón R1 y R2, por ejemplo, por HYPERMU<sup>TM</sup> MuA transposasa), comprendiendo dicho extremo de transposón ADN bicatenario que consiste en una cadena transferida y en una cadena no transferida, que combinadas, presentan las secuencias del extremo de transposón bicatenario, en donde la cadena transferida presenta la secuencia de un primer rótulo,

4. un oligonucleótido rotulador de ligación que tiene un extremo 5' que es capaz de ser ligado al 3' hidroxilo de una molécula de ADN y que presenta una secuencia de un segundo rótulo; y

5. una ligasa de ácido nucleico;

Incubar el ADN objetivo con la transposasa y el extremo de transposón bajo condiciones y durante un tiempo suficientes, en donde la inserción, catalizada por transposasa, del extremo de transposón en el ADN objetivo genera fragmentos de ADN 5-rotulados, cada uno de los cuales tiene el rótulo en su extremo- 5'; e

Incubar los fragmentos de ADN rotulados en 5' con la ligasa de ácido nucleico y el oligonucleótido rotulador de ligación bajo condiciones y durante un tiempo suficientes en donde el segundo rótulo se une a sus extremos 3's y se genera una biblioteca de fragmentos ADN 5'- y 3'-rotulados, cada uno de los cuales fragmentos de ADN rotulados tiene el primer rótulo en el extremo '5 y el rótulo sobre el extremo 3'.

En algunas formas de realización de la divulgación, la cadena transferida presenta solamente la secuencia extrema de transposón transferida y, por ello, el primer rótulo que se halla presente en los fragmentos de ADN rotulados presenta solamente la secuencia extrema de transposón transferida. En algunas formas de realización de la invención, la cadena transferida comprende o consiste en una porción 3' y una porción 5', en donde la porción 3' presenta la secuencia extrema de transposón transferida y la porción 5' presenta cualquier otra secuencia deseada, en las cuales formas de realización el primer rótulo comprende o consiste en tanto la porción 3' como en la porción 5'. En formas de realización en las que la cadena transferida comprende o consiste en una porción 3' y una porción 5', la cadena no transferida puede, sin que sea necesario, presentar una secuencia que se complementa con la porción 5' de la cadena transferida.

En algunas formas de realización en donde la cadena transferida comprende o consiste en una porción 3' y una porción 5', la porción 5' presenta la secuencia de un rótulo de secuenciación (por ejemplo, un primer rótulo de secuenciación Roche 454, por ejemplo, el Roche 454A) y la porción 3' presenta la secuencia de cadena transferida del extremo de transposón. Un "rótulo de secuenciación", tal como se utiliza en la presente, significa un rótulo que se une al 5'-extremo o al 3'-extremo de un fragmento de ADN monocatenario generado a partir de la molécula de ADN objetivo, el cual rótulo tiene la finalidad de facilitar el secuenciado de dicho fragmento de ADN. Por ejemplo, en algunas formas de realización, el rótulo de secuenciación provee un sitio para capturar dicho fragmento de cadena de ADN sobre una superficie y/o para cebar la síntesis de ADN de dicho fragmento de ADN o el complemento de dicho fragmento de

ADN (por ejemplo, como los rótulos de secuenciación Roche 454A y 454B para el genoma Roche 454 Genome. Se utiliza el sistema Sequencer FLX System). Por lo tanto, cuando la porción 5' de la cadena transferida presenta la secuencia de un rótulo de secuenciación, los fragmentos de ADN rotulados en 5' tienen un primer rótulo que comprende o consiste en el rótulo de secuenciación (por ejemplo, el rótulo de secuenciación Roche 454).  
 5 Seguidamente, la ligasa de ácido nucleico liga un oligonucleótido rotulador de ligación que tiene un segundo rótulo que comprende o que consiste en un segundo rótulo de secuenciación (por ejemplo, el rótulo de secuenciación Roche 454B) a los fragmentos de ADN rotulados en 5', con lo que se genera la biblioteca de fragmentos de ADN 5' y 3' rotulados con rótulos de secuenciación en cada extremo (por ejemplo, los rótulos de secuenciación Roche 454A y Roche 454B). Los fragmentos de ADN rotulados en 5' y 3' se generan de manera que tengan un intervalo de tamaños deseado adecuado destinado a ser utilizado como plantillas para la siguiente generación de secuenciación mediante la utilización del sistema Genome Sequencer FLX System Roche 454. En otras formas de realización, se genera una biblioteca de fragmentos de ADN que comprenden fragmentos de ADN rotulados de un tamaño y con rótulos primero y segundo que son adecuados destinados a ser utilizados como rótulos de secuenciación para el secuenciado de generación siguiente mediante la utilización de otra plataforma de secuenciación (por ejemplo, mediante la utilización de la plataforma de secuenciación ROCHE 454, plataforma de secuenciación ILLUMINA™ SOLEXA™, plataforma de secuenciación SOLID™ de LIFE TECHNOLOGIES/APPLIED BIOSYSTEMS, plataforma de secuenciación SMRT™ de PACIFIC BIOSCIENCES, plataforma de secuenciación POLLONATOR Polony, plataforma de secuenciación COMPLETE GENOMICS, plataforma de secuenciación INTELLIGENT BIOSYSTEMS o la plataforma de secuenciación HELICOS).

#### 20 Ligamiento dependiente de plantilla, del segundo rótulo

En algunas formas de realización preferidas, el método comprende proporcionar un oligonucleótido rotulador de ligación que comprende o consiste en una porción 3' y una porción 5', en donde la porción 3' presenta un segundo rótulo que comprende o consiste en cualquier secuencia que se desee unir al extremo 3' de los fragmentos de ADN rotulados en 5' (es decir, una secuencia arbitraria) y la porción 5' tiene un grupo 5'-monofosfato y presenta una secuencia aleatoria (por ejemplo, una secuencia aleatoria que consiste en aproximadamente tres a aproximadamente ocho nucleótidos) en su extremo 5'. En algunas formas de realización preferidas, el oligonucleótido rotulador de ligación tiene una porción-5' de cuatro nucleótidos (por ejemplo, en algunas formas de realización descritas en la presente en donde la transposasa es una transposasa TN5 de tipo salvaje o mutante (por ejemplo, EZ-Tn5™ transposasa, o por ejemplo, HYPERMU™ MuA transposasa, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos) y la ligasa de ácido nucleico es ADN ligasa de E. coli).

La invención no se limita a un oligonucleótido rotulador de ligación que tiene una porción 5' que presenta una secuencia aleatoria que consiste en aproximadamente tres a aproximadamente ocho nucleótidos. Por ejemplo, la invención también incluye métodos en donde el oligonucleótido rotulador de ligación tiene una porción 5' que presenta una secuencia aleatoria consistente en solamente dos nucleótidos; presenta una secuencia aleatoria que consiste en más de ocho nucleótidos; presenta una secuencia semialeatoria en lugar de una secuencia totalmente aleatoria; o que presenta una secuencia que comprende uno o más nucleótidos degenerados (por ejemplo, un nucleótido inosina) en lugar de una secuencia totalmente aleatoria. Sin embargo, se prefiere un oligonucleótido rotulador de ligación que tiene una porción 5' que presenta una secuencia aleatoria que consiste en aproximadamente tres a aproximadamente ocho nucleótidos.

En algunas formas de realización preferidas, la ligación del oligonucleótido rotulador de ligación al extremo 3' de los fragmentos de ADN rotulados en 5' tiene lugar solamente en la presencia de una plantilla de ADN que presenta una secuencia que es exactamente complementaria a la unión de ligación; en tales formas de realización, la plantilla a la cual se ligan ambas moléculas de ácido nucleico fusionados recibe en la presente la denominación de "plantilla de ligación" y la ligación recibe la denominación de "ligación dependiente de plantilla". En algunas formas de realización en las que la ligación tiene lugar solamente en la presencia de una plantilla de ligación, la ligasa de ácido nucleico es una ADN ligasa que requiere una plantilla de ligación, y en la presente recibe la denominación de "ligasa dependiente de plantilla" (por ejemplo, una ADN ligasa dependiente de plantilla de tipo N tal como, pero sin limitación, ADN ligasa de E. coli, Tth ADN ligasa, Tfl ADN ligasa, o AMPLIGASE® ADN ligasa, provistos por EPICENTRE Biotechnologies, Madison, WI, Estados Unidos). En algunas otras formas de realización en donde la ligación tiene lugar solamente en la presencia de una plantilla de ligación, la ligasa de ácido nucleico es una ADN ligasa, que si bien no requiere una plantilla de ligación para la ligación, cataliza la ligación más eficientemente en la presencia de la plantilla de ligación que en su ausencia (por ejemplo, una ADN ligasa dependiente de plantilla de tipo ATP tal como, pero sin limitación, T4 ADN ligasa o FASTLINK™ ADN ligasa, provistos por EPICENTRE Biotechnologies, Madison, WI, Estados Unidos). Si la ligación tiene lugar sobre una plantilla de ligación, la ligación recibe la denominación de "ligación dependiente de plantilla" en la presente, aun si la ligasa podría también catalizar la ligación independientemente de la plantilla. En formas de realización preferidas, la secuencia aleatoria del oligonucleótido rotulador de ligación es breve, en las cuales formas de realización se prefiere un ácido nucleico que cataliza ligasa dependiente de plantilla a una temperatura inferior (por ejemplo, inferior a una temperatura aproximadamente igual a 40 °C, inferior o igual a 37 °C, inferior o igual a 30 °C, inferior a o igual a aproximadamente 25 °C, o inferior a o igual a aproximadamente 20 °C). En algunas formas de realización preferidas, la ligasa dependiente de plantilla es ADN ligasa de E. coli.

La invención no está limitada con respecto al método de ligación utilizado salvo que con respecto a las formas de realización que comprenden ligación dependiente de plantilla, la ligación debería tener lugar de manera eficiente en la

presencia de una secuencia objetivo al que el oligonucleótido rotulador de ligación y los fragmentos de ADN rotulados en 5' se fusionan de manera continua y la ligación debería tener lugar raramente o no transcurrir en absoluto en la ausencia de una secuencia objetivo. Como se utiliza en la presente, la expresión "ligación dependiente de plantilla" se refiere a cualquier método adecuado para unir los extremos 5' y 3' adyacentes de los oligonucleótidos rotuladores de ligación y los fragmentos de ADN rotulados en 5', respectivamente, que son adyacentes a o contiguos o que hacen tope entre sí cuando se fusionan a una secuencia objetivo.

La inserción catalizada por transposasa del extremo de transposón en ambas cadenas del ADN objetivo tiene como resultado la fragmentación del ADN objetivo, la unión del extremo de transposón transferido a cada cadena del ADN objetivo, y la generación de una región de base 3' de ADN objetivo monocatenario del sitio de unión del extremo de transposón al mismo, lo que resulta en una región de brecha o huelgo en la cadena opuesta del ADN objetivo (por ejemplo, FIG. 3). La región monocatenaria de ADN objetivo corriente abajo del extremo de transposón transferido puede servir como una plantilla de ligación para fusionar la porción aleatoria del oligonucleótido rotulador de ligación. Entre la totalidad de las secuencias representadas por el oligonucleótido rotulador de ligación que presenta la secuencia aleatoria en su porción 5' (lo que incluye todas las secuencias posibles), por lo menos una de ellas presenta una secuencia en su extremo 5' que es capaz de fusionarse a cada región monocatenaria en el ADN objetivo por lo que el extremo 5'-fosforilado de dicho oligonucleótido rotulador de ligación es adyacente a, y hace tope con, el extremo 3' del ADN objetivo que es complementario del fragmento de ADN 5-rotulado, en cuyo caso, la ligasa de ácido nucleico puede entonces catalizar la ligación dependiente de plantilla del oligonucleótido rotulador de ligación a dicho extremo 3'. En algunas formas de realización, la cadena transferida del extremo de transposón es insertada en cadenas opuestas del ADN objetivo en dos ubicaciones que se hallan en una proximidad relativamente estrecha, generándose dos fragmentos de ADN rotulados en 5' como se muestra en in FIG. 2, y dos fragmentos de ADN rotulados en 5' y 3' como se muestra en in FIG. 7. Al tener lugar la desnaturalización de los dos fragmentos de ADN 5'- y 3'- rotulados, se los puede utilizar para una variedad de aplicaciones, lo que incluye para la amplificación (por ejemplo, por PCR mediante la utilización de un primer cebador de PCR que es complementario con respecto al segundo rótulo y un segundo cebador de PCR que es complementario con respecto al primer rótulo).

En algunas formas de realización el método comprende la utilización de una ligasa de ácido nucleico y ligación dependiente de plantilla para unir un segundo rótulo a los fragmentos de ADN 5' rotulados, los fragmentos de ADN 5'- y 3'- comprenden un primer rótulo 5' que presenta una secuencia extrema de transposón y un segundo rótulo 3' que no presenta una secuencia de extremo de transposón (si bien la porción 3' de oligonucleótido rotulador de ligación puede comprender o consistir en un segundo rótulo que presenta cualquier secuencia deseada, lo que incluye, si se desea, una secuencia extrema de transposón). Por ejemplo, en algunas formas de realización, la porción 3' del oligonucleótido rotulador de ligación presenta la secuencia del rótulo de secuenciación Roche 454 o de un rótulo de secuenciación para otra plataforma de secuenciación (por ejemplo, mediante la utilización de la plataforma de secuenciación ROCHE 454, la plataforma de secuenciación ILLUMINA™ SOLEXA™, la plataforma de secuenciación SOLID™ de LIFE TECHNOLOGIES/APPLIED BIOSYSTEMS, la plataforma de secuenciación SMRT™ de PACIFIC BIOSCIENCES, la plataforma de secuenciación POLLONATOR Polony, la plataforma de secuenciación COMPLETE GENOMICS, la plataforma de secuenciación de INTELLIGENT BIOSYSTEMS, o la plataforma de secuenciación HELICOS).

A título de otro ejemplo, en algunas formas de realización, el segundo rótulo en la 3'- porción del oligonucleótido rotulador de ligación presenta la secuencia de un promotor ARN polimerasa promotor (por ejemplo, un promotor de ARN polimerasa de tipo T7; por ejemplo, a promotor T7, T3, SP6, o fago N4 MINI-VrM ARN polimerasa; EPICENTRE Biotechnologies, Madison, WI, Estados Unidos; en general, si la ARN polimerasa requiera un promotor polimerasa bicatenaria, el segundo rótulo en estas formas de realización presenta la "secuencia de sentido promotora de polimerasa de ARN", lo que significa la secuencia del promotor polimerasa que se une al extremo 3' de la plantilla de cadena ADN que es transcrita por la ARN polimerasa, en las cuales formas de realización, la "secuencia de promotor antisentido de ARN polimerasa" complementaria ha de ser provista o sintetizada durante una etapa o un método para generar un promotor polimerasa bicatenaria que será reconocida por la ARN polimerasa. En algunas formas de realización, el segundo rótulo en la porción 3' del oligonucleótido rotulador de ligación también presenta la secuencia de un "rótulo de dirección" lo que significa una secuencia que permite la identificación de un rótulo de dirección simple específico (por ejemplo, mediante la utilización de un rótulo de dirección en un oligonucleótido rotulador de ligación que presenta una secuencia diferente para cada muestra de ADN objetivo). En algunas formas de realización, la porción 3' del oligonucleótido rotulador de ligación también presenta la secuencia de uno o más rótulos para una finalidad particular en el método.

En algunas formas de realización preferidas, la porción 5' del oligonucleótido rotulador de ligación es una secuencia aleatoria de una longitud que es capaz de fusionarse a las porciones monocatenarias de los fragmentos de ADN rotulados en 5' generados a partir de la inserción catalizada de los extremos de transposón en ambas cadenas del ADN objetivo. La secuencia aleatoria en la porción 5' del oligonucleótido rotulador de ligación se fusiona a esta región de brecha monocatenaria adyacente al extremo 3' de los fragmentos de ADN rotulados en 5', que sirve como una plantilla de ligación para la ligación dependiente de plantilla del oligonucleótido rotulador de ligación a los extremos 3' de la cadena complementaria del ADN objetivo. En formas de realización en las que se utiliza EZ-Tn5™ transposasa, la inserción de los oligonucleótidos de extremo de transposón 19-bp EZ-Tn5™ en ambas cadenas del ADN objetivo genera fragmentos de ADN rotulados en 5' que presentan brechas de base 9 que consisten en regiones monocatenarias que están opuestas a los sitios de inserción del extremo de transposón transferido. Sin embargo, el

tamaño de la brecha en donde el oligonucleótido rotulador de ligación puede fusionarse, varía para diferentes enzimas transposasa. Por ejemplo, la MuA transposasa genera una región monocatenaria de ADN objetivo corriente abajo del sitio de inserción del extremo de transposón transferido que es de solamente cinco nucleótidos. En algunas formas de realización, la secuencia aleatoria del oligonucleótido rotulador de ligación comprende o consiste en entre

5 aproximadamente tres y aproximadamente ocho nucleótidos aleatorios. Sin embargo, la longitud de la porción de secuencia aleatoria del oligonucleótido rotulador de ligación puede variar para diferentes enzimas transposasa en base al tamaño de las regiones monocatenarias generadas y otros factores, tales como la longitud de la secuencia aleatoria que es la que más eficientemente se liga con la respectiva ligasa de ácido nucleico y en base a las condiciones de ligación utilizadas. Por ejemplo, los solicitantes han observado que, mediante la utilización de fragmentos de ADN

10 rotulados en 5' generados mediante la utilización de EZ-Tn5<sup>TM</sup> transposasa, un oligonucleótido rotulador de ligación con una porción 5' que consiste en una secuencia aleatoria de cuatro nucleótidos generados da buenos rendimientos de fragmentos de ADN 5' y 3' rotulados mediante la utilización de ADN ligasa de E. coli como la ligasa de ácido nucleico. Sin embargo, este oligonucleótido rotulador de ligación con una porción 5' que consiste en una secuencia aleatoria de cuatro nucleótidos no fue ligada de manera eficiente por ligasas ADN dependientes, tal como ligasa

15 termoestable AMPLIGASE® bajo condiciones de ligación similares. En formas de realización preferidas, los fragmentos de ADN rotulados en 5' y 3' comprenden o consisten en la totalidad de los fragmentos de ADN generados a partir de la muestra de ADN.

En algunas formas de realización, la ligasa de ácido nucleico para la ligación dependiente de plantilla es una ligasa que utiliza NAD como un co-factor. En algunas formas de realización, la ligasa de ácido nucleico para la ligación dependiente de plantilla se selecciona de entre las siguientes ligasas de ADN de tipo NAD: ADN ligasa de E. coli, Tth

20 ADN ligasa, Tfi ADN ligasa, y Ampligase® ADN ligasa (todos provistos por EPICENTRE Biotechnologies, Madison, WI, Estados Unidos), y Tsc ADN ligasa (Roche Applied Systems, Indianápolis, IN, Estados Unidos). En algunas formas de realización, la ligasa de ácido nucleico para la ligación dependiente de plantilla es una ADN ligasa de tipo ATP. En algunas formas de realización, la ADN ligasa de tipo ATP se selecciona entre: T4 ADN ligasa y FASTLINK<sup>TM</sup> ADN

25 ligase (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos). En algunas formas de realización preferidas, la ligasa de ácido nucleico se selecciona de entre E. coli ADN ligase u otra ligasa de ADN bacteriano mesófilo como un co-factor. En algunas formas de realización preferidas, la selección del tamaño y la purificación de los fragmentos de ADN 5' de tamaño seleccionado se lleva a cabo mediante la Utilización de una ADN ligasa dependiente de plantilla de ADN (por ejemplo, ADN ligasa de E. coli).

#### 30 Ligación independiente de plantilla del segundo rótulo

En algunas formas de realización el método comprende unir el segundo rótulo al extremo 3' de los fragmentos de ADN rotulados en 5' mediante la utilización de una ligasa de ácido nucleico, el oligonucleótido rotulador de ligación que presenta el segundo rótulo es ligado directamente al extremo 3' de los fragmentos de ADN rotulados en 5' sin fusionar

35 el oligonucleótido rotulador de ligación a una plantilla de ligación adyacente a los extremos-3 de los fragmentos de ADN rotulados en 5'. En estas formas de realización, el oligonucleótido rotulador de ligación no presenta una secuencia aleatoria, sino que presenta solamente la secuencia del segundo rótulo que se desea unir a los fragmentos de ADN rotulados en 5'. En estas formas de realización, el oligonucleótido rotulador de ligación se liga a los extremos 3 de los fragmentos de ADN rotulados en 5' monocatenarios sin utilizar una plantilla de ligación. En estas formas de realización del método, la ligasa de ácido nucleico es una ligasa de ácido nucleico que es capaz de ligar una molécula de ADN

40 monocatenaria que un grupo 3'-hidroxilo a una molécula de ADN monocatenaria que tiene un grupo 5'-monofosfato en la ausencia de fusión a una secuencia complementaria en la unión de ligación (por ejemplo, seleccionado entre T4 ARN ligasa 1, T4 ARN ligasa 2, ARN ligasa TS2126 bacteriófago termoestable, y ARN ligasa CIRCLIGASE<sup>TM</sup>, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos; y el método comprende adicionalmente la etapa de: desnaturalizar dsADN que comprenden los fragmentos de ADN rotulados en 5' antes de incubar los fragmentos de

45 ADN rotulados en 5' con la ligasa de ácido nucleico y el oligonucleótido rotulador de ligación.

La invención no se limita a una ligasa particular ligasa de ácido nucleico y a los métodos que comprenden incubar los fragmentos de ADN rotulados en 5' con una ligasa de ácido nucleico bajo condiciones y durante un tiempo suficiente en donde el segundo rótulo se une a sus extremos 3 y a una biblioteca de fragmentos de ADN 5' y 3' rotulados se entenderá que también comprende el uso de otras composiciones en lugar de la ligasa de ácido nucleico para la unión

50 dependiente de plantilla o independiente de plantilla. A título de ejemplo, pueden utilizarse otros métodos de ligación tales como, sin limitación, el uso de un oligonucleótido rotulador de ligación que comprende una parte topoisomerasa, en donde la ligación comprende la ligación mediada por topoisomerasa (por ejemplo, la patente de los Estados Unidos No. 5.766.891) si bien en la mayoría de los casos no se prefiere la ligación mediada por topoisomerasa.

#### 55 IV. Generación de fragmentos de ssADN circulares rotulados a partir de Ds-ADN objetivo mediante una transposasa y una ligasa (30842)

La presente invención comprende métodos y composiciones y se refiere a kits para generar una biblioteca que comprende una población de fragmentos de ssADN circulares rotulados a partir de ADN objetivo en una muestra destinada a ser utilizada como plantilla en la secuenciación de ADN o de reacciones de amplificación de ácidos nucleicos. En general, cada fragmento de ssADN circular rotulado en la biblioteca presenta una secuencia contigua

60 de una porción del ADN objetivo y de un rótulo.

En pocas palabras, en determinadas formas de realización, el método comprende: incubar el ADN objetivo, que es por lo general dsADN, con una transposasa y una composición de extremo del transposón en una reacción de transposición *in vitro* para al mismo tiempo fragmentar y rotular el ADN objetivo, con lo que se genera una población de fragmentos de ADN rotulados; seguido por desnaturalización de los fragmentos de ADN rotulados para generar fragmentos de ssADN rotulados en 5', y seguidamente incubar los fragmentos de ssADN rotulados en 5' con una ligasa de ácido nucleico independiente de plantilla o no homóloga que es capaz de catalizar la ligación intramolecular independiente de plantilla (es decir, la circularización) de ssADN para generar una biblioteca de fragmentos de DNAss circulares rotulados. En algunas formas de realización, los fragmentos de ssADN circulares rotulados son linealizados mediante la fusión de un oligodesoxirribonucleótido que se fusiona a un sitio de restricción dentro del rótulo, después de lo cual se trata con la endonucleasa de restricción para generar fragmentos de ssADN lineales que tienen a porción del rótulo en sus extremos 5' y la porción restante del rótulo en sus extremos 3' (los cuales fragmentos de ssADN lineales en la presente se designan como "fragmentos de ssADN lineales dirrotulados" o simplemente como "fragmentos de ssADN dirrotulados").

En algunas formas de realización los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados se utilizan como plantillas de ADN en la amplificación de ácidos nucleicos y/o reacciones de secuenciación de ADN. En algunas formas de realización, el método comprende además la etapa de amplificar y/o secuenciar la biblioteca de los fragmentos de ssADN circulares rotulados (por ejemplo, para amplificar y/o determinar la secuencia del ADN objetivo). En algunas formas de realización, el método comprende además la etapa de la secuenciación del ADN que es complementario con respecto al ADN objetivo obtenido mediante amplificación de los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados. En algunas formas de realización, por lo menos una porción del ADN objetivo en cada uno de los fragmentos de DNAss circulares rotulados o de los fragmentos de ssADN lineales dirrotulados es secuenciado mediante una ADN polimerasa y por lo menos un cebador que es complementario con respecto al rótulo (por ejemplo, para la secuenciación por síntesis). En algunas formas de realización, por lo menos una porción del ADN objetivo en cada uno de los fragmentos de ssADN circulares rotulados o de los fragmentos de ssADN lineales dirrotulados es secuenciado mediante una ligasa dependiente de plantilla para ligarse a por lo menos un oligodesoxirribonucleótido que es complementario con respecto al rótulo y por lo menos otro oligodesoxirribonucleótido que se fusiona a la porción de la secuencia objetivo (por ejemplo, secuenciación mediante ligación). En algunas formas de realización, por lo menos una porción del ADN objetivo en cada uno de los fragmentos de ssADN circulares rotulados o de los fragmentos de ssADN lineales dirrotulados es secuenciado mediante la fusión de oligodesoxirribonucleótidos que se fusionan o hibridan al rótulo y a una porción de la secuencia objetivo (por ejemplo, secuenciación por hibridación). En algunas formas de realización, el ssADN que es complementario con respecto a los fragmentos de ssADN circulares rotulados o de los fragmentos de ssADN lineales dirrotulados es secuenciado mediante secuenciación por síntesis, secuenciación por ligamiento o secuenciación por hibridación.

Por lo tanto, una forma de realización preferida de la presente invención es un método para generar una biblioteca que comprende una población de fragmentos de ssADN circulares rotulados de ADN objetivo en una muestra destinada a ser utilizada como plantillas en la secuenciación de ADN o en reacciones de amplificación de ácidos nucleicos, cada uno de los cuales fragmentos de ssADN circulares rotulados presenta la secuencia de una porción del ADN objetivo y la secuencia de un rótulo que se une a la porción la secuencia objetivo, que comprende:

Proporcionar:

1. ADN objetivo que comprende o que consiste en una o más moléculas bicatenarias (dsADN) (por ejemplo, ADN genómico eucariota y/o procariota o ADNc bicatenario preparado mediante transcripción inversa de ARN mediante la utilización de una ADN polimerasa dependiente de ARN o de transcriptasa inversa para generar ADNc de primera cadena y seguidamente extender un cebador fusionado al ADNc de primera cadena para generar dsADN),

2. una transposasa (por ejemplo, una transposasa de tipo salvaje o mutante; por ejemplo, Tn5 transposasa de tipo salvaje o mutante, por ejemplo, EZ-Tn5<sup>TM</sup> transposasa, por ejemplo, HYPERMU<sup>TM</sup> MuA transposasa, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos), y

3. una composición de extremo del transposón que es capaz de formar un complejo funcional con la transposasa en una reacción de transposición (por ejemplo, que comprende o que consiste en el 19-bp OE (outer end) extremo de transposón, el 19 bp IE (inner end) extremo de transposón, o el 19-bp ME ("mosaic end") extremo de transposón reconocido por una Tn5 transposasa de tipo salvaje o mutante, por ejemplo, mediante EZ-Tn5<sup>TM</sup> transposasa), en donde dicha composición de extremo del transposón comprende o consiste en una cadena transferida y una cadena no transferida, que, en combinación, presentan las secuencias del extremo de transposón bicatenario, en donde la cadena transferida presenta la secuencia del rótulo.

4. una ligasa de ácido nucleico independiente de plantilla o no homóloga, que es capaz de ligación intramolecular independiente de plantilla o circularización de ssADN que tiene un grupo 5'-monofosfato y un grupo 3'-hidroxilo (por ejemplo, ARN ligasa termoestable de fago TS2126, por ejemplo, en donde una elevada proporción de la ARN ligasa está adenilada);

Incubar el ADN objetivo con la transposasa y la composición de extremo del transposón bajo condiciones y durante un tiempo suficientes en donde la inserción catalizada por transposasa, de la cadena transferida en el ADN objetivo

genera fragmentos de ADN rotulados en 5' (por ejemplo, FIG. 2); y

Desnaturalizar el ADN objetivo que comprende fragmentos de ADN rotulados en 5' de manera de obtener fragmentos de ssADN rotulados en 5'; y

5 Incubar los fragmentos de ssADN rotulados en 5' con la ligasa de ácido nucleico bajo condiciones y durante un tiempo suficientes en donde los fragmentos de ssADN rotulados en 5' son ligados intramolecularmente para generar una biblioteca de fragmentos de ssADN circulares rotulados, cada uno de los cuales presenta la secuencia de una porción del ADN objetivo y del rótulo.

10 En algunas formas de realización, antes de la etapa de la ligación, el método comprende adicionalmente una o más etapas para remover ADN objetivo que no se rotula durante la reacción de transposición y/o para remover los componentes de la composición de extremo del transposón que no están unidos al ADN objetivo.

15 En algunas formas de realización, el método adicionalmente comprende tratar la biblioteca que contiene los fragmentos de ssADN circulares rotulados con exonucleasa I a efectos de remover el ssADN lineal no ligado. En algunas formas de realización, el método adicionalmente comprende la etapa de tratar la mezcla de reacción con exonucleasa I y exonucleasa III (EPICENTRE Biotechnologies, Madison, WI) para remover el ssADN lineal no ligado. La exonucleasa III ayuda en la remoción de parte del ssADN lineal mediante la digestión de las regiones bicatenarias de moléculas de ssADN lineal que resultan de la fusión intramolecular o intermolecular. En algunas formas de realización preferidas, el método adicionalmente comprende tratar la biblioteca de fragmentos de ssADN circulares rotulados con T5 exonucleasa (EPICENTRE Biotechnologies, Madison, WI) para remover el ssADN y el ssADN no ligados (por ejemplo, los fragmentos de ADNs que están "pinzados" y/o contienen regiones monocatenarias).

20 En algunas formas de realización de la divulgación, el método comprende además amplificar los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados por transcripción, comprendiendo el método : (a) fusionar la secuencia de promotor de sentido a un oligodesoxirribonucleótido que presenta una secuencia de promotor antisentido, o fusionar los fragmentos de ssADN circulares rotulados o a los fragmentos de ssADN lineales dirrotulados, un cebador que es complementario con respecto a los mismos y extender el cebador con una ADN polimerasa bajo condiciones en las que se sintetiza un promotor polimerasa bicatenaria; y (b) incubar los productos de ssADN con una ARN polimerasa que liga el promotor ARN polimerasa bajo condiciones en las que se sintetiza ARN.

30 En algunas formas de realización preferidas en las que la cadena transferida o un cebador de PCR presenta una secuencia promotora de polimerasa de ARN, el promotor ARN polimerasa es un promotor ARN polimerasa de tipo T7 y el método comprende además la etapa de transcribir los fragmentos de ssADN circulares rotulados in vitro mediante la utilización de una ARN polimerasa de tipo T7 que reconoce el promotor. Lo más preferible es que la ARN polimerasa y el promotor sean elegidos de entre T7 RNAP, T3 RNAP y SP6 RNAP y los correspondientes promotores cognatos. Sin embargo, las etapas de transcripción de un método de la divulgación pueden utilizar cualquier ARNP para el exista o pueda obtenerse una adecuada secuencia de promotor que permita la transcripción con un elevado carácter específico. Los kits y enzimas para la transcripción in vivo pueden obtenerse en comercio de muchos vendedores y la mezcla adecuada para llevar a cabo las etapas de la presente divulgación que comprenden la transcripción in Vitro pueden utilizar aquellos productos descritos por los fabricantes. Por ejemplo, la transcripción in Vitro mediante la utilización de T7 RNAP puede ser llevado a cabo mediante el kit AMPLISCRIBE™ T7-FLASH™ Transcription Kit o el kit AMPLISCRIBE™ T7 High Yield Transcription Kit de EPICENTRE Biotechnologies, Madison, WI como se describe en la literatura del producto. De manera similar, si se utiliza T3 RNAP o SP6 RNAP en un método de la divulgación para la transcripción in vitro, puede utilizarse un kit AMPLISCRIBE™ T3-FLASH™ High Yield Transcription Kit o el kit AMPLISCRIBE™ SP6 High Yield Transcription Kit (EPICENTRE Biotechnologies, Madison, WI), respectivamente, como se describe.

45 En algunas formas de realización, la cadena transferida, el oligonucleótido rotulador de ligación, o un cebador de PCR presenta, además de la secuencia del promotor de ARN polimerasa, secuencias adicionales para la traducción, tal como pero sin limitación, un sitio de ligación de ribosoma y un codón de iniciación de traducción (que en la presente también recibe la denominación de "señal de iniciación de la traducción"), y el método adicionalmente comprende traducir el ARN transcrito. En algunas de estas formas de realización, el método comprende además la etapa de la traducción in vitro de los transcritos de ARN resultantes. Los sistemas y kits para la traducción in vitro de los transcritos de ARN también están disponibles en el comercio de muchas fuentes y pueden utilizarse para la presente divulgación. Por ejemplo, los sistemas de lisato de reticulocitos de conejo, extracto de germen de trigo, y extracto de E. coli S30 de Promega Corporación, Madison, WI pueden utilizarse para la presente divulgación. Más aun, kits para la transcripción in vitro acoplada y para la traducción in Vitro también están disponibles en el comercio y pueden utilizarse, tales los sistemas TNT® Quick Coupled Transcription/ Translation de Promega.

55 En algunas otras formas de realización, el método comprende además la etapa de amplificar y/o secuenciar el ADN objetivo in los fragmentos de ssADN circulares rotulados mediante la utilización de una ADN polimerasa y por lo menos una primer que es complementario al rótulo. En algunas formas de realización, la etapa de amplificar los fragmentos de ssADN circulares rotulados mediante ADN polimerasa comprende la replicación por círculo de rodadura. En algunas formas de realización, la etapa de amplificar los fragmentos de ssADN circulares rotulados mediante la utilización de una ADN polimerasa comprende PCR amplificación mediante la utilización de una ADN polimerasa termoestable y un

primer cebador de PCR que es complementario a por lo menos una porción del rótulo y un segundo cebador de PCR que es complementario con respecto a por lo menos una porción del complemento del rótulo. En algunas formas de realización, el método comprende además la etapa de amplificar los fragmentos de ssADN circulares rotulados mediante la utilización de una ARN polimerasa.

- 5 Por lo tanto, en algunas otras formas de realización, el método comprende además amplificar los fragmentos de ssADN circulares rotulados mediante RCR (rolling circle replication, replicación de círculo de rodadura), comprendiendo el método: (a) fusionar un cebador que es complementario con respecto a los fragmentos de ssADN circulares rotulados; y (b) extender el primer fusionado a los fragmentos de ssADN circulares rotulados mediante la utilización de una ADN polimerasa que desplaza cadena (por ejemplo, phi29 ADN polimerasa o Bst ADN, gran fragmento de polimerasa (EPICENTRE) o DISPLACEACE™ ADN polimerasa (EPICENTRE). En estas formas de realización, los productos de amplificación de RCR son moléculas de ssADN concatamérico que son complementarias con respecto a los fragmentos de ssADN circulares rotulados. En algunas formas de realización en donde los fragmentos de ssADN circulares rotulados presentan una secuencia de promotor de antisentido, los productos de amplificación concataméricos por RCR presentan una secuencia de promotor de sentido y el método comprende además preparar el promotor ARN polimerasa bicatenario (por ejemplo, mediante fusión a la secuencia de promotor de sentido de un oligodesoxirribonucleótido complementario que presenta una secuencia de promotor de antisentido, después de lo cual se transcribe el ADN concatamérico mediante la utilización de una ARN polimerasa que se liga al promotor polimerasa bicatenaria e inicia la transcripción a partir del mismo.

- En algunas formas de realización de la divulgación, la composición de extremo del transposón comprende una cadena transferida que presenta solamente la secuencia extrema de transposón transferida y, por ello, el rótulo presenta solamente la secuencia extrema de transposón transferida. En algunas formas de realización de la invención, la composición de extremo del transposón comprende una cadena transferida que comprende o consiste en una porción 3' y a porción 5', en donde la porción 3' presenta la secuencia extrema de transposón transferida y la porción 5' presenta cualquier otra secuencia deseada, en las cuales formas de realización el rótulo comprende o consiste en tanto la porción 3' como la porción 5'. En algunas formas de realización en donde la composición de extremo del transposón comprende una cadena transferida que comprende o consiste en una porción 3' y una porción 5', la cadena no transferida presenta una secuencia que se complementa con la porción 5' de la cadena transferida. Sin embargo, en algunas formas de realización preferidas de la composición de extremo del transposón, la cadena no transferida no presenta una secuencia que se complementa con la porción 5' de la cadena transferida. En algunas formas de realización preferidas, la cadena no transferida presenta solamente la secuencia extrema de transposón no transferida. En algunas formas de realización preferidas, la cadena no transferida presenta una secuencia que no es complementaria con respecto a la cadena transferida de la secuencia de transposón de extremo 3' no transferida.

- En algunas formas de realización de cualquiera de los métodos en donde la composición de extremo del transposón comprende una cadena transferida que comprende o consiste en una porción 5' y una porción 3', la porción 5' presenta la secuencia de un dominio de rótulo de secuenciación o de un dominio de rótulo de captura (por ejemplo, un dominio de rótulo de secuenciación o un dominio de rótulo de captura para el Roche 454 Genome Sequencer FLX System, por ejemplo, ya que los rótulos Roche 454A y 454B se utilizan para secuenciación mediante la utilización del sistema Roche 454 Genome Sequencer FLX System) y la porción 3' presenta la secuencia extrema de transposón transferida. Por lo tanto, cuando la composición de extremo del transposón comprende una cadena transferida que tiene un dominio de rótulo de secuenciación o un dominio de rótulo de captura, los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados tienen un rótulo que comprende el dominio de rótulo de secuenciación o el dominio de rótulo de captura (por ejemplo, el rótulo Roche 454A o 454B utilizado para la secuenciación mediante la utilización del sistema Roche 454 Genome Sequencer FLX System). Los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados generados que tengan el intervalo de tamaños deseado se utilizan como plantillas para la secuenciación de próxima generación mediante la utilización del sistema Roche 454 Genome Sequencer FLX System. En otras formas de realización, los fragmentos de ssADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados se generan que comprenden uno o más dominios de sitios de restricción, dominios de rótulo de secuenciación, dominios de rótulo de captura, dominios de rótulo de amplificación, dominios de rótulo de detección y/o dominios de dirección de rótulo destinado a ser utilizados en la secuenciación (por ejemplo, mediante la utilización de la plataforma de secuenciación ROCHE 454, la plataforma de secuenciación ILLUMINA™ SOLEXA™, la plataforma de secuenciación SOLID™ de LIFE TECHNOLOGIES/APPLIED BIOSYSTEMS, la plataforma de secuenciación SMRT™ de PACIFIC BIOSCIENCES, la plataforma de secuenciación POLLONATOR Polony, la plataforma de secuenciación COMPLETE GENOMICS, la plataforma de secuenciación INTELLIGENT BIOSYSTEMS, o la plataforma de secuenciación HELICOS).

- 55 No hay un límite para la utilización de secuencias adicionales en la porción 5' de la cadena transferida o en la porción 3' de la cadena no transferida, las cuales secuencias pueden utilizarse para llevar a cabo cualquier finalidad deseada. En algunas formas de realización, la porción 5' de la cadena transferida o la 3' porción de la cadena no transferida presenta una o más secuencias de dominios de rótulo.

- En algunas formas de realización, el método comprende además las etapas de extender las cadenas transferidas que comprenden los fragmentos de ADN rotulados en 5' generados en la reacción de transposición mediante la utilización de una ADN polimerasa que carece de actividad de desplazamiento de cadena y de actividad de exonucleasa 5' a -3' (por ejemplo, T4 ADN polimerasa, EPICENTRE), y seguidamente mediante la utilización de una ADN ligasa

dependiente de plantilla (por ejemplo, E. coli DNA ligasa) para ligar el extremo 3' de cada producto de extensión de ADN al extremo 5' de una cadena no transferida que comprende el fragmento rotulado de DNAs mediante la utilización de la cadena opuesta como una plantilla de ligación; en estas formas de realización, los extremos 5' de las cadenas no transferidas de la composición del extremo tienen un grupo 5'-monofosfato. Esta forma de realización del método genera fragmentos de ssADN dirrotulados.

Los trabajos llevados a cabo durante el desarrollo de formas de realización de la presente invención llevaron a la observación que la transposición tiene lugar en ssADN. Por ello, en algunas formas de realización preferidas, la composición de extremo del transposón comprende o consiste en una cadena no transferida que presenta solamente la secuencia no extrema de transposón transferido de manera tal que la porción 5' de la cadena transferida es monocatenaria (por ejemplo, a efectos de minimizar la probabilidad o frecuencia de inserción de la cadena transferida en porciones bicatenarias de sí misma durante la reacción de transposición in vitro). En algunas formas de realización preferidas, en las que la cadena no transferida presenta una porción 3' que es complementaria con respecto a la porción 5' de la cadena transferida, el tamaño de la porción 5' de la cadena transferida se minimiza a efectos de minimizar la probabilidad o frecuencia de inserción de la cadena transferida en sí misma durante la reacción de transposición in vitro. Por ejemplo, en algunas formas de realización, el tamaño o magnitud de la porción 5' de la cadena transferida (y de la porción 3' complementaria no transferida) es inferior a aproximadamente 150 nucleótidos, inferior a aproximadamente 100 nucleótidos, inferior a aproximadamente 75 nucleótidos, inferior a aproximadamente 50 nucleótidos, inferior a aproximadamente 15 nucleótidos, o inferior a aproximadamente 15 nucleótidos.

En algunas formas de realización preferidas, el extremo 5' de la cadena transferida de la composición de extremo del transposón tiene un grupo 5'-monofosfato. En algunas formas de realización preferidas, el extremo 5' de la cadena no transferida tiene un grupo 5'-monofosfato. En algunas formas de realización preferidas, tanto la cadena transferida como la cadena no transferida tienen un grupo 5'-monofosfato. En formas de realización en las que la cadena transferida no tiene un grupo 5'-monofosfato, el método comprende además la etapa de fosforilar el extremo 5' del nucleótido extremo de transposón transferido (por ejemplo, mediante la utilización de polinucleótido quinasa; por ejemplo, T4 polinucleótido quinasa) antes de la etapa de ligación del método.

La inserción catalizada por transposasa, del extremo de transposón en el ADN objetivo tiene como resultado la unión del extremo de transposón transferido al extremo 5' de una cadena del ADN objetivo y la rotura o fragmentación de dicha cadena en el sitio en donde la secuencia extrema de transposón transferido se une al ADN objetivo, con la generación concomitante de una región de base 9 de ADN objetivo monocatenario situado en 3' con respecto al sitio de la unión del extremo de transposón transferido al ADN objetivo debido a la región de brecha de base 9 en la cadena opuesta del ADN objetivo. Por ejemplo, la FIG. 1 muestra los resultados de dos eventos de inserción independientes del extremo de transposón transferido en cadenas opuestas del ADN objetivo. Como se muestra en in FIG. 2, a veces tienen lugar inserciones independientes del extremo de transposón transferido en cadenas opuestas del ADN objetivo en lugares en el ADN objetivo que se hallan en una proximidad relativamente estrecha, generándose dos fragmentos de ADN rotulados en 5'. Al tener lugar la desnaturalización, se liberan dos fragmentos de ssADN rotulados en 5'.

Ligación independiente de templado, de fragmentos de ssADN rotulados en 5'

En algunas formas de realización, una ligasa de ácido nucleico independiente de plantilla o no homóloga (por ejemplo, que lleva a cabo una ligación intramolecular, una circularización de ssADN que tiene un grupo 3'-hidroxilo y un grupo 5'-monofosfato) se utiliza en un método de la invención para circularizar fragmentos de ssADN lineales rotulados en 5'. En algunas formas de realización preferidas, la ligasa de ácido nucleico es una ARN ligasa termoestable (por ejemplo, seleccionada entre ARN ligasa de TS2126 bacteriófago termoestable (patente de los Estados Unidos No. 7,303,901 y Blondal et al., Nucleic Acids Res 33: 135-142, 2005), CIRCLIGASE™ ssADN ligasa (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos), y una ARN ligasa (por ejemplo, Methanobacterium thermoautotrophicum ARN ligasa 1 o "MthRnl"; Torchia, C et al., Nucleic Acids Res. 36:6218-6227, 2008). La expresión "ligasa independiente de plantilla" o "ligasa no homóloga" da a entender que una ligasa que tiene como resultado la ligación de ssADN en la ausencia de fusión de una secuencia complementaria a los extremos del ssADN que han de ser unidos o ligados (es decir, los dos extremos no son fusionados a una secuencia complementaria a efectos de mantenerlos adyacentes entre sí durante la etapa de ligación). En estas formas de realización, el método comprende la etapa de: desnaturalizar los fragmentos de ADN rotulados en 5' fusionados generados en la reacción de transposición in vitro mediante la incubación de los fragmentos de ssADN lineales rotulados en 5' con la ligasa de ácido nucleico. Mediante la expresión "ligación intramolecular" se da a entender que los dos extremos de una molécula de ssADN se ligan entre sí de manera de generar fragmentos de ssADN circulares, en lugar de ser ligados a los extremos de otras moléculas de ADN.

En algunas formas de realización preferidas, la reacción de ligación no homóloga se lleva a cabo en una "mezcla de reacción de ligación mejorada" lo que en la presente significa una mezcla de reacción de ligación que comprende: a) los fragmentos de ssADN lineales rotulados; b) un tampón que mantiene el pH; b) cationes  $Mn^{2+}$ ; y (c) una composición de moléculas de ARN termoestables en donde una elevada proporción de las moléculas de ARN ligasa termoestable están adeniladas; en donde la concentración de las moléculas de ARN ligasa termoestable RNA es por lo menos igual a la molaridad de los fragmentos de ssADN lineales, y en donde no se añaden ATP ni cationes  $Mg^{2+}$  a la mezcla de reacción de la ligación.

Mediante la expresión que “una elevada proporción de las moléculas de ARN ligasa termoestables están adeniladas”, se da a entender que por lo menos aproximadamente el 50% de la totalidad de las moléculas de ARN ligasa termoestables presentes en la mezcla de reacción mejoradas de la ligación están adeniladas. En algunas formas de realización de la mezcla mejorada de reacción de ligación, más de aproximadamente el 60% de la totalidad de las termoestable RNA ligasa moléculas están adeniladas. En algunas formas de realización de la mezcla mejorada de reacción de ligación, más de aproximadamente el 70% de la totalidad de moléculas de ARN ligasa termoestables están adeniladas. En algunas formas de realización de la mezcla mejorada de reacción de ligación, más de aproximadamente el 80% de la totalidad de las moléculas de ARN ligasa termoestables están adeniladas. En algunas formas de realización preferidas de la mezcla mejorada de reacción de ligación, más de aproximadamente el 90% de la totalidad de las moléculas de ARN ligasa termoestables están adeniladas. En algunas formas de realización preferidas de la mezcla mejorada de reacción de ligación, más de aproximadamente el 95% de la totalidad de las moléculas de ARN termoestables están adeniladas. En algunas formas de realización preferidas, las moléculas de ARN ligasa termoestables se adenilan para preparar una composición en donde una elevada proporción de las moléculas de ARN ligasa termoestables se adenilan mediante la incubación de la enzima con ATP durante o después del proceso de purificación. Por ejemplo, un protocolo que puede utilizarse para adenilar la ARN ligasa termoestable consiste en incubar la enzima en una solución que contiene Tris-HCl 50 mM, pH 8,0, MgCl<sub>2</sub> 2 mM, NaCl 100 mM y ATP 0,5 mM durante 15 minutos a 50 °C; seguidamente detener la reacción mediante la adición de EDTA hasta una concentración final de 5 mM; y a continuación retirar los componentes de la reacción mediante diálisis o filtración en gel. El porcentaje de ARN ligasa termoestable adenilado puede estimarse mediante análisis y SDS-PAGE. En algunas formas de realización preferidas, la ARN ligasa termoestable en donde una elevada proporción de las moléculas de ARN ligasa termoestables están adeniladas es la ARN ligasa termoestable de TS2126 bacteriófago. En algunas formas de realización de la mezcla mejorada de reacción de ligación, el tampón mantiene el pH entre 6,5 y 8,0. En algunas formas de realización preferidas de la mezcla mejorada de reacción de ligación, el tampón mantiene el pH entre 7,0 y 8,0. En algunas formas de realización preferidas de la mezcla mejorada de reacción de ligación, el tampón que mantiene el pH entre pH 7.0 y 8.0 es un tampón Tris tampón. En algunas formas de realización de la mezcla de reacción de ligación mejorada, la concentración de los cationes Mn<sup>2+</sup> es de entre 0,5 y 10 mM. En algunas formas de realización de la mezcla mejorada de reacción de ligación, la concentración de los cationes Mn<sup>2+</sup> es de entre 1 y 10 mM. En algunas formas de realización de la mezcla mejorada de reacción de ligación, la concentración de cationes Mn<sup>2+</sup> es de entre 1 y 5 mM. En algunas formas de realización preferidas de la mezcla mejorada de reacción de ligación, la concentración de cationes Mn<sup>2+</sup> es de 2,5 mM. En algunas formas de realización preferidas de la mezcla de reacción mejorada de la ligación, los cationes Mn<sup>2+</sup> se proveen en forma de MnCl<sub>2</sub>. En algunas formas de realización, la concentración de las moléculas de ARN termoestables adeniladas en la mezcla mejorada de reacción de ligación es de por lo menos dos veces la molaridad de los fragmentos de ssADN lineales. En algunas formas de realización, la concentración de las moléculas de ARN ligasa termoestables adeniladas en la mezcla mejorada de reacción de ligación es de por lo menos el quintuplo de la molaridad de los fragmentos de ssADN lineales. En algunas formas de realización, la concentración de las moléculas de ssADN termoestables adeniladas ligasa en la mezcla mejorada de reacción de ligación es de por lo menos el decuplo de los fragmentos de ADN lineales. En algunas formas de realización preferidas, la mezcla mejorada de reacción de ligación adicionalmente comprende una sal tal como cloruro de potasio o acetato de potasio (por ejemplo, con una concentración de aproximadamente a aproximadamente 100 mM). En algunas formas de realización preferidas, la mezcla mejorada de reacción de ligación adicionalmente comprende un reactivo de reducción tal como ditiotreitól (DTT) (por ejemplo, con una concentración de aproximadamente 0,5 o 1 mM). En algunas formas de realización, la mezcla mejorada de reacción de ligación adicionalmente comprende trimetil glicina (botaina) zwitteriónico con una concentración de 0,25 y 5,2 M. En algunas formas de realización, la mezcla mejorada de reacción de ligación comprende adicionalmente trimetil glicina (betaína) zwitteriónico con una concentración de entre 0,5 y 2 M. En algunas formas de realización, la mezcla mejorada de reacción de ligación comprende adicionalmente trimetilglicina (betaína) zwitteriónico con una concentración de aproximadamente 1 M.

En algunas formas de realización preferidas, la mezcla mejorada de reacción de ligación comprende: a) los fragmentos de ADN lineales que tienen grupos 5'-fosforilo y 3'-hidroxilo (por ejemplo, 0,5 micromolar); b) TRIS acetato 33 mM con un pH 7,8; c) 2,5 mM de cationes Mn<sup>2+</sup>; y (c) una composición de moléculas de ARN ligasa termoestables en donde >70% de las moléculas de ARN termoestables están adeniladas; en donde la concentración de las moléculas de ARN ligasa termoestables adeniladas es por lo menos igual a la molaridad de los fragmentos de ADN lineales (por ejemplo, aproximadamente 1 micromolar de ARN ligasa termoestable adenilado para 0,5 micromolar de los fragmentos de ssADN lineales rotulados), y en donde no se añade ATP ni cationes Mg<sup>2+</sup> a la mezcla de reacción de la ligación. En algunas formas de realización preferidas, la concentración de las moléculas de ARN termoestable adenilado es de por lo menos 5 veces, de por lo menos 10 veces, o de por lo menos 20 veces la molaridad de los fragmentos de ssADN lineales (por ejemplo, aproximadamente 2,5 micromolar, aproximadamente 5 micromolar, o aproximadamente 10 micromolar de ARN ligasa termoestable adenilado para 0,5 micromolar de los fragmentos de ADN lineales rotulados). En algunas formas de realización preferidas, la mezcla mejorada de reacción de ligación comprende adicionalmente acetato de potasio y DTT 0,5 mM. En algunas formas de realización preferidas, la mezcla mejorada de reacción de ligación comprende adicionalmente betaína 1 M.

En algunas formas de realización preferidas del método, la ligación intramolecular de fragmentos de ADN lineales rotulados en 5' para sintetizar fragmentos de ssADN circulares rotulados se lleva a cabo en la mezcla de reacción de la ligación a una temperatura entre aproximadamente 40 °C y aproximadamente 70 °C durante un tiempo suficiente

(por ejemplo, de aproximadamente 1 hora a aproximadamente 72 horas), en donde se sintetizan fragmentos de ssADN circulares rotulados. En algunas formas de realización preferidas, la ligación intramolecular se lleva a cabo a una temperatura de reacción de aproximadamente 60 °C durante un tiempo suficiente en donde se sintetizan fragmentos de ssADN circulares.

5 La invención no se limita a solamente la ligasa de ácido nucleico particular descrita en la presente. Las personas expertas en el arte comprenderán que la ligación intramolecular puede llevarse a cabo mediante la utilización de cualquier ligasa de ácido nucleico que tenga una actividad similar a las descritas en la presente, lo que significa que tiene como resultado una ligación intramolecular no homóloga de ssADN que tienen un grupo 3'-hidroxilo y un grupo 5'-monofosfato.

10 V. Fragmentación y rotulación de ADN-ds mediante transposición in vitro de extremos de transposón en horquilla (30963)

En pocas palabras, en algunas formas de realización, el método comprende: incubar el ADN objetivo, que es dsADN, con una transposasa y una composición de extremo del transposón en horquilla en una reacción de transposición in vitro para al mismo tiempo fragmentar y rotular el ADN objetivo, con lo que se genera una biblioteca que comprende una población de fragmentos de ADN 5-rotulado; después de lo cual se une el extremo 3' de cada fragmento de ADN 5' rotulado, que comprende una porción de una cadena de ADN objetivo al extremo 5' de otro fragmento de ADN rotulado en 5' que comprende una porción complementaria (es decir, la cadena opuesta del ADN objetivo), con lo que se genera una biblioteca de fragmentos de ADN circulares rotulados cerrados de manera covalente (por ejemplo, que presentan estructuras de forma circular o acampanada monocatenarias). En algunas formas de realización preferidas, la etapa de unir comprende: extender los extremos 3' de los fragmentos de ADN rotulados en 5' con una ADN polimerasa que carece de exonucleasa 5' a 3' (lo que incluye nucleasa 5' dependiente de la estructura) y actividades de desplazamiento de cadena de modo de generar productos de extensión rotulados en 5' de ADN y ligar el extremo 3' de cada extremo de dichos productos de extensión de fragmentos de ADN al extremo 5' del producto de extensión de fragmento complementario 5' rotulado mediante la utilización de una ligasa dependiente de plantilla (por ejemplo, E. coli DNA ligasa o una ligasa dependiente de plantilla de una bacteria siicrófila o un bacteriófago sicrófilo). En otras formas de realización preferidas, la etapa de unir comprende: incubar oligodesoxirribonucleótidos de secuencia aleatoria (por ejemplo, oligodesoxirribonucleótidos de secuencia aleatoria o semialeatoria que están 5'-monofosforilados o 5'-adenilados) que son de uno o más tamaños adecuados para llenar exactamente las brechas monocatenarias que resultan de la reacción de transposición in vitro (por ejemplo, oligodesoxirribonucleótidos 5'-monofosforilados de secuencia aleatoria que comprenden o que consiste en 9-mer de secuencia aleatoria o semialeatoria o en un 4-mer de secuencia aleatoria y un 5-mero de secuencia aleatoria para rellenar las brechas monocatenarias que resultan de la transposición in vitro mediante la utilización de EZ-Tn5™ transposasa) y una ADN ligasa dependiente de plantilla (por ejemplo, ADN ligasa de E. coli o una ADN ligasa de una bacteria sicrófila o de un bacteriófago sicrófilo) con los fragmentos de ADN rotulados en 5' bajo condiciones y durante un tiempo suficiente en donde los oligonucleótidos de secuencia aleatoria se fusionan de manera de llenar las brechas monocatenarias en los fragmentos de ADN 5' rotulados y se ligan, con lo que se genera una población de moléculas de ADN circulares rotulados.

Por lo tanto, una forma de realización preferida de la invención es un método para generar una biblioteca que comprende una población de fragmentos de ADN circulares rotulados de ADN objetivo bicatenario destinado a ser utilizado como plantillas en la secuenciación de ADN o en reacciones de amplificación de ácidos nucleicos, en donde cada uno dichos fragmentos de ADN circulares rotulados presenta las secuencias de ambas cadenas de una porción del ADN objetivo y la secuencia del rótulo, que comprende:

Proporcionar:

45 1. ADN objetivo que comprenden o que consiste en una o más moléculas bicatenarias (dsADN) (por ejemplo, dsADN genómico, mitocondrial, cloroplasto u otro dsADN de una célula eucariota y/o ADN genómico o episomal ADN de una célula procariota, o ADNc bicatenario preparado mediante transcripción inversa de ARN de una célula eucariota o procariota para generar ADNc de primera cadena y seguidamente se extiende un cebador fusionado al ADNc de primera cadena);

50 2. una transposasa (por ejemplo, una transposasa de tipo salvaje o mutante; por ejemplo, Tn5 transposasa de tipo salvaje o mutante, por ejemplo, EZ-Tn5™ transposasa, por ejemplo, HYPERMU™ MuA transposasa, EPICENTRE Biotechnologies, Madison, WI, Estados Unidos); y

55 3. una composición de extremo del transposón en horquilla que es capaz de formar un complejo funcional con la transposasa en una reacción de transposición y que presenta la secuencia del rótulo, en donde dicha composición de extremo del transposón en horquilla comprende o consiste en un oligonucleótido que contiene 5'-fosfato que presenta un extremo no transferido de secuencia de transposón en su extremo 5' (por ejemplo, que en la presente recibe la denominación de "MENTS" con respecto a la secuencia extrema de transposón no transferida EZ-Tn5™), una secuencia extrema de transposón transferida en su extremo 3' (por ejemplo, que en la presente recibe la denominación de "METS" con respecto a la secuencia extrema de transposón transferido EZ-Tn5™), y una secuencia interviniente (por ejemplo, para cualquiera finalidad que se desee, tal como la de proporcionar un rótulo) entre el extremo no

- transferido de secuencia de transposón y la secuencia extrema de transposón transferido que es lo suficientemente larga para permitir la formación de un tallo-bucle intramolecular, en donde el tallo presenta las secuencias del extremo de transposón bicatenario con las que la transposasa forma un complejo que es funcional para la transposición (por ejemplo, en donde el tallo presenta las secuencias del extremo 19-bp OE (*outer end*) de transposón, el extremo 19-bp E (*inner end*) de transposón, o el 19-bp ME (*mosaic end*) de transposón reconocido por una transposasa de tipo salvaje o mutante, por ejemplo, por EZ-Tn5<sup>TM</sup> transposasa) (o, por ejemplo, los extremos R1 y R2 MuA transposón para la MuA transposasa) y el bucle presenta la secuencia interviniente, que puede ser cualquier secuencia arbitraria;
- 5 4. (a) una ADN polimerasa que carece de 5' nucleasa (lo que incluye exonucleasa 5' a 3' y actividad de 5' nucleasa dependiente de la estructura) y actividades de desplazamiento de cadenas (por ejemplo, T4 ADN polimerasa); o (b)
- 10 uno o más tamaños de oligonucleótidos de secuencia aleatoria, que por sí solos, o en combinación, tienen la misma longitud que las brechas monocatenarias en los fragmentos de ADN rotulados en 5' que resultan después de una reacción de transposición con la transposasa y la composición extrema del transposón en horquilla; y
5. una ligasa dependiente de plantilla (por ejemplo, E. coli DNA ligasa o una ligasa dependiente de plantilla procedente de una bacteria sicrofílica o de bacteriófago sicrofílo);
- 15 incubar el ADN objetivo en una reacción de transposición in vitro con la transposasa y la composición de extremo del transposón en horquilla bajo condiciones y durante un tiempo suficiente en donde la inserción de la composición de 5' extremo del transposón en horquilla en el ADN objetivo genera una población de fragmentos de ADN rotulados en 5' (véase, por ejemplo, FIG 2 y FIG 3);
- 20 incubar los fragmentos de ADN rotulados en 5' bajo condiciones y durante un tiempo suficiente en donde se llenan las brechas monocatenarias en los fragmentos de ADN y el extremo 3' de cada fragmento de ADN rotulado en 5' experimenta una extensión y unido al extremo 5' de otro fragmento de ADN rotulado en 5' que comprende una porción complementaria ADN objetivo, con lo que se genera una biblioteca de fragmentos de ADN circulares rotulados, cada uno de cuales presenta las secuencias de ambas cadenas de una porción del ADN objetivo y la secuencia del rótulo.
- 25 En algunas formas de realización preferidas del método (como se diagrama, por ejemplo, en las FIGS. 7 y 8), la etapa de unir comprende: (1) incubar los fragmentos de ADN rotulados en 5' con la ADN polimerasa que carece de actividad de nucleasa bajo condiciones en las que el extremo 3' de fragmento de ADN rotulado en 5' experimenta una extensión para generar una población de productos de extensión de fragmentos de ADN rotulado 5'; e (2) incubar los productos de extensión de fragmentos de ADN rotulados en 5' con la ligasa dependiente de plantilla bajo condiciones y durante un tiempo suficiente en donde los productos de extensión de fragmento de ADN rotulado en 5' son ligados, con lo que
- 30 se genera la biblioteca de fragmentos de ADN circulares rotulados. En algunas formas de realización, la ADN polimerasa que carece de actividades de 5' nucleasa y de desplazamiento de cadenas, y la ligasa dependiente de plantilla se proveen en una mezcla y la etapa de la unión se lleva a cabo en una única mezcla de reacción.
- 35 En algunas otras formas de realización preferidas del método, la etapa de unir comprende: incubar los fragmentos de ADN rotulados en 5' con los uno o más tamaños de los oligonucleótidos de secuencia aleatoria y la ligasa dependiente de plantilla bajo condiciones y durante un tiempo suficiente en donde los oligonucleótidos de secuencia aleatoria se fusionan de manera de llenar las regiones de brechas monocatenarias en los fragmentos de ADN rotulados en 5' y en donde dichos oligonucleótidos de secuencia aleatoria fusionados son ligados o entre sí o a un extremo adyacente de los fragmentos de ADN rotulados en 5', con lo que se genera los fragmentos de ADN circulares rotulados.
- 40 En algunas formas de realización, el método comprende adicionalmente, después de la etapa de ligación, una o más etapas para remover los oligonucleótidos de secuencia aleatoria, ADN objetivo lineal y/o la composición de extremo del transposón en horquillas que no se unen al ADN objetivo.
- 45 En algunas formas de realización preferidas, el método comprende adicionalmente: tratar la mezcla de reacción que contiene los fragmentos de ADN circulares rotulados con T5 exonucleasa (EPICENTRE Biotechnologies, Madison, WI) para remover el ssADN lineal no ligado y el dsADN (por ejemplo, fragmentos de ADN que están nicked y/o contiene regiones monocatenario).
- 50 En algunas formas de realización el método comprende adicionalmente: desdoblar los fragmentos de ADN circulares rotulados en cada una de las estructuras de bucle para generar fragmentos lineales bicatenarias de ADNs, teniendo cada cadena de ADNs una porción del rótulo en su extremo 5' y una porción del rótulo en su extremo 3' (los cuales fragmentos lineales de ADN en la presente reciben la denominación de "fragmentos de dsADN lineales dirrotulados abanico" o "fragmentos de dsADN en abanico").
- 55 En algunas formas de realización, el método desdoblar comprende: fusionar a los fragmentos de ADN circulares rotulados un oligonucleótido que se fusiona a un sitio de restricción dentro del rótulo, y seguidamente incubar con la endonucleasa de restricción que se desdobla en el sitio bicatenario de restricción para generar los fragmentos de ADNsds de abanico.
- En algunas otras formas de realización, la composición de extremo del transposón en horquilla tiene uno o más sitios clivables (por ejemplo, en la estructura de bucle) que son clivables mediante la utilización de una composición de enzima de clivaje (por ejemplo, un sitio clivable que consiste en un residuo dUMP que es clivable mediante la utilización

de una composición de enzima de clivaje que comprende uracil-N-glicosilasa y una endonucleasa, tal como endonucleasa III o endonucleasa IV de *E. coli*; o, por ejemplo, un sitio clivable que consiste en un residuo 8-oxo-guanina-2'-desoxirribonucleósido-monofosfato que es clivable mediante la utilización de una composición de enzima de clivaje que comprende FPG proteína + una AP endonucleasa, tal como endonucleasa III o endonucleasa IV de *E. coli*), y el método de clivaje comprende: incubar los fragmentos de ADN circulares rotulados con la composición de enzima de clivaje bajo condiciones y durante un tiempo suficiente en donde los fragmentos de ADN circulares rotulados son desdoblados en los sitios clivables de manera de generar los fragmentos ADNsa de cola de milano. En algunas formas de realización del método, se utiliza un nucleótido no canónico diferente para proporcionar el sitio clivable y se utiliza una N-glicosilasa diferente en la composición de enzima de clivaje. La composición de extremo del transposón en horquilla es sintetizada (por ejemplo, mediante la utilización de un sintetizador de oligonucleótidos) para contener un sitio clivable que consiste en un nucleótido no canónico en lugar del nucleótido canónico (por ejemplo, dUMP como el nucleótido no canónico en lugar de TMP cuando se utiliza uracilN-glicosilasa como un componente de la composición de enzima de clivaje, o, por ejemplo, 8oxo-GMP como el nucleótido no canónico en lugar de GMP cuando se utiliza la proteína FPG como un componente de enzima de clivaje) en el o los sitios en los que se desea desdoblarse los fragmentos de ADN circulares rotulados en la biblioteca (por ejemplo, en donde el sitio en el que se desea desdoblarse los fragmentos de ADN circulares rotulados dentro de la estructura de bucle de la composición de extremo del transposón en horquilla que se inserta en los fragmentos de ADN circulares rotulados).

Por lo tanto, en algunas formas de realización preferidas en las que la composición de extremo del transposón tiene uno o más sitios clivables, la composición de enzima de clivaje utiliza una N-glicosilasa (o "ADN glicosilasa") para generar un sitio abásico o apirimidínico/apurínico (AP). Como se define en la presente, una "N-glicosilasa" es una enzima que cataliza la hidrólisis de la unión entre una base de ácido no canónico y un azúcar en ADN para generar un sitio abásico (AP). Tales enzimas se hallan presentes en muchas especies. Un ejemplo de *Escherichia coli* es el uracil N-glicosilasa (UNG), también denominado ADN glicosilasa (UDG). El UNG cataliza el clivaje de la base uracilo a partir de la desoxiribosa en ADN (Lindahl, *Prog. Nucl. Acid Res. Mol. Biol.* 22:135-192, 1979), pero no cataliza el clivaje de uracilo a partir de dUTP libre, deoxiuridina libre o ARN (Duncan, in *The Enzymes*, Boyer ed., pp. 565-586, 1981). Otros ejemplos de N-glicosilasas que pueden utilizarse como enzimas de clivaje se describen en Demple and Harison (*Annu. Rev. Biochem.* 63: 915-48, 1994) y en Duncan ("DNA Glycosylases," in *The Enzymes*, Boyer ed., p. 565-586, 1981). La expresión "N-glicosilasa" o "ADN-glicosilasa" se refiere a una enzima con actividad de N-glicosilasa, tanto si la enzima recibe la denominación formal de glicosilasa o tiene una actividad de glicosilasa combinada con otras actividades enzimáticas. Las glicosilasas reciben a veces el nombre de "glicosidasas," y con ello nos referimos al a definición de N-glicosilasa para abarcar las N-glicosidasas. Por ejemplo, la proteína FPG es también una N-glicosilasa como se define en la presente. La proteína FPG (formamidopirimidina ADN N-glicosilasa) es una enzima de reparación de escisiones que reconoce bases de ácidos nucleicos diversas pero estructuralmente relacionadas tales como 8-hidroxiguanina (que también recibe la denominación de 7-hidro-8-oxoguanina u 8-oxoguanina, que se refiere al tautómero favorecido 6,8-diceto bajo pH fisiológico) (Tchou, et al., *Proc. Natl. Acad. Sci. USA* 88: 4690-4694, 1991), derivados de anillo de imidazol abierto de adenina o guanina, designados 4,6- diamino-5-formamidopirimidina y 2,6-diamino-4-hidroxi-5-formamidopirimidina, respectivamente (Chetsanga, et al., *Biochemistry* 20: 5201-5207, 1981; y Breimer, *Nucl. Acids Res.* 12: 6359-6367, 1984), N.sub.7 - metilformamidopirimidinas, 5-hidroxiuracilo y 5-hidroxicitosina (Hatahet, et al., *J. Biol. Chem.* 269:18814-18820, 1994) y cataliza el clivaje del enlace N-glicosilo entre la base modificada y la columna vertebral desoxirribosa- fosfodiéster en el ADN, generándose un sitio AP. Además, la proteína FPG también posee una actividad de AP liasa. La actividad AP-liasa de la enzima cataliza las reacciones de beta, delta-eliminación, dejando una brecha de un solo nucleótido en el ADN (Bailly, et al., *Biochem. J.* 261: 707-713, 1989). La proteína FPG y el ADN de 8 hidroxiguanina han demostrado ser idénticos (Chung, MH et al., *Mutation Research* 254: 1-12, 1991). El tratamiento de ADN con azul de metileno más luz visible (Floyd, et al., *Arch Biochem Biophys* 273: 106-111, 1989) o con rosa bengali más luz ultravioleta (Friedmann y Brown, *Nucleic Acids Research* 5: 615-622, 1978) induce una modificación específica para guanina que es clivable mediante proteína FPG. Otras N-glicosilasas específicas se encuentran disponibles y son conocidos de la persona experta en el arte. Para determinar si una N-glicosilasa es adecuada o no para la presente invención, se empezaría por incorporar el nucleótido no canónico en el ADN y se determinaría si la base no canonical base puede ser específicamente removido por la N-glicosilasa candidato de una manera similar a la remoción de uracilo o de 8 oxo-guanina por la proteína UNG y FPG, respectivamente. Una vez que se ha creado un sitio abásico o AP, en la técnica se conocen varios métodos para desdoblarse el sitio abásico. Puede utilizarse el calor y/o condiciones básicas para romper la molécula de ADN en los sitios abásicos. Por ejemplo, podría utilizarse el siguiente protocolo: los ácidos nucleicos que contiene sitios abásicos (AP) después de la remoción de las bases no canónicas se calientan en una solución tampón que contiene una amina, por ejemplo, 25 mM Tris-HCl e iones de 1 a 5 mM, durante un período de 10 a 30 minutos de 70 °C a 90 °C. Como alternativa, puede utilizarse el siguiente tratamiento para romper el ADN en los sitios abásicos: piperidina < 1 M, una base, se añade a ADN que ha sido precipitado con etanol y secado al vacío. Seguidamente se calienta la solución durante 30 minutos a 90 °C y se trata por liofilización para remover la piperidina. En algunas formas de realización preferidas, se utiliza el tratamiento enzimático mediante la utilización de un endonucleasa apurínica/apirimidínica (AP endonucleasa) conocida en la técnica (Lindahl, *Prog. Nucl. Acid Res. Mol. Biol.* 22: 135-192, 1979; Demple and Harison *Annu. Rev. Biochem.* 63: 915-48, 1994) para romper el polímero de ADN en el sitio abásico. Tal como se define en la presente, una AP endonucleasa es cualquier enzima que catalice el clivaje de ADN en sitios abásicos (AP). Tales enzimas se hallan presentes en muchas especies. Los ejemplos de AP endonucleasas a partir de *E. coli* incluyen, sin limitación, endonucleasa III y endonucleasa IV. Asimismo, la exonucleasa III de *E. coli* en la presencia de iones calcio es una AP endonucleasa. Las enzimas útiles en la presente invención incluyen cualquier enzima con una actividad

similar a AP endonucleasa, tanto si lleva este nombre u otro.

En algunas formas de realización preferidas, el método comprende adicionalmente: desnaturalizar los fragmentos de dsADN de cola de milano para generar una biblioteca de fragmentos dirrotulados de ADN lineales (por ejemplo, destinados a ser utilizado como plantillas para el ADN o para la amplificación de ADN).

5 En algunas formas de realización los fragmentos de ADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados en la biblioteca generados mediante la utilización de los métodos se utilizan como plantillas de ADN en la amplificación de ácidos nucleicos y/o reacciones de secuenciación de ADN. En algunas formas de realización, el método comprende además la etapa de amplificar y/o secuenciar el ADN objetivo en los fragmentos de ADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados. En algunas formas de realización, el método comprende además la etapa de la secuenciación de ADN que es complementario con respecto al ADN objetivo obtenido mediante la amplificación de los fragmentos de ADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados. En algunas formas de realización, por lo menos una porción del ADN objetivo en cada uno de los fragmentos de ADN circulares rotulados o los fragmentos de ssADN lineales dirrotulados es secuenciado mediante la utilización de una ADN polimerasa y por lo menos un cebador que es complementario con respecto al rótulo (por ejemplo, para la secuenciación por síntesis). En algunas formas de realización, por lo menos una porción del ADN objetivo en cada uno de los fragmentos de ADN circulares rotulados o de los fragmentos de ADN lineales dirrotulados es secuenciado mediante la utilización de una ligasa dependiente de plantilla para ligar por lo menos un oligodesoxirribonucleótido que es complementario con respecto al rótulo y por lo menos otro oligodesoxirribonucleótido que se fusiona a la porción de la secuencia objetivo (por ejemplo, para la secuenciación por ligación). En algunas formas de realización, por lo menos una porción del ADN objetivo en cada uno de los fragmentos de ADN circulares rotulados o de los fragmentos de ADN lineales dirrotulados es secuenciado mediante la fusión de oligodesoxirribonucleótidos que se fusionan o hibridan con el rótulo y a una porción de la secuencia objetivo (por ejemplo, para la secuenciación por hibridación). En algunas formas de realización, el ADN que es complementario con respecto a los fragmentos de ADN circulares rotulados con respecto a los fragmentos de ssADN lineales dirrotulados es secuenciado mediante la utilización de la secuenciación por ligamiento o secuenciación por hibridación.

Por ejemplo, en algunas formas de realización preferidas, la secuencia extrema de transposón transferido presentado por la composición de extremo del transposón de horquilla que se provee en un kit o que se utiliza en un método de la presente invención es una secuencia extrema de transposón transferido reconocida por una Tn5 transposasa. En algunas formas de realización preferidas, la secuencia extrema de transposón transferida es una secuencia reconocida por EZ-Tn5<sup>TM</sup> transposasa (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos).

En general, los fragmentos de ADN circulares rotulados, los fragmentos dsADN cola de milano, y los fragmentos de ssADN lineales dirrotulados generados mediante la utilización de una composición de extremo del transposón en horquilla en los métodos de la presente invención presentan tanto la secuencia extrema de transposón transferido como la secuencia extrema de transposón no transferida, y secuencias adicionales que comprenden o derivan de la porción de bucle no complementaria de una porción de la composición de extremo del transposón en horquilla. Por lo tanto, en algunas formas de realización, la composición de extremo del transposón en horquilla presenta una más secuencias 5'- de la secuencia extrema de transposón transferido y 3' de la secuencia extrema de transposón no transferida, las cuales una o más secuencias nucleótidas adicionales son también presentados por el rótulo. Por lo tanto, además de las secuencias de extremos de transposón, el rótulo de la composición de extremo del transposón en horquilla puede tener una o más porciones de rótulo o dominios de rótulo.

En algunas formas de realización en donde la composición de extremo del transposón en horquilla comprende uno o más dominios de sitio de restricción, el método comprende además: fusionar un oligodesoxirribonucleótido que es complementario con respecto al sitio monocatenario de restricción de los fragmentos de ADN circulares rotulados y seguidamente el clivaje de los fragmentos de ADN circulares rotulados en el sitio de restricción mediante la utilización de la endonucleasa de restricción que reconoce el sitio de restricción. Por lo tanto, en algunas formas de realización, el método comprende linealizar los fragmentos de ADN circulares rotulados para generar fragmentos de ssADN cola de milano, o después de la desnaturalización, fragmentos de ssADN lineales dirrotulados.

En algunas formas de realización, el método comprende además la etapa de ligar los fragmentos de ADN lineales desdoblados por endonucleasa de restricción rotulado a una o más moléculas de ADN (por ejemplo, para unir un rótulo).

Por lo tanto, en algunas formas de realización, el método comprende además: amplificar los fragmentos de ADN circulares rotulados o los fragmentos de ssADN lineales cola de milano por transcripción, comprendiendo el método : (a) fusionar a la secuencia de promotor de sentido un oligodesoxirribonucleótido que presenta una secuencia de promotor de antisentido complementaria o fusionar a los fragmentos de ADN circulares rotulados o los fragmentos de dsADN lineales dirrotulados a un cebador que complementario con el mismo y extender el cebador con una ADN polimerasa bajo condiciones bajo las que se sintetiza una dsADN, lo que incluye un promotor polimerasa bicatenaria; y (b) incubar los productos de dsADN con una ARN polimerasa que liga el promotor de ARN polimerasa bajo condiciones en las que se sintetiza ARN.

En algunas formas de realización preferidas en las que la composición de extremo del transposón en horquilla o un

5 cebador de PCR presenta una secuencia promotora de polimerasa de ARN, la ARN polimerasa promotor es una RNA polimerasa promotor de tipo T7 y el método comprende además la etapa de transcribir los fragmentos de ADN circulares rotulados in vitro mediante la utilización de una ARN polimerasa de tipo T7 que reconoce el promotor. Lo más preferible es que, la ARN polimerasa y el promotor sean elegidos de entre T7 RNAP, T3 RNAP y SP6 RNAP y los correspondientes promotores cognato. Sin embargo, las etapas de transcripción de un método de la invención pueden utilizar cualquier ARNP para el que se conozca o pueda obtenerse una secuencia promotora adecuada que permita la transcripción con un elevado carácter específico. Los kits y enzimas para la transcripción in vitro pueden obtenerse en el comercio provistos por muchos vendedores, y las mezclas de reacción y condiciones adecuadas para llevar a cabo las etapas de la presente invención que comprenden la transcripción in vitro pueden utilizar aquellos productos descritos por los fabricantes. A título de ejemplo, la transcripción in vitro mediante la utilización de T7 RNAP puede llevarse a cabo utilizando el kit AMPLISCRIBE™ T7-FLASH™ Transcripción Kit o el kit AMPLISCRIBE™ T7 High Yield Transcription Kit de EPICENTRE Biotechnologies, Madison, WI como se describen la literatura del producto. De manera similar, si se utiliza T3 RNAP o SP6 RNAP en un método de la invención para la transcripción in vitro, puede utilizarse un kit AMPLISCRIBE™ T3-FLASH™ High Yield Transcription Kit o el kit AMPLISCRIBE™ SP6 High Yield Transcription Kit (EPICENTRE Biotechnologies, Madison, WI), respectivamente, como se describe.

10 En algunas otras formas de realización, el método comprende además la etapa de amplificar y/o secuenciar el ADN objetivo en los fragmentos de ADN circulares rotulados mediante la utilización de una ADN polimerasa y de por lo menos un cebador que es complementario con respecto al rótulo. En algunas formas de realización, el método comprende adicionalmente: amplificar los fragmentos de ADN circulares rotulados mediante replicación de rodadura por medio de una ADN polimerasa que desplaza cadenas. En algunas otras formas de realización, el método comprende adicionalmente: amplificar los fragmentos de ADN circulares rotulados por PCR mediante la utilización de una ADN polimerasa termoestable, un primer cebador de PCR que es complementario con respecto a por lo menos una porción del rótulo, y un segundo cebador de PCR que es complementario con respecto a por lo menos una porción del complemento del rótulo.

25 En algunas formas de realización, el método comprende además: amplificar los fragmentos de ADN circulares rotulados mediante PCR (replicación por círculo de rodadura), que comprende: (a) fusionar un cebador que es complementario con respecto a los fragmentos de ADN circulares rotulados; y (b) extender el cebador fusionado a los fragmentos de ADN circulares rotulados mediante la utilización de una ADN polimerasa que desplaza cadenas (por ejemplo, phi29 ADN polimerasa, gran fragmento de rBst ADN polimerasa o r DISPLACEACE™ ADN polimerasa (EPICENTRE). En estas formas de realización, los productos de amplificación de RCR son moléculas de ssADN concataméricas que son complementarios con respecto a los fragmentos de ADN circulares rotulados. En algunas formas de realización en donde los fragmentos de ADN circulares rotulados presentan una secuencia de promotor de antisentido, los productos de amplificación por RCR concataméricos presentan una secuencia de promotor de sentido y el método comprende además preparar la ARN polimerasa promotor bicatenario (por ejemplo, mediante la fusión de la secuencia de promotor de sentido a un oligodesoxirribonucleótido complementario que presenta una secuencia de promotor antisentido, y seguidamente se transcribe el ADN concatamérico mediante la utilización de una ARN polimerasa que religa al promotor de polimerasa bicatenaria e inicia la transcripción a partir de ella.

40 En algunas formas de realización preferidas, la porción de tallo de la composición de extremo del transposón en horquilla presenta solamente las secuencias transferida y no transferida extrema de transposón y el bucle es monocatenario (por ejemplo, a efectos de minimizar la probabilidad o frecuencia de inserción de la composición de extremo del transposón en horquilla en porciones bicatenarias de sí mismo durante la reacción de transposición in vitro). En algunas otras formas de realización, la porción de tallo de la composición de extremo del transposón en horquilla presenta, además de las secuencias transferida y no transferida de extremo de transposón, secuencias adicionales que se hallan inmediatamente a la posición 5' de la secuencia extrema de transposón transferida e inmediatamente a la posición 3' de la secuencia extrema de transposón no transferida. Sin embargo, en estas formas de realización, el tamaño o magnitud de las secuencias adicionales en la porción de tallo de la composición de extremo del transposón en horquilla se minimiza a efectos de minimizar la probabilidad o frecuencia de la inserción de la composición de extremo del transposón en horquilla en sí misma durante la reacción de transposición in vitro. Por ejemplo, en algunas formas de realización, la longitud del tallo en la composición de extremo del transposón en horquilla es inferior a aproximadamente 75 nucleótidos; inferior a aproximadamente 50 nucleótidos, o inferior a aproximadamente 30 nucleótidos.

55 En algunas formas de realización, la porción de bucle de la composición de extremo del transposón en horquilla presenta la secuencia de un dominio de rótulo de secuenciación o un dominio de rótulo de captura (por ejemplo, un dominio de rótulo de secuenciación y/o un dominio de rótulo de captura para el sistema Genome Sequencer FLX System Roche 454, por ejemplo, que presenta las secuencias del dominio de rótulo de secuenciación de los rótulos Roche 454A y 454B que se utilizan para la secuenciación mediante la utilización del sistema Genome Sequencer FLX System Roche 454). En algunas formas de realización en donde la composición de extremo del transposón en horquilla tiene un dominio de rótulo de secuenciación o un dominio de rótulo de captura, los fragmentos de ADN circulares rotulados o los fragmentos de dsADN cola de milano o los fragmentos de ADN dirrotulados tienen un rótulo que comprende el dominio de rótulo de secuenciación y/o el dominio de rótulo de captura (por ejemplo, el rótulo Roche 454A o 454B utilizado para la secuenciación mediante la utilización del sistema Genome Sequencer FLX System Roche 454). Después de aislar los fragmentos de ADN circulares rotulados o los fragmentos de dsADN de cola de milano o los fragmentos de ssADN lineales dirrotulados en el intervalo de tamaños deseado, se los utiliza como

plantillas para la secuenciación de próxima generación mediante la utilización del sistema Genome Sequencer FLX System Roche 454. En otras formas de realización, los fragmentos de ADN circulares rotulados generados o los fragmentos de ssADN lineales dirrotulados, tienen uno o más dominios de sitio de restricción, dominios de rótulos de secuenciación, dominios de rótulos de amplificaciones, dominios de rótulos de capturas, dominios de rótulos de detecciones y/o dominios de rótulos de direcciones destinados a ser utilizados en la secuenciación (por ejemplo, mediante la utilización de la plataforma de secuenciación ROCHE 454, plataforma de secuenciación ILLUMINA™ SOLEXA™, plataforma de secuenciación SOLID™ de LIFE TECHNOLOGIES/APPLIED BIOSYSTEMS, plataforma de secuenciación SMRT™ de PACIFIC BIOSCIENCES, plataforma de secuenciación POLLONATOR Polony, plataforma de secuenciación GENOMICS, plataforma de secuenciación INTELLIGENT BIOSYSTEMS, o el plataforma de secuenciación HELICOS.

En algunas formas de realización, la composición de extremo del transposón en horquilla presenta una o más secuencias de dominio de rótulo, las cuales secuencias pueden utilizarse para llevar a cabo cualquier secuencia deseada. No hay límite en cuando a cuáles secuencias adicionales se utilizan para una o más secuencias adicionales en la porción de bucle de la composición de extremo del transposón en horquilla.

En algunas formas de realización preferidas, el extremo 5' de la composición de extremo del transposón en horquilla tiene u grupo 5'-monofosfato. En formas de realización en las que la composición de extremo del transposón en horquilla no tiene un grupo 5'-monofosfato, el método comprende además la etapa de fosforilar el extremo 5' de la composición de extremo del transposón en horquilla (por ejemplo, mediante la utilización de polinucleótido quinasa; por ejemplo, T4 polinucleótido quinasa) antes de la etapa de ligación del método.

La inserción catalizada por transposasa del extremo de transposón en el objetivo tiene como resultado la unión del extremo 3' de la secuencia extrema de transposón transferida la posición 5' de un nucleótido en una cadena del ADN objetivo, resultando una rotura o fragmentación de dicha cadena en el sitio de dicha cadena donde la secuencia extrema de transposón transferida se une al ADN objetivo y generación concomitante de una región de base-9 de ADN objetivo monocatenario situado en el sitio de la unión del extremo de transposón transferido al ADN objetivo debido a una región de brecha de base 9 en la cadena opuesta del ADN objetivo. Por ejemplo, la FIG. 16 muestra un resultado posible de dos eventos de inserción independiente de la composición de extremo del transposón en horquilla en el ADN objetivo, como se muestra en la FIG. 16, a veces se presentan inserciones independientes del extremo de transposón en horquilla en cadenas opuestas del ADN objetivo en ubicaciones en el ADN objetivo, generándose dos fragmentos de ADN rotulados, como se muestra en in FIG. 16. Al tener lugar la unión, se genera un fragmento de ADN circular rotulado.

La invención no se limita a solamente la ligasa de ácido nucleicos particular descrita en la presente. Las personas expertas en el arte comprenderán de inmediato que cualquier ligasa de ácido nucleico dependiente de plantilla que tenga una actividad a las enzimas descritas en la presente, y los métodos y condición para el uso de tales enzimas para la ligación dependiente de plantilla son conocidos y de fácil obtención en la especialidad.

#### Ejemplos experimentales

La presente invención también se define en los siguientes Ejemplos. Se ha de entender que estos Ejemplos, si bien indican formas de realización preferidas de la invención, se brindan sólo a modo de ilustración. A partir de la discusión anterior y de estos Ejemplos, un experto en la técnica puede determinar las características esenciales de esta invención, y sin apartarse del espíritu y alcance de la misma, puede realizar diversos cambios y modificaciones de la invención para adaptarla a diversos usos y condiciones.

Las técnicas biológicas moleculares estándar usadas son bien conocidas en el arte y se describen por Sambrook, J., Fritsch, E. F. y Maniatis, T., Molecular Cloning: A Laboratory Manual, Second Edition, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, N. Y. (1989).

Definiciones, nomenclatura y abreviaturas usadas en los Ejemplos:

"pMETS" se refiere al oligonucleótido de extremo del transposón monocatenario que contiene 19 bases de 5'-fosfato oligonucleótido que exhibe la secuencia del extremo del transposón EZ-Tn5™:

5' pAGA TGT GTA TAA GAG ACAG 3' (SEQ ID NO:1)

"METS" se refiere al oligonucleótido del extremo del transposón monocatenario de 19 bases que exhibe la secuencia del extremo del transposón EZ-Tn5™:

5' AGA TGT GTA TAA GAG ACAG 3' (SEQ ID NO:1)

"pMENTS" se refiere al oligonucleótido del extremo del transposón monocatenario con contenido de 5'-fosfato de 19 bases que exhibe la secuencia del extremo del transposón EZ-Tn5™:

5' pCTG TCT CTT ATA CAC ATCT 3' (SEQ ID NO:2)

"pMEDS" se refiere al extremo del transposón bicatenario de 19 pares de bases EZ-Tn5™, en donde ambos extremos

5' contienen fosfatos:

5' pAGA TGT GTA TAA GAG ACAG 3' (SEQ ID NO:1)

3' TCT ACA CAT ATT CTC TGTCp 5' (SEQ ID NO:2)

5 El extremo del transposón pMEDS EZ-Tn5<sup>TM</sup> se prepara por fusión del oligonucleótido del extremo del transposón pMETS con el oligonucleótido del extremo del transposón pMENTS.

"MEDS" se refiere al extremo del transposón bicatenario de 19 pares de bases EZ-Tn5<sup>TM</sup>, en donde sólo la cadena no transferida (pMENTS) contiene un 5'-fosfato:

5' AGA TGT GTA TAA GAG ACAG 3' (SEQ ID NO:1)

3' TCT ACA CAT ATT CTC TGTCp 5' (SEQ ID NO:2)

10 El extremo de transposón MEDS EZ-Tn5<sup>TM</sup> se prepara por fusión del oligonucleótido del extremo del transposón METS con el oligonucleótido del extremo del transposón pMENTS.

"p454.1METS" se refiere a la cadena transferida monocatenaria con contenido de 5'-fosfato de 36 bases que tiene una porción 5' que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo, que está adjunta al extremo 5' de la secuencia del extremo del transposón transferida EZ-Tn5<sup>TM</sup> de 19 bases subrayada (pMETS):

15 5'pGCC TTG CCA GCC CGC TCA GAT GTG TAT AAG AGA CAG 3' (SEQ ID NO:4)

"p454.1 MEDS": la composición de extremo del transposón EZ-Tn5<sup>TM</sup> se prepara por fusión de la cadena transferida p454.1METS (SEQ ID NO: 4) con la cadena no transferida pMENTS (SEQ ID NO: 2)

5'pGCC TTG CCA GCC CGC TCA GAT GTG TAT AAG AGA CAG 3'

3'T CTA CAC ATA TTC TCT GTCp 5'

20 "pc454.1" se refiere al oligonucleótido monocatenario que contiene 5'-fosfato de 18 bases que es complementario a la porción 5' de p454.1 METS y tiene la secuencia:

5' pTGA GCG GGC TGG CAA GGC 3' (SEQ ID NO:5)

25 "A-METS" se refiere a la cadena transferida monocatenaria de 38 bases que tiene una porción 5' que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo, que está anexa al extremo 5' de la secuencia del extremo del transposón transferida EZ-Tn5<sup>TM</sup> subrayada de 19 bases (METS):

5' GCC TCC CTC GCG CCA TCA GAG ATG TGT ATA AGA GAC AG 3' (SEQ ID NO:7)

"A-MEDS" la composición de extremo del transposón EZ-Tn5<sup>TM</sup> se prepara por fusión de la cadena transferida A-METS (SEQ ID NO: 7) con la cadena no transferida pMENTS (SEQ ID NO:2):

5' GCC TCC CTC GCG CCA TCA GAG ATG TGT ATA AGA GAC AG 3' 3'TCTACACATATTCTCTGTGTCp5'

30 "B-METS" se refiere a la cadena transferida monocatenario de 38 bases que tiene una porción 5' que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo, que se anexa al extremo 5' de la secuencia del extremo del transposón transferida EZ-Tn5<sup>TM</sup> de 19 bases subrayada (METS):

5' GCC TTG CCA GCC CGC TCA GAG ATG TGT ATA AGA GAC AG 3' (SEQ ID NO:8)

35 "B-MEDS" la composición de extremo del transposón EZ-Tn5<sup>TM</sup> se prepara por fusión de la cadena transferida B-METS (SEQ ID NO: 8) con la cadena no transferida pMENTS (SEQ ID NO:2)

5' GCC TTG CCA GCC CGC TCA GAG ATG TGT ATA AGA GAC AG 3' 3' TC TAC ACA TAT TCT CTG TCp 5'

"FLX-A" se refiere al oligonucleótido monocatenario de 19 bases que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo:

5' GCC TCC CTC GCG CCA TCA G 3' (SEQ ID NO:9)

40 "FLX-B" se refiere al oligonucleótido monocatenario de 19 bases que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo:

5' GCC TTG CCA GCC CGC TCA G 3' (SEQ ID NO:10)

"A-MID2-METS" se refiere a la cadena transferida monocatenaria de 48 bases que tiene una porción 5' que consiste en un rótulo de secuenciación Roche 454 y secuencia de código de barras (MID2, *itálica*) que exhibe la secuencia de

abajo, que se anexa al extremo 5' de la secuencia del extremo del transposón transferida EZ-Tn5™ subrayada de 19 bases (METS):

5' GCC TCC CTC GCG CCA TCA G ACGCTCGACA AG ATG TGT ATA  
AGA GAC AG 3' (SEQ ID NO:11)

5 "Ti A-METS" se refiere a la cadena transferida monocatenario de 49 bases que tiene una porción 5' que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo, que se anexa al extremo 5' de la secuencia del extremo del transposón transferida EZ-Tn5™ subrayada de 19 bases (METS):

5' CCA TCT CAT CCC TGC GTG TCT CCG ACT CAG AGA TGT GTA  
TAA GAG ACA G 3' (SEQ ID NO:12)

10 "Ti B-METS" se refiere a la cadena transferida monocatenaria de 49 bases que tiene una porción 5' que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo, que se anexa al extremo 5' de la secuencia del extremo del transposón transferida EZ-Tn5™ de 19 bases subrayada (METS):

5' CCT ATC CCC TGT GTG CCT TGG CAG TCT CAG AGA TGT GTA  
TAA GAG ACA G 3' (SEQ ID NO:13)

"Ti A" se refiere al oligonucleótido monocatenario de 26 bases que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo:

5' CCA TCT CAT CCC TGC GTG TCT CCG AC 3' (SEQ ID NO:14)

15 "Ti B" se refiere al oligonucleótido monocatenario de 26 bases que consiste en un rótulo de secuenciación Roche 454 que exhibe la secuencia de abajo:

5' CCT ATC CCC TGT GTG CCT TGG CAG TC 3' (SEQ ID NO:15)

"BP1-A" se refiere al oligonucleótido monocatenario de 48 bases que tiene una porción 5' que consiste en un rótulo de PCR Illumina bridge que exhibe la secuencia de abajo, que se anexa a la secuencia FLX-A (subrayado abajo):

5' AAT GAT ACG GCG ACC ACC GAG ATC TAC ACG CCT CCC TCG  
CGC CAT CAG 3' (SEQ ID NO:16)

20 "BP2-B" se refiere al oligonucleótido monocatenario de 49 bases que tiene una porción 5' que consiste en un rótulo de PCR Illumina bridge que exhibe la secuencia de abajo, que se anexa a la secuencia FLX-B (subrayada debajo):

5' CAA GCA GAA GAC GGC ATA CGA GAT CGG TCT GCC TTG CCA  
GCC CGC TCA G 3' (SEQ ID NO:17)

25 "BP2-ID1-B" se refiere al oligonucleótido monocatenario de 49 bases que tiene una porción 5' que consiste en un rótulo de PCR Illumina bridge y una secuencia de código de barras (ID2, itálica) que exhibe la secuencia de abajo, que se anexa a la secuencia FLX-B (subrayada debajo):

5' CAA GCA GAA GAC GGC ATA CGA GAT *GCATGTCGG* TCT GCC TTG  
CCA GCC CGC TCA G 3' (SEQ ID NO:18)

"BP1" se refiere al oligonucleótido monocatenario de 20 bases que consiste en un rótulo adaptador Illumina bPCR que exhibe la secuencia de abajo:

30 5' AAT GAT ACG GCG ACC ACC GA 3' (SEQ ID NO:19)

## ES 2 743 029 T3

“BP2” se refiere al oligonucleótido monocatenario de 21 bases que consiste en un rótulo adaptador Illumina bPCR que exhibe la secuencia de abajo:

5' CAA GCA GAA GAC GGC ATA CGA 3' (SEQ ID NO:20)

5 “pMETS-N-MENTS” se refiere a una composición de extremo del transposón en horquilla que comprende o que consiste en: un oligonucleótido con contenido de 5'-fosfato que exhibe la secuencia del extremo del transposón no transferida EZ-Tn5™ en el extremo 5' y la secuencia del extremo del transposón transferida EZ-Tn5™ en el extremo 3', conectado por una secuencia arbitraria interviniente representada por “(N)x”. La secuencia interviniente entre las secuencias METS y MENTS consiste en una cantidad suficiente de nucleótidos para permitir la formación de tallo-bucle:

10 5' PCTGTCTCTTATACACATCT-(N)<sub>x</sub>-AGATGTGTATAAGAGACAG 3' (SEQ ID NO:3)

La fusión intramolecular de la secuencia del extremo del transposón transferida pMETS con la secuencia del extremo del transposón no transferida pMENTS dentro de un oligonucleótido pMETS-N-MENTS forma una composición de extremo del transposón en horquilla EZ-Tn5™. Por ejemplo, si x = 6;

```

NN
N  AGATGTGTATAAGAGACAG  3'
N  TCTACACATATTCTCTGTCp 5'  (SEQ ID NO:3)
NN
    
```

15 “TSasa ” se refiere a la transposasa hiperactiva EZ-Tn5™ Tn5 (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos) en 50 mM de cloruro de Tris pH 7,5, 50% de glicerol, 0,1 mM de EDTA, 1 mM de DTT, 500 mM de cloruro de sodio, 0,5% v/v de NP-40, 0,5% v/v de Tween-20.

20 “Transposoma” se refiere a la transposasa hiperactiva EZ-Tn5™ Tn5 (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos) preincubada con ADN de transposón bicatenario en condiciones que sustentan la formación del complejo no covalente. El ADN de transposón bicatenario puede consistir, sin limitación, en Tn5 ADN, una porción de Tn5 ADN, una composición de extremo del transposón, una mezcla de composiciones de extremo del transposón, una mezcla de composiciones de extremo del transposón u otros ADN bicatenarios capaces de interactuar con la transposasa hiperactiva EZ-Tn5™.

10 X TA tampón de reacción:	330 mM de acetato de Tris, pH 7,8
	100 mM de acetato de magnesio
	660 mM de acetato de potasio

5 X TA-DMF tampón de reacción:	165 mM de acetato de Tris, pH 7,8
	50 mM de acetato de magnesio 3
	30 mM de acetato de potasio
	50% v/v de dimetilformamida

10 X TMgCl tampón de reacción:	100 mM de cloruro de Tris, pH 8,0
	50 mM de cloruro de magnesio

5 X TMgCl-DMF tampón de reacción:	50 mM de cloruro de Tris, pH 8,0
	25 mM de cloruro de magnesio
	50% v/v de dimetilformamida

10 X TMgAc tampón de reacción:	100 mM de acetato de Tris, pH 7,6
--------------------------------	-----------------------------------

## ES 2 743 029 T3

50 mM de cloruro de magnesio

5 X TMgAc-DMF tampón de reacción: 50 mM de acetato de Tris, pH 7,6  
25 mM de cloruro de magnesio  
50% v/v de dimetilformamida

“ADN objetivo” se refiere al ADN sometido a transposición. En el ejemplo siguiente, ADN de bacteriófago T7D111 se usa como el ADN objetivo.

5 “TSasa” se refiere a la transposasa hiperactiva EZ-Tn5™ Tn5 (EPICENTRE Biotechnologies, Madison, WI, Estados Unidos).

10 X Transposasa tampón de reacción: 330 mM de acetato de Tris, pH 7,8  
100 mM de acetato de magnesio  
660 mM de acetato de potasio

### EJEMPLO 1

10 Fragmentación de ADN mediada por transposición in vitro y rotulación 5' usando transposasa EZ-Tn5™ y el extremo de transposón EZ-Tn5™

La siguiente mezcla de reacción se colocó:

x	agua hasta un volumen final de 50 microlitros
5 microlitros	10 X tampón de transposición EZ-Tn5™
1 microgramo	ADN objetivo en 1 a 40 microlitros
2 microlitros	pMEDS (25 micromolar)*
2 microlitros	transposasa EZ-Tn5™ (a 10 unidades por microlitro)
50 microlitros	

\* En algunas formas de realización, dos diferentes extremos de transposón pMEDS que exhiben cada uno adicionalmente una secuencia arbitraria diferente en su respectiva porción 5' del extremo del transposón transferido, 5' de la secuencia del extremo del transposón transferida (FIG. 4).

Después de mezclar, la reacción se incubó durante 1 hora a 37 °C. La reacción se detuvo con 10 microlitros de solución de detención (15% de sacarosa, 66 mM de EDTA, 20 mM de TRIS, pH 8,0, 0,1% de SDS, 0,9% de Orange G [Sigma 0-7252] y Proteinasa K a 100 microgramos por ml), se mezcló y se calentó a 50 °C durante 10 minutos.

15 El ADN se analizó por electroforesis en gel de agarosa al 1% en tampón TAE. LMP agarosa se usó para aislar el 5 ADN en clases de tamaños. Los geles se tiñeron con SYBR Gold y el ADN se visualizó con luz no UV. Trozos de gel para geles LMP se incubaron a 70 °C durante 5 minutos para licuar el gel. Después de 5 minutos a 37 °C, se añadió 100% en volumen de solución de digestión de agarosa Gelase™ (EPICENTRE Biotechnologies). La reacción se mezcló y se incubó durante 1 hora a 37 °C.

20 El ADN objetivo se fragmentó en una medida similar y en un rango de tamaño similar tal como se describe en los Ejemplos 3 y 4 usando cantidades y concentraciones comparables de la transposasa EZ-Tn5™ Tn5 y los extremos del transposón. El ADN del procedimiento de tamaño se usó en el EJEMPLO 2 para rotular los extremos 3' de los fragmentos de ADN con rótulo en 5'.

### EJEMPLO 2

25 Rango de tamaño de productos de transposición de fragmentos de ADN con rótulo en 5' usando diferentes concentraciones de transposasa EZ-Tn5™ Tn5

La transposasa Tn5 hiperactiva EZ-Tn5™ (EPICENTRE) a una concentración de 90 unidades por microlitro se diluyó hasta concentraciones finales de 45, 22,5, 11,3 y 9 unidades por microlitro. 2 microlitros de la enzima en cada

concentración se incubaron con 1 microgramo de ADN objetivo de fago T7 D111 (con un tamaño de aproximadamente 39 Kbp) y 1 micromolar del extremo del transposón pMEDS en tampón TA en un volumen de mezcla de reacción final de 50 microlitros durante 1 hora a 37 °C.

5 Las reacciones se detuvieron con 10 microlitros de una solución de detención con 15% de sacarosa, 66 mM de EDTA, 20 mM de Tris/HCl pH 8,0, 0,1% de SDS, 0,9% de Orange G y 100 microgramos por ml de proteinasa K. Después de mezclar e incubar a 50 °C durante 10 min, se sometieron alícuotas de 10 microlitros a electroforesis en gel de agarosa al 1% en tampón TAE durante 1 hora a 100 voltios. El gel se tiñó con SYBR Gold y se fotografió con transiluminación A340.

10 Una concentración final de aproximadamente 0,9 unidades por microlitro de Tn5 transposasa en la mezcla de reacción dio una máxima fragmentación del ADN objetivo del fago T7 D111. Mayores concentraciones de Tn5 transposasa eran inhibitorias y las menores concentraciones desplazaron el rango de tamaño del fragmento hacia arriba. A una concentración final de aproximadamente 0,9 unidades de Tn5 transposasa por microlitro, la mayor parte del ADN objetivo del fago T7 D111 se fragmentó en ADN que migró en el gel en tamaños de entre aproximadamente 150 bp y aproximadamente 1,5 Kbp en base a las bandas marcadoras. A una concentración final de aproximadamente 0,45 unidades de Tn5 transposasa por microlitro, la mayor parte del ADN objetivo del fago T7 D111 se fragmentó en ADN que migró al gel en tamaños de entre aproximadamente 400 bp y aproximadamente 3,5 Kbp en base a las bandas marcadoras.

### EJEMPLO 3

20 Rango de tamaño de productos de transposición del fragmento de ADN con rótulo en 5' usando diferentes concentraciones del extremo de transposón pMEDS

Una solución madre 25 micromolar del extremo de transposón pMEDS se diluyó serialmente 2, 4 y 8 veces con tampón T10E1. A continuación, se incubaron 2 microlitros de cada dilución final de transposón y un control de tampón final no transposón, en reacciones de 50 microlitros que contenían 1 X tampón TA, 1 microgramo de ADN objetivo de fago T7 D111 y 0,4 unidades por microlitro de transposasa hiperactiva Tn5 durante 1 hora a 37 °C.

25 Las reacciones se detuvieron y las muestras se analizaron por electroforesis en gel de agarosa al 1 % tal como se describió en el EJEMPLO 2.

30 La dilución de 4 veces de la solución madre de 25 uM, que dio como resultado una concentración final de 0,25 micromolar del extremo de transposón pMEDS en la mezcla de reacción, dio como resultado una buena fragmentación del ADN objetivo y probablemente era más eficaz en términos de uso del extremo de transposón pMEDS. En esta concentración, la mayoría del ADN objetivo del fago T7 D111 se fragmentó en ADN que migró al gel en tamaños de entre 400 bp y aproximadamente 3,5 Kbp en base a las bandas marcadoras. En concentraciones de 0,5 y 1 micromolar del extremo de transposón pMEDS, los tamaños del ADN fragmentado se desplazaron ligeramente hacia abajo entre aproximadamente 200-300 bp y aproximadamente 3 Kbp.

### EJEMPLO 4

35 Rango de tamaño de los productos de transposición del fragmento de ADN etiquetado en 5' usando diferentes concentraciones de transposoma

40 Rango de tamaño de productos de transposición del fragmento de ADN con rótulo en 5' usando diferentes concentraciones de transposoma "A-MEDS transposomas" y "B-MEDS transposomas" se formaron preincubando 12,5 µM de TSasa con 12,5 µM de A-MEDS o 12,5 µM de composiciones de extremo de transposón B-MEDS, respectivamente, durante 60 minutos a 37 °C. Los transposomas A-MEDS y B-MEDS se combinaron en iguales relaciones para formar "A/B Transposomas".

Luego se usaron transposomas a 125 µM o se diluyeron hasta 10 µM, 7,5 µM, 5 µM, 2,5 µM o 1 µM con tampón de almacenamiento TSasa (50 mM de cloruro de Tris, pH 7,5, 50% de glicerol, 0,1 mM de EDTA, 1 mM de DTT, 500 mM de cloruro de sodio, 0,5% v/v de NP-40, 0,5% v/v de Tween-20).

45 ADN genómico de E. coli se rotuló en el extremo 5' y se fragmentó usando A/B transposomas en las siguientes reacciones:

Reactivo	Volumen
5 X TMgCl-DMF tampón de reacción	4 µl
50 ng/µl de ADN genómico de E. coli	1 µl
A/B Transposoma (12,5, 10, 7,5, 5, 2,5 o 1 µM)	1 µl
Agua	14 µl

Volumen final:	20 µl
----------------	-------

Las reacciones se incubaron durante 5 minutos a 55 °C. A continuación, las reacciones se detuvieron con 5 microlitros de solución de detención (15% de sacarosa, 66 mM de EDTA, 20 mM de TRIS, pH 8,0, 0,1% de SDS, 0,9% de Orange G [Sigma 0-7252] y Proteinasa K a 100 microgramos por ml), se mezclaron y se calentaron a 70 °C durante 10 minutos.

- 5 El ADN se analizó por electroforesis en gel de agarosa al 1% en tampón TAE. Los geles se tiñeron con SYBR Gold y el ADN se visualizó con luz no UV.

- 10 El grado de fragmentación de ADN objetivo es proporcional a la cantidad de transposoma añadida durante la dilución de 12,5 veces de la solución madre de 12,5 µM de transposoma. A altas concentraciones de transposoma, la mayor parte de los fragmentos de ADN migró al gel en tamaños inferiores a 1000 bp (Figure 5, carriles 3 y 9). A bajas concentraciones de transposomas, los fragmentos de ADN migraron al gel predominantemente entre 500 bp y 6000 bp (Figure 5, carriles 8 y 14). La flecha en bloque indica la migración libre de composición de extremo del transposón en gel.

#### EJEMPLO 5

Fragmentación de ADN objetivo y rotulación en 5' a 55 °C y 37 °C en presencia de dimetilformamida

- 15 Para evaluar el efecto de la dimetilformamida sobre la fragmentación del ADN objetivo y la rotulación 5', el ADN genómico de HeLa se fragmentó y se rotuló con ME transposomas o A/B transposomas de la siguiente manera.

“ME Transposomas” se formaron preincubando 12,5 µM de TSasa con 12,5 µM de composiciones de extremo del transposón MEDS durante 60 minutos a 37 °C.

Se fijaron reacciones por duplicado de la siguiente manera:

Reactivo	TA	TA-DMF	TMgCl	TMgCl-DMF
10 X TA tampón de reacción	2 µl	---	---	--
5 X TA-DMF tampón de reacción	---	4 µl	---	--
10 X TMgCl tampón de reacción	---	---	2 µl	--
5 X TMgCl-DMF tampón de reacción	---	---	---	4 µl
50 ng/µl de ADN genómico de HeLa	1 µl	1 µl	1 µl	1 µl
ME Transposoma (12,5 µM)	1 µl	1 µl	1 µl	1 µl
Agua	16 µl	14 µl	16 µl	14 µl
Volumen final:	20 µl	20 µl	20 µl	20 µl

20

“A-MEDS Transposomas” y “B-MEDS Transposomas” se formaron preincubando 12,5 µM de TSasa con 12,5 µM de A-MEDS o 12,5 µM de composiciones de extremo de transposón B-MEDS, respectivamente, durante 60 minutos a 37 °C. Se combinaron A-MEDS y B-MEDS transposomas en relaciones iguales para formar “A/B Transposomas”.

Se fijaron reacciones por duplicado de la siguiente manera:

Reactivo	TA	TA-DMF	TMgCl	TMgCl-DMF
10 X TA tampón de reacción	2 µl	---	--	
5 X TA-DMF tampón de reacción	---	4 µl	---	--
10 X TMgCl tampón de reacción	---	---	2 µl	
5 X TMgCl-DMF tampón de reacción	---	---	---	4 µl
50 ng/µl de ADN genómico de HeLa	1 µl	1 µl	1 µl	1 µl
MA/B Transposoma (12,5 µM)	1 µl	1 µl	1 µl	1 µl

Agua	16 µl	14 µl	16 µl	14 µl
Volumen final:	20 µl	20 µl	20 µl	20 µl

Las reacciones se incubaron durante 5 minutos a 37 °C o durante 5 minutos a 55 °C. Las reacciones se detuvieron con 5 microlitros de solución de detención (15% de sacarosa, 66 mM de EDTA, 20 mM de TRIS, pH 8,0, 0,1% de SDS, 0,9% de Orange G [Sigma 0-7252] y Proteinasa K a 100 microgramos por ml), se mezclaron y se calentaron a 70 °C durante 10 minutos.

El ADN se analizó por electroforesis en gel de agarosa al 1% en tampón TAE. Los geles se tiñeron con SYBR Gold y el ADN se visualizó con luz no UV.

La dimetilformamida mejoró la eficacia de la fragmentación y la reacción de rotulación en 5' tal como se juzgó por medio de la reducción en la distribución de MW del producto de reacción (Figura 6, comparar los carriles 4, 6, 8, 10, 13, 15, 17 y 19 a carriles 3, 5, 7, 9, 12, 14, 16 y 18, respectivamente). De modo similar, las reacciones en presencia de tampón de reacción TMgCl eran más eficaces que las reacciones en presencia del tampón de reacción TA. Finalmente, pareció que las reacciones a 55 °C mejoran la eficacia general de las reacciones en comparación con las reacciones a 37 °C (Figura 6, comparar los carriles 4-10 con los carriles 12-19). La flecha en bloque indica la migración de la composición de extremo del transposón libre en gel.

#### EJEMPLO 6

Rotulación por fragmentación de ADN objetivo usando MuA Transposasa.

HyperMu™ MuA transposasa (EPICENTRE) a una concentración final de 1 unidad por microlitro y luego un rango de concentraciones de proteína de MuA transposasa de entre aproximadamente 0,01 microgramos y aproximadamente 0,5 microgramos de proteína por microlitro de mezcla de reacción se incubó en una reacción de 50 microlitros con tampón de reacción de MuA transposasa (EPICENTRE), 1 microgramo de ADN objetivo de fago T7 D111 y 1 micromolar de extremo de transposón pR1R2 MuA durante 1 hora a 37 °C.

La reacción se detuvo y los productos se analizaron por electroforesis en gel de agarosa tal como se describió en el EJEMPLO 3.

La fragmentación del ADN objetivo del fago T7 D111 era mucho menor que el observado usando la transposasa EZ-Tn5™ Tn5 en todos los niveles de MuA transposasa ensayada. Una cantidad muy pequeña de fragmentación se observó solamente con la máxima concentración de MuA transposasa ensayada. Así, el uso de MuA transposasa y el extremo del transposón pR1R2 MuA era mucho menos eficiente para la fragmentación y rotulación en 5' del ADN objetivo que EZ-Tn5™, la transposasa Tn5 hiperactiva y el extremo del transposón EZ-Tn5™ Tn5 ME.

#### EJEMPLO 7

Rotulación de los extremos 3' de los fragmentos de ADN rotulados en 5'

A. PCR de dos cebadores

A fin de rotular los extremos 3' de los fragmentos generados por transposición y los fragmentos de ADN con rótulo en 5' con la secuencia del extremo del transposón transferida, se llevó a cabo la siguiente reacción:

22 microlitros de productos de transposición rotulados en 5' seleccionados por tamaño de 0,5-1 Kbp

25 microlitros de Failsafe™ PCR premezcla C

1 microlitro de Failsafe™ ADN polimerasa (EPICENTRE)

2 microlitros 5 micromolar de cada cebador oligonucleótido de PCR, del que es complementario a la porción 5' de cada una de las porciones 5' de los dos diferentes extremos del transposón transferidos.

50 microlitros de volumen de reacción total

Como FailSafe ADN polimerasa tiene una actividad de desplazamiento de cadena y de 5' nucleasa, la polimerización del método se lleva a cabo incubando la reacción durante 10 minutos a 70 °C (etapa de extensión de 3' ADN polimerasa), generando así fragmentos de ADN rotulados en 5' y 3' (FIG. 8).

A continuación, la reacción se incubó a 94 °C durante 5 minutos para desnaturalizar el ADN.

La amplificación de los fragmentos de ADN rotulados en 5' y 3' se lleva a cabo por amplificación por PCR de los fragmentos de ADN rotulados en 5' y 3' usando los dos cebadores de PCR donde cada uno es complementario a la porción 5' de uno de los dos diferentes extremos del transposón transferidos.

La mezcla de reacción de PCR anterior se somete a PCR para 20 ciclos con las siguientes condiciones de ciclación:

94 °C 10 seg.

55 °C 10 seg.

72 °C 2 min.

- 5 El análisis en gel indicó que se produjeron los productos de PCR del rango de tamaño esperado (0,5 - 1 Kbp).

También se llevan a cabo reacciones de control: si los productos de transposición seleccionados por tamaño se desnaturalizan con calor antes de la PCR y sin la etapa de extensión de 3' ADN polimerasa, no se produce ningún producto de PCR de 0,5 - 1 Kbp.

B. PCR con cebador simple

- 10 A fin de rotular los extremos 3' de los fragmentos generados por transposición y los fragmentos rotulados en 5' ME con la secuencia ME, se llevó a cabo la siguiente reacción:

23 microlitros de productos de transposición rotulados en 5' seleccionados por tamaño de 0,5-1 Kbp

25 microlitros de Failsafe™ PCR premezcla C

1 microlitro de Failsafe™ ADN polimerasa (EPICENTRE)

- 15 1 microlitro 5 micromolar de cada cebador oligonucleótido de PCR.

50 microlitros de volumen de reacción total

Como FailSafe ADN polimerasa tiene una actividad de desplazamiento de cadena y de 5' nucleasa, la polimerización del método se llevó a cabo incubando la reacción durante 10 minutos a 70 °C (etapa de extensión de 3' ADN polimerasa), generando así fragmentos de ADN rotulados en 5' y 3' (FIG. 7).

- 20 A continuación, la reacción se incubó a 94 °C durante 5 minutos para desnaturalizar el ADN.

La amplificación de los fragmentos de ADN rotulados en 5' y 3' se llevó a cabo por amplificación por PCR de los fragmentos de ADN rotulados en 5' y 3' usando pMETs como el único cebador oligonucleótido de PCR. La mezcla de reacción de PCR anterior se sometió a PCR para 20 ciclos con las siguientes condiciones de ciclación:

94 °C 10 seg.

- 25 55 °C 10 seg.

72 °C 2 min.

El análisis en gel indicó que se produjeron los productos de PCR del rango de tamaño esperado (0,5 - 1 Kbp).

- 30 También se llevaron a cabo reacciones de control: si los productos de transposición seleccionados por tamaño se desnaturalizan con calor antes de la PCR y sin la etapa de extensión de 3' ADN polimerasa, no se produjo ningún producto de PCR de 0,5 - 1 Kbp.

#### EJEMPLO 8

Amplificación y secuenciación profunda de una biblioteca de fragmentos de ADN

- 35 A fin de generar una biblioteca de fragmentos de ADN no selectiva que se puede amplificar antes de la preparación de la biblioteca, se generaron fragmentos de ADN y se rotularon en el extremo 3' y el extremo 5' usando "ME Transposomas".

Se formaron "ME Transposomas" preincubando 10 µM de TSasa con 10 gM de composiciones de extremo del transposón MEDS durante 10 minutos sobre hielo.

Un ADN de cósmido de 43 kb se fragmentó y se rotuló en 5' en la siguiente reacción:

Reactivo	Volumen
10 X TA tampón de reacción	5 µl
142 ng/µl 43 kb de cósmido de ADN	7 µl

ME Transposoma (10 µM)	5 µl
Agua	33 µl
Volumen final:	50 µl

La reacción se incubó durante 2 horas a 37 °C. Se añadieron 5 µl adicionales de 10 µM de ME Transposoma a la reacción y se incubó durante 2 horas más a 37 °C.

5 A fin de rotular los extremos 3' de los fragmentos generados por transposición y los fragmentos de ADN con rótulo en 5' con la secuencia del extremo del transposón transferida, los productos de reacción se incubaron con una mezcla de polimerasa de desplazamiento de cadena (FailSafe™) y dNTPs.

10 Una porción de los fragmentos generados por transposición y los fragmentos de ADN con rótulo en 5' se diluyó 1:10 antes de la rotulación en el extremo 3' y amplificación para caracterizar la amplificación de 4 ng de la plantilla de la biblioteca de ADN. A fin de amplificar no selectivamente la población entera de fragmentos de ADN rotulados usando un cebador de PCR simple, se llevó a cabo la siguiente reacción con el cebador de PCR METS, que hibrida sólo en la secuencia del extremo de transposón y no contiene información adicional de secuencia en 3'.

Reactivo	Volumen
2 X FailSafe™ tampón E de PCR	12,5 µl
Fragmentos de ADN con rótulo en 5' (diluido 1:10)	2 µl
Cebador de PCR METS (25 µM)	1 µl
FailSafe™ enzima de PCR, 2,5 U/µl	1 µl
Agua	8,5 µl
Volumen final:	25 µl

La reacción se incubó de la siguiente manera:

- 72 °C/2:00\*
- 15 - 98 °C/1:00
- 25 ciclos de (98 °C/0:10, 55 °C/0:10, 72 °C/1:00)
- 4 °C de mantenimiento

20 \*- A fin de rotular los extremos 3' de los fragmentos generados por transposición y los fragmentos de ADN con rótulo en 5' con la secuencia del extremo del transposón transferida, los productos de reacción se incubaron con una mezcla de polimerasa de desplazamiento de cadena (FailSafe™) y dNTPs antes de la etapa de desnaturalización (véase la FIG. 7).

Los productos de reacción amplificados y no amplificados se purificaron usando una columna QIAGEN PCR-Clean-up según las instrucciones del fabricante y se usaron como entrada para la etapa 3.4 del protocolo de preparación de la biblioteca estándar Roche/454 FLX según las instrucciones del fabricante (USM00048.A, octubre de 2008).

25 Una secuenciación profunda de las bibliotecas fragmentadas con transposones produjo un cóntigo simple del tamaño esperado con longitud de lectura, precisión y cobertura que era comparable con una biblioteca de control producida usando nebulización (Figura 9). Estos datos son consistentes con la amplificación no selectiva y masivamente paralela de una biblioteca de fragmentos de ADN.

#### EJEMPLO 9

30 Preparación de bibliotecas de secuenciación compatibles con código de barras Roche/454 FLX por adición de información de rotulación de secuenciación en 5' y 3' adicional usando PCR con oligonucleótidos adaptadores

35 A fin de generar una biblioteca con código de barras de fragmentos de ADN que se puede usar directamente en 5 emPCR para secuenciación 454 GS FLX, se fragmentó el lambda ADN genómico y se rotuló en 5' con ME transposomas. Se usaron oligonucleótidos adaptadores no selectivos durante la PCR para anexar la biblioteca de ADN con 454 FLX emPCR y adaptador de secuenciación y secuencia con código de barras (FIG. 10).

## ES 2 743 029 T3

Se formaron "ME Transposomas" preincubando 12,5  $\mu\text{M}$  de TSasa con 12,5  $\mu\text{M}$  de composiciones de extremo del transposón MEDS durante 60 minutos a 37 °C.

Lambda ADN genómico se fragmentó y se rotuló en 5' en la siguiente reacción:

Reactivo	Volumen
10 X TA tampón de reacción	5 $\mu\text{l}$
500 ng/ $\mu\text{l}$ Lambda ADN	2 $\mu\text{l}$
ME Transposoma (12,5 $\mu\text{M}$ )	2 $\mu\text{l}$
Agua	39 $\mu\text{l}$
Volumen final:	48 $\mu\text{l}$

- 5 La reacción se incubó durante 2 horas a 37 °C. Se añadieron 2  $\mu\text{l}$  adicionales de 12,5  $\mu\text{M}$  de ME Transposoma a la reacción y se incubó durante 2 horas más a 37 °C.

Los productos de reacción se purificaron usando columna QIAGEN PCR-Clean-Up según las instrucciones del fabricante.

- 10 A fin de amplificar no selectivamente y anexar la biblioteca de fragmentos de ADN con adaptadores compatibles con Roche/454 FLX emPCR y secuenciación, la PCR se realizó usando adaptadores oligos que hibridan en la secuencia del extremo del transposón y no contiene información de secuencia 3' adicional (FIG. 10).

Reactivo	Volumen
2 X FailSafe™ tampón E de PCR	25 $\mu\text{l}$
Fragmentos de ADN rotulados en 5' (20 ng/ $\mu\text{l}$ )	0,5 $\mu\text{l}$
A-MID2-METS cebador de PCR (2,5 $\mu\text{M}$ )	1 $\mu\text{l}$
B-METS cebador de PCR (2,5 $\mu\text{M}$ )	1 $\mu\text{l}$
FLX-A cebador de PCR (50 $\mu\text{M}$ )	1 $\mu\text{l}$
FLX-B cebador de PCR (50 $\mu\text{M}$ )	1 $\mu\text{l}$
FailSafe™ enzima de PCR, 2,5 U/ $\mu\text{l}$	1 $\mu\text{l}$
Agua	24,5 $\mu\text{l}$
Volumen final:	50 $\mu\text{l}$

La reacción se incubó de la siguiente manera:

- 15
- 72 °C/5:00\*
  - 98 °C/2:00
  - 4 ciclos de (98 °C/0:10, 37 °C/0:30, 72 °C/3:00)
  - 6 ciclos de (98 °C/0:10', 64 °C/3:00)
  - 4 °C de mantenimiento

- 20 \*- A fin de rotular los extremos 3' de los fragmentos generados por transposición y los fragmentos de ADN con rótulo en 5' con la secuencia del extremo del transposón transferida, los productos de reacción se incubaron con una mezcla de polimerasa de desplazamiento de cadena (FailSafe™) y dNTPs antes de la etapa de desnaturalización (véase la FIG. 7).

Las reacciones de control omitieron los cebadores de PCR A-MID2-METS y B-METS y contenían 20 ng de fragmentos de ADN con rótulo en 5'.

- 25 La reacción de PCR produjo una biblioteca de emPCR compatible con la distribución de MW esperado y era similar a la de los fragmentos de ADN generados por transposición y rotulados en 5' (Figura 11, carriles 3 y 4). La falta de productos de amplificación detectables en reacciones que carecen de cebadores adaptadores (A-MID2-METS y B-

METS) es consistente con la amplificación específica de una biblioteca de ADN rotulados FLX-A y FLX-B (Figura 11, carril 5).

EJEMPLO 10

5 Secuenciación profunda de biblioteca de secuenciación compatible con Roche/454 FLX Titanium a partir de 50 ng de cADN de amplicón viral

A fin de generar una biblioteca de fragmentos de ADN no selectiva que se puede usar directamente en la emPCR para secuenciación de Roche/454 FLX Titanium, el ADN de amplicón se fragmentó y se rotuló en 5' con ME transposomas. Se usaron oligonucleótidos adaptadores no selectivos para anexar la biblioteca de ADN con emPCR Roche/454 FLX Titanium y secuencia adaptadora de la secuenciación (FIG. 10).

10 "ME Transposomas" se formaron preincubando 12,5 µM de TSasa con 12,5 µM de composiciones de extremo del transposón MEDS durante 60 minutos a 37 °C. Esta solución madre se diluyó hasta 7,5 µM en tampón de almacenamiento TSasa (50 mM de cloruro de Tris pH 7,5, 50% de glicerol, 0,1 mM de EDTA, 1 mM de DTT, 500 mM de cloruro de sodio, 0,5% v/v de NP-40, 0,5% v/v de Tween-20).

El cADN de amplicón viral se fragmentó y se rotuló en 5' en la siguiente reacción:

15

Reactivo	Volumen
10 X TA tampón de reacción	5 µl
9,4 ng/gl de cADN de amplicón viral	5,5 µl
ME Transposoma (7,5 µM)	1 µl
Agua	38,5 µl
Volumen final:	50 µl

La reacción se incubó durante 15 minutos a 55 °C. Se añadió 1 µl adicional de 7,5 µM de ME Transposoma a la reacción y se incubó durante 15 minutos más a 55 °C.

20 Los productos de reacción se purificaron usando columna QIAGEN PCR-Clean-Up según las instrucciones del fabricante usando dos eluciones de 11 µL que se reunieron.

A fin de amplificar no selectivamente y anexar la biblioteca de fragmentos de ADN con adaptadores compatibles con emPCR Roche/454 FLX Titanium y secuenciación, se llevó a cabo una PCR usando oligos adaptadores que hibridan en la secuencia del extremo del transposón y no contienen información adicional de secuencia 3' (FIG. 10).

Reactivo	Volumen
2 X FailSafe™ tampón E de PCR	25 µl
Fragmentos de PCR rotulados en 5' (-20 µl recuperados)	5 µl
Ti A-METS cebador de PCR (0,5 µM)	1 µl
Ti B-METS cebador de PCR (0,5 µM)	1 µl
Ti A cebador de PCR (10 µM)	1 µl
Ti B cebador de PCR (10 µM)	1 µl
FailSafe™ enzima de PCR, 2,5 U/µl	1 µl
Agua	15 µl
Volumen final:	50 µl

25

La reacción se incubó de la siguiente manera:

- 72 °C/5:00\*
  - 98 °C/2:00
  - 4 ciclos de (98 °C/0:10, 55 °C/0:30, 72 °C/3:00)
- 5    • 6 ciclos de (98 °C/0:10, 64 °C/3:00)
- 4 °C de mantenimiento

\*- A fin de rotular los extremos 3' de los fragmentos de ADN generados por transposición y fragmentos de ADN con rótulo en 5' con la secuencia del extremo del transposón transferida, los productos de reacción se incubaron con una mezcla de polimerasa de desplazamiento de cadena (FailSafe™) y dNTPs antes de la etapa de desnaturalización (véase la FIG. 7).

La reacción de PCR produjo una biblioteca compatible con emPCR con la distribución esperada de MW y era similar a la de los fragmentos de ADN generados por transposición y los fragmentos de ADN con rótulo en 5' (Figura 12, carriles 2 y 3).

La secuenciación profunda de la biblioteca en 1/8 de una placa de Roche/454 FLX Titanium proporcionó aproximadamente 80.000 lecturas con aproximadamente el 95% de las lecturas de mapeo del rendimiento genómico viral de referencia con la cobertura esperada (datos no mostrados). Estos datos son consistentes con la amplificación no selectiva y masivamente paralela de una biblioteca de fragmentos de ADN.

EJEMPLO 11

Preparación de bibliotecas de secuenciación compatible con Illumina GAI de código de barras por adición de información de rotulación de secuenciación en 5' y 3' adicional usando PCR con oligonucleótidos adaptadores

A fin de generar una biblioteca de fragmentos de ADN que se pueden usar directamente en bPCR para secuenciación Illumina GAI, el lambda ADN genómico se fragmentó y se rotuló en 5' con A/B transposomas. Se usaron oligonucleótidos adaptadores no selectivos para anexar la biblioteca de ADN con adaptador de bPCR Illumina GAI y secuencia de código de barras (FIG. 13).

“A-MEDS Transposomas” y “B-MEDS Transposomas” se llevaron a cabo preincubando 12,5 µM de TSasa con 12,5 µM de A-MEDS o 12,5 µM de composiciones de extremo del transposón B-MEDS, respectivamente, durante 60 minutos a 37 °C.

Lambda ADN genómico se fragmentó y se rotuló en 5' en la siguiente reacción:

Reactivo	Volumen
10 X TA tampón de reacción	5 µl
500 ng/µl de lambda ADN	2 µl
A-MEDS Transposoma (12,5 µM)	2 µl
B-MEDS Transposoma (12,5 µM)	2 µl
Agua	39 µl
Volumen final:	50 µl

La reacción se incubó durante 2 horas a 37 °C. Se añadieron 2 µl más de 12,5 µM de A-MEDS Transposoma y 2 µl de 12,5 µM de B-MEDS Transposoma a la reacción y se incubó durante 2 horas más a 37 °C.

Los productos de reacción se purificaron usando columna QIAGEN PCR-Clean-Up según las instrucciones del fabricante.

A fin de amplificar no selectivamente y anexar la biblioteca de fragmentos de ADN con adaptadores compatibles con PCR en puente Illumina GAI y secuenciación, se llevó a cabo una PCR usando oligos adaptadores que hibridan en la secuencia del extremo del transposón y no contienen información adicional de secuencia en 3' (FIG. 13).

Reactivo	Volumen
----------	---------

2 X FailSafe™ tampón E de PCR	25 µl
Fragmentos de ADN rotulados en 5' (20 ng/µl)	0,5 µl
BPI-A cebador de PCR (0,5 µM)	1 µl
BP2-ID1-B cebador de PCR (0,5 µM)	1 µl
BPI cebador de PCR (10 µM)	1 µl
BP2 cebador de PCR (10 µM)	1 µl
FailSafe™ enzima de PCR, 2,5 U/µl	1 µl
Agua	24,5 µl
Volumen final:	50 µl

La reacción se incubó de la siguiente manera:

- 72 °C/5:00\*
  - 98 °C/2:00
- 5    • 10 ciclos de (98 °C/0:10, 58 °C/0:30, 72 °C/3:00)
- 4 °C de mantenimiento

\*- A fin de rotular los extremos 3' de los fragmentos de ADN generados con transposición y los fragmentos de ADN con rótulo en 5' con la secuencia del extremo del transposón transferida, se incubaron los productos de reacción con una mezcla de polimerasa de desplazamiento de cadena (FailSafe™) y dNTPs antes de la etapa de desnaturalización (véase la FIG. 8).

Las reacciones de control omitieron los cebadores BP1-A y BP2-ID1-B de PCR y contenían 20 ng de fragmentos de ADN rotulados en 5'.

La reacción de PCR produjo una biblioteca bPCR-compatible con la distribución esperada de MW y era similar a la de los fragmentos de ADN generados por transposición y fragmentos de ADN con rótulo en 5' (Figura 14, carriles 3 y 4). La falta de productos de amplificación detectables en reacciones que carecen de cebadores adaptadores (BP1-A y BP2-ID1-B) es consistente con la amplificación específica de una biblioteca de ADN rotulado en BP1 y BP2 (Figura 14, carril 5).

#### EJEMPLO 12

Secuenciación profunda de una biblioteca de secuenciación compatible con Illumina GAI

A fin de generar una biblioteca de fragmentos de ADN no selectiva que se puede usar directamente en bPCR para secuenciación de Illumina GAI, el ADN genómico CC118 de E. coli se fragmentó y se rotuló en 5' con A/B transposomas. Se usaron oligonucleótidos adaptadores no selectivos para anexar la biblioteca de ADN con adaptadores de bPCR Illumina GAI (FIG. 13).

El ADN genómico CC118 de E. coli se fragmentó, se rotuló en 5' y se purificó tal como se describió en el EJEMPLO D2.

A fin de amplificar no selectivamente y anexar la biblioteca de fragmentos de ADN con adaptadores compatibles con PCR en puente Illumina GAI y secuenciación, se llevó a cabo una PCR usando oligos adaptadores que hibridan en la secuencia del extremo del transposón y no contienen información adicional de secuencia en 3'.

Reactivo	Volumen
2 X FailSafe™ tampón E de PCR	25 µl
Fragmentos de ADN rotulados en 5' (20 ng/µl)	0,5 µl
BP1-A cebador de PCR (0,5 µM)	1 µl



Después de mezclar, la reacción se incubó durante 2 horas a 37 °C. La reacción se detuvo con un volumen igual de solución de detención (15% de sacarosa, 66 mM de EDTA, 20 mM de TRIS, pH 8,0, 0,1% de SDS, 0,9% de Orange G [Sigma 0-7252]), se mezcló y se calentó a 70 °C durante 10 minutos.

- 5 El ADN se analizó por electroforesis en gel de agarosa al 1% en tampón TAE. Los geles se tiñeron con SYBR Gold y ADN se visualizaron con luz no UV.

La fragmentación de ADN objetivo es proporcional a la cantidad de transposoma añadida (FIG. 17).

#### EJEMPLO 15

Fragmentos de ADN circulares rotulados son resistentes a T5 Exonucleasa

- 10 Se aislaron fragmentos de ADN rotulados en 5' del EJEMPLO 14 usando PCR Clean-up NucleoTraP®CR 15 (Macherey-Nagel, GmbH) según las instrucciones del fabricante. A fin de crear fragmentos de ADN circulares rotulados, el ADN recuperado se incubó durante con una polimerasa de no desplazamiento de cadena que carece de actividad de 5' a 3' exonucleasa (por ejemplo T4 ADN polimerasa) y una ligasa dependiente de la plantilla (por ejemplo, ligasa de E. coli) en presencia de dNTPS y β-NAD.

- 15 La siguiente mezcla de reacción se colocó:

14 microlitros de fragmentos de ADN rotulados en 5' del Ejemplo 1

2 microlitros de 10 X EZ-Tn5™ tampón de transposición

1 microlitro 2 mM de β-NAD

1 microlitro 4 mM de dNTPS

- 20 1 microlitro ADN Ligasa de E. coli (10 U/μl)

1 microlitro T4 ADN Polimerasa (0,5 U/μl)

20 microlitros

- 25 La reacción se incubó durante 15 minutos a temperatura ambiente. Las reacciones se detuvieron por incubación durante 20 minutos a 75 °C. Una porción de la reacción (10 μl) se incubó con 10 unidades de T5 Exonucleasa durante 5 minutos a 37 °C para degradar fragmentos de ADN lineales (no circularizados).

Todas las reacciones se trataron con 2 μl de solución de detención (15% de sacarosa, 66 mM de EDTA, 20 mM de TRIS, pH 8,0, 0,1% de SDS, 0,9% de Orange G [Sigma 0-7252]), se mezclaron y se calentaron a 70 °C durante 5 minutos.

- 30 El ADN se analizó por electroforesis en gel de agarosa al 1% en tampón TAE. Los geles se tiñeron con SYBR Gold y el ADN se visualizó con luz no UV.

El tratamiento con T4 ADN polimerasa y ligasa de E. coli convirtió una porción de los fragmentos de ADN en moléculas de fragmentos de ADN circular rotulado resistentes a T5 exonucleasa que se detectan con facilidad (FIG. 18, carril 9). Más aún, la distribución del peso molecular del ADN resistente a T5 exonucleasa es comparable con el ADN de entrada, lo que demuestra la falta de bias de peso molecular en esta reacción.

- 35 También se llevaron a cabo reacciones de control. Cuando los fragmentos de ADN no se trataron, se trataron con T4 ADN polimerasa sola o se trataron con ligasa de E. coli sola, no se detectaron fragmentos de ADN circulares rotulados resistentes a T5 exonucleasa (Figura 18; carriles 3, 5 y 7).

#### EJEMPLO 16

- 40 Rotulación de los extremos 3' de los fragmentos de ADN rotulados en 5' generados por transposición usando una ligasa de ácido nucleico y un oligonucleótido de rotulación de ligamiento

ADN fragmentados rotulados en 5' del procedimiento de tamaño anterior (Ejemplo 1) se usaron para rotular los extremos 3' de los fragmentos de ADN con rótulo en 5' usando una ligasa de ácido nucleico y un oligonucleótido de rotulación de ligamiento (FIG. 21).

- 45 A fin de rotular los extremos 3' de los fragmentos de ADN con rótulo en 5' generados por transposición con un segundo rótulo que comprende el rótulo de secuenciación Roche 454 (4N454B), se llevó a cabo la siguiente reacción.

43 microlitros fragmentos de ADN con rótulo en 5' de tamaño seleccionado de 0,5-1-Kb del EJEMPLO 1

## ES 2 743 029 T3

5 microlitros de 10X tampón de reacción de ligasa (0,2 M de Tris-HCl pH 8,3, 100 mM de MgCl<sub>2</sub>, 250 mM de KCl, 5 mM de β-NAD)

1 micromolar de oligonucleótido 4N454B

(5' pNNNNCTGAGCGGGCTGGCAAGGC 3' (SEQ ID NO: 6))

5 1 microliter de ADN Ligasa de E. coli (10 unidades por microlitro)

Después de mezclar, la reacción se incubó durante 1 hora a temperatura ambiente. A continuación, la reacción se detuvo, la ligasa se inactivó y el ADN se desnaturalizó por incubación a 95 °C durante 5 min. Luego se enfrió la reacción inmediatamente en un baño con agua helada.

10 El análisis de PCR se usó para mostrar que los extremos 5' de los fragmentos de ADN exhibían la secuencia del extremo del transposón transferida del extremo de EZ-transposón Tn5 y los extremos 3' exhibían el rótulo de secuenciación Roche 454 (4N454B).

Se llevó a cabo la siguiente reacción de PCR de la siguiente manera:

21 microlitros Agua

1 microlitro fragmentos de ADN rotulados duales 5' y 3'

1 microlitro cebador de PCR 1 5 micromolar (pMETS)

1 microlitro cebador de PCR 2 5 micromolar (5' ATA GGC GCG CCG CCT TGC CAG CCC GCT CAG 3' (SEQ ID NO:21))

1 microlitro FailSafe™ ADN polimerasa

25 microlitros FailSafe™ 2 X premezcla C de PCR

50 microlitros

La PCR se llevó a cabo para 20 ciclos en las siguientes condiciones:

15 94 °C 10 seg.

55 °C 10 seg.

72 °C 2 min.

El análisis en gel indicó que se produjeron los productos de PCR del rango de tamaño esperado de 0,5-1,0 KB (datos no mostrados).

20 También se llevan a cabo reacciones de control. Cuando la reacción de ligamiento se llevó a cabo sin la ligasa o son el rótulo de secuenciación Roche 454 (4N454B), no se produjeron productos de PCR. Cuando se midió ya sea el cebador de PCR 1 pMETS o cebador de PCR 2 de la reacción de PCR, no se produjeron productos de 0,5 – 1 KB. Cuando se llevó a cabo una reacción de ligamiento sin la ligasa o sin el rótulo de secuenciación Roche 454 (4N454B), no se produjeron productos de PCR. Cuando se omitió el cebador de PCR 1 pMETS o cebador de PCR 2 de la reacción de PCR, no se produjeron productos de 0,5 – 1 KB (datos no mostrados).

25 EJEMPLO 17

Circularización de fragmentos de ssADN rotulados de fragmentación de ADN mediada por transposición in vitro y rotulación en 5' usando la composición del extremo del transposón p454.1MEDS y EZ-Tn5™ Transposasa

30 ADN genómico T7 D111 se rotuló en el extremo 5' y se fragmentó usando la composición de extremo del transposón p454MEDS EZ-Tn5™ en la siguiente reacción:

x agua hasta un volumen final de 50 microlitros

5 microlitros 10 X de tampón de transposición EZ-Tn5™

1 microgramo ADN objetivo en 1 a 40 microlitros

2 microlitros composición de extremo del transposón p454.1MEDS (25 μM)

## ES 2 743 029 T3

<u>2 microlitros</u>	<u>EZ-Tn5™ Transposasa (a 10 unidades por microlitro)</u>
50 microlitros	volumen final de reacción

5 Después de mezclar, la reacción se incubó durante 1 hora a 37 °C. A continuación, la reacción se detuvo con 10 microlitros de solución de detención (15% de sacarosa, 66 mM de EDTA, 20 mM de TRIS, pH 8,0, 0,1% de SDS, 0,9% de Orange G [Sigma 0-7252] y Proteínasa K a 100 microgramos por ml), se mezcló y se calentó a 50 °C durante 10 minutos.

El ADN se analizó por electroforesis en gel de agarosa al 1% en tampón TAE. Los geles se tiñeron con SYBR Gold y el ADN se visualizó con luz no UV.

10 El ADN objetivo se fragmentó en una medida similar y en un rango de tamaño similar al descrito en el Ejemplo 3, usando cantidades y concentraciones comparables de EZTn5™ transposasa y el extremo de transposón pMEDS (FIG. 22B, carril 5). Estos datos indican que la extensión del oligo pMETS con el rótulo de secuenciación Roche 454 adicional no altera significativamente la eficacia de la fragmentación y la rotulación del ADN por EZ-Tn5™ transposasa. La omisión ya sea de la composición de extremo del transposón p454.1MEDS o la EZ-Tn5™ transposasa no dio como resultado una fragmento detectable del ADN (FIG. 22, carriles 3 y 4).

15 El ADN fragmentado rotulado en 5' de la FIG. 22, carril 5, se desnaturalizó con calor y se circularizó usando ligasa independiente de la plantilla (véase la FIG. 22A) en la siguiente reacción:

x	agua hasta un volumen final de 20 microlitros
1 microlitro	330 mM de acetato de Tris pH 7,8, 660 mM de KOAc
1 microlitro	50 mM de MnCl <sub>2</sub>
4 microlitros	5 M de betaína
10 microlitros	20 µg/ml de ADN fragmentado desnaturalizado rotulado en 5'
<u>4 microlitros</u>	<u>100 U/µl de lipasa de ssADN CIRCLIGASE™ (EPICENTRE)</u>
20 microlitros	volumen final de reacción

La reacción se incubó durante 2 horas a 60 °C. A continuación, los productos de reacción se trataron con 18 unidades de Exo I y 20 unidades de Exo III durante 1 hora a 37 °C para eliminar ADN lineal no circularizado

### EJEMPLO 18

20 Análisis de PCR de ssADN circular rotulado

Se realizó un análisis de PCR usando pMETS y pc454.1 como cebadores para demostrar la circularización; se puede amplificar sólo ssADN circular ligado para generar un producto de dsADN lineal que corresponde al tamaño del ssADN circular. La reacción de PCR se llevó a cabo de la siguiente manera:

21 microlitros	agua
1 microlitro	reacción de CircLigasa tratada con exonucleasa (1:1000)
1 microlitro	5 µM de oligonucleótido pMETS como un cebador
1 microlitro	5 µM de oligonucleótido pc454.1 como un cebador
1 microlitro	FailSafe™ ADN polimerasa
<u>25 microlitros</u>	<u>FailSafe™ 2 X premezcla C de PCR</u>
50 microlitros	volumen final de reacción

25 La PCR se llevó a cabo para 29 ciclos, en las siguientes condiciones:

94 °C 10 seg.

50 °C 10 seg.

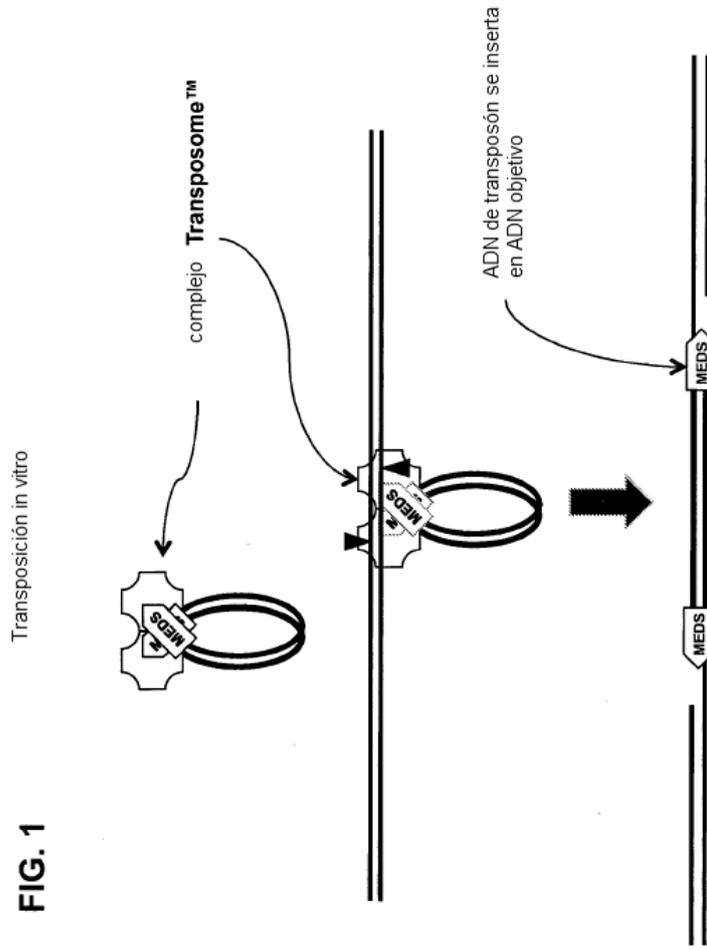
72 °C 1 min.

El análisis en gel indicó que el rango de tamaño de los productos de PCR producidos eran comparables con el ADN fragmentado con rótulo en 5' ADN (FIG. 23).

- 5 También se llevaron a cabo reacciones de control. Cuando se omitió la ssADN Ligasa CIRCLIGASE™ de la reacción de ligamiento, se generaron productos de PCR que indicaron la circularización de la cadena transferida p454.1METS, pero no los fragmentos de ssADN lineales con rótulo en 5' (FIG. 23, carril 1). Cuando se omitió ya sea el oligonucleótido pMETS (como un cebador de PCR) o el cebador de PCR pc454.1 de la reacción de PCR, no se produjeron productos (datos no mostrados).
- 10 El hecho de que los productos de PCR tuvieran la misma distribución de tamaño que los fragmentos de ssADN lineales con rótulo en 5' (comparar la Figura 22, carril 5 y la Figura 23, carril 2) indica que: 1) la composición de extremo del transposón p454.1MEDS puede rotular eficazmente en 5' y fragmentar el ADN objetivo; 2) los fragmentos de ssADN con rótulo en 5' complementario fusionado se pueden desnaturalizar con calor para lograr fragmentos de ssADN lineales rotulados desnaturalizados que son sustratos para el ligamiento independiente de la plantilla; y 3) los
- 15 fragmentos de ssADN lineales con rótulo en 5' se pueden convertir eficazmente en los fragmentos de ssADN circulares rotulados son un bias detectable (confirmado por amplificación de PCR después del tratamiento con exonucleasa I y exonucleasa III).

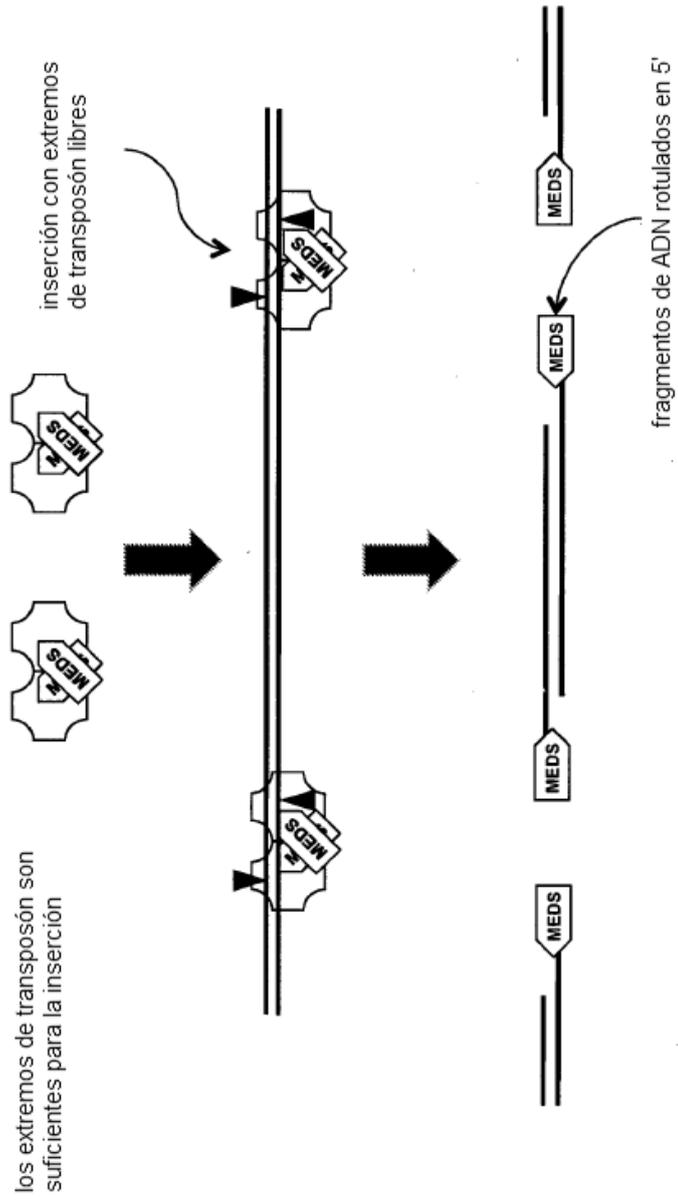
**REIVINDICACIONES**

1. Una composición adecuada para la fragmentación por transposición in vitro que comprende una pluralidad de complejos de transposomas, en donde:
- 5 (1) al menos un complejo de transposoma comprende una transposasa y al menos un primer polinucleótido que comprende una secuencia final de transposón y una primera secuencia de rótulo; y
- (2) al menos un complejo de transposoma comprende una transposasa y al menos un segundo polinucleótido que comprende una secuencia final de transposón y una segunda secuencia de rótulo.
2. Una composición como se reivindica en la reivindicación 1, en donde los complejos de transposomas de (1) y (2) comprenden diferentes secuencias de polinucleótidos.
- 10 3. Una composición como se reivindica en la reivindicación 1 o la reivindicación 2, en donde la primera y la segunda secuencia final del transposón comprenden cada una una cadena transferida y una cadena no transferida, en donde las cadenas transferidas comprenden porciones 5' y porciones 3', y en donde las porciones 5' de las cadenas transferidas difieren entre sí por al menos un nucleótido, y en donde las porciones 3' de las cadenas transferidas comprenden la misma secuencia final del transposón.
- 15 4. Una composición como se reivindica en la reivindicación 3, en donde las cadenas transferidas comprenden dominios de rótulo 5'.
5. Una composición como se reivindica en la reivindicación 4, en donde los dominios de rótulo comprenden uno o más de un dominio de sitio de restricción, un dominio de rótulo de captura, un dominio de rótulo de secuenciación, un dominio de rótulo de amplificación, un dominio de rótulo de detección, un dominio de rótulo de dirección y un dominio
- 20 promotor de la transcripción.
6. Una composición como se reivindica en la reivindicación 5, en donde los dominios de rótulo comprenden un dominio de rótulo de captura, y el dominio de rótulo de captura opcionalmente comprende una porción 5' que está unida a un grupo químico que comprende una molécula de unión por afinidad o una biotina.
7. Una composición como se reivindica en cualquiera de las reivindicaciones 1-6, en donde los extremos del transposón comprenden extremos del transposón Mu y la transposasa es la transposasa Mu.
- 25 8. Una composición como se reivindica en cualquiera de las reivindicaciones 1-6, en donde los extremos del transposón comprenden extremos del transposón Tn5 y la transposasa es una transposasa Tn5.
9. Una composición como se reivindica en la reivindicación 8, en donde la transposasa Tn5 es una transposasa Tn5 hiperactiva.
- 30 10. Un método in vitro para fragmentar y rotular ADN objetivo, que comprende incubar ADN objetivo con una composición como se reivindica en cualquiera de las reivindicaciones 1-9, en donde ADN objetivo se fragmenta y una cadena transferida de la composición se une a un extremo 5' de un fragmento del ADN objetivo para producir un fragmento de ADN objetivo rotulado en 5'.
- 35 11. Un método como se reivindica en la reivindicación 10, que comprende además incubar el fragmento de ADN rotulado en 5' con una ADN ligasa dependiente de plantilla u homóloga y un ligamiento de oligonucleótido de rotulación que comprende un tercer rótulo, en donde el ligamiento de oligonucleótido de rotulación se une a los extremos 3' de los fragmentos de ADN rotulados en 5', generando una biblioteca de fragmentos de ADN rotulados en 5' y 3'.

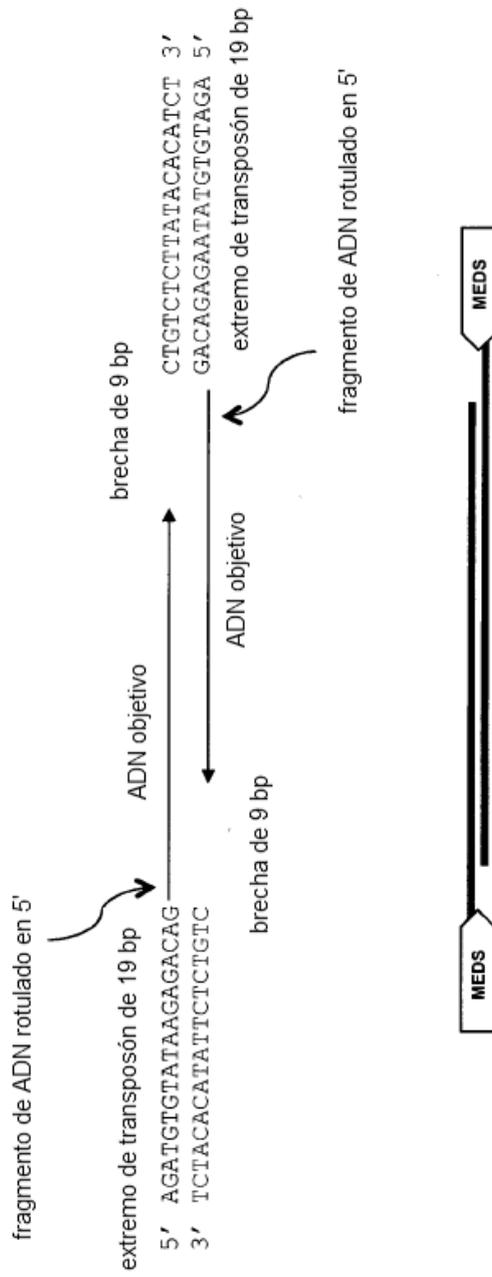


**FIG. 1**

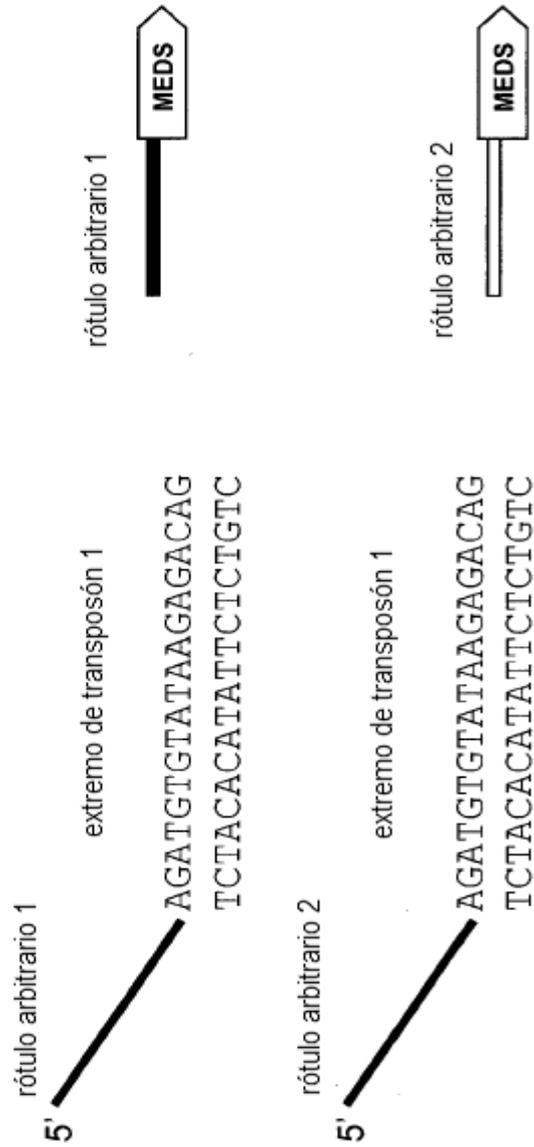
**FIG. 2** Fragmentación por transposición in vitro



**FIG. 3** Producto de rotulación del extremo 5' y fragmentos de dsADN por EZ-Tn5<sup>TM</sup> Transposasa y extremos de transposón de 19 bp

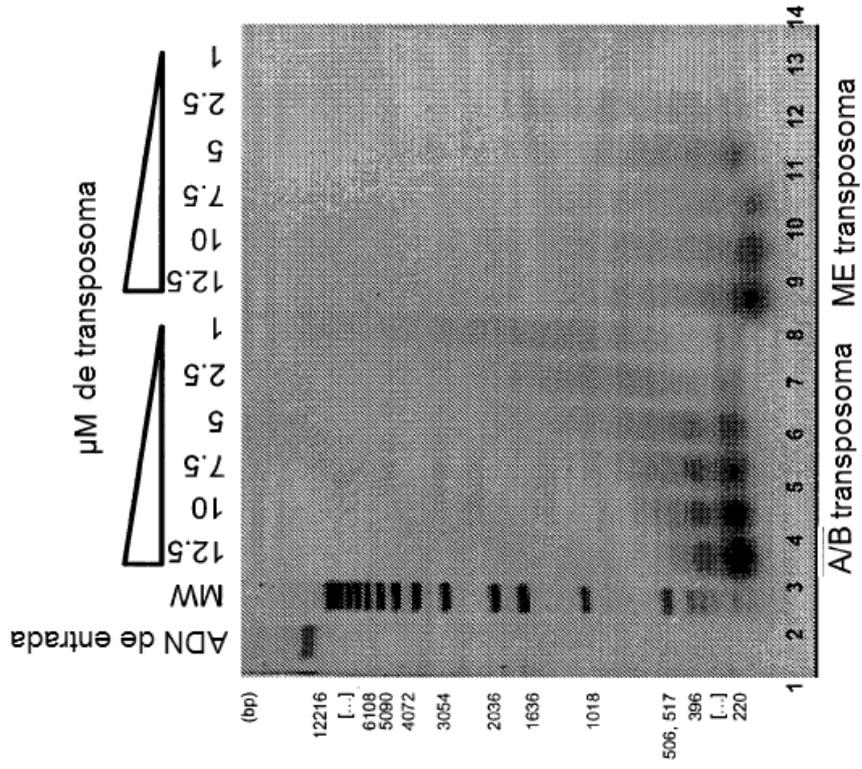


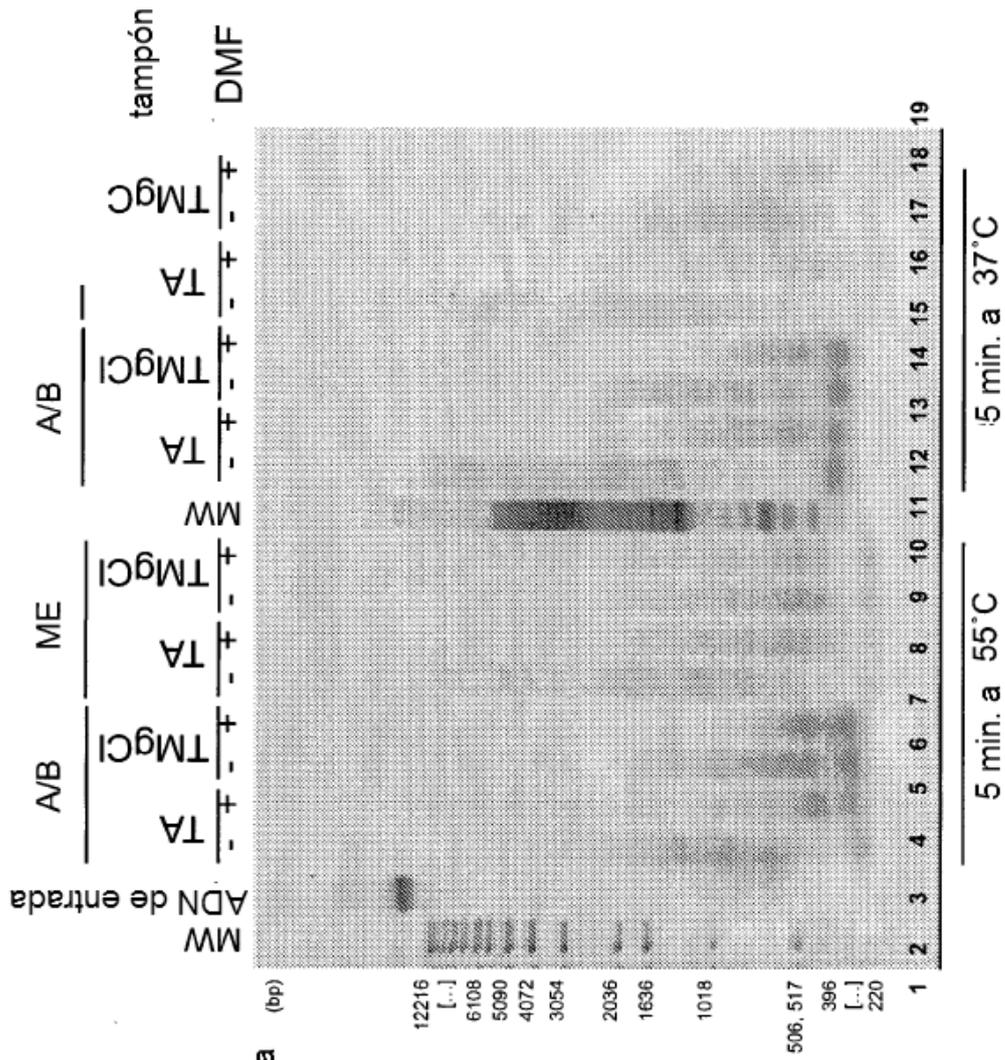
**FIG. 4** Extremos de transposón con oligonucleótidos de extremo de transposón con diferentes rótulos en sus porciones 5'



El primer extremo de transposón y el segundo extremo de transposón comprenden el mismo oligonucleótido de extremo de transposón no transferido, pero los oligonucleótidos de extremo de transposón transferidos exhiben diferentes secuencias de rotulación en sus porciones 5'. Este ejemplo ilustra el uso de una secuencia de extremo de transposón bicatenaria para EZ-Tn5™ Transposasa.

**FIG. 5** Rango de tamaño de productos de transposición de fragmentos de ADN rotulados en 5' usando diferentes concentraciones de transposoma



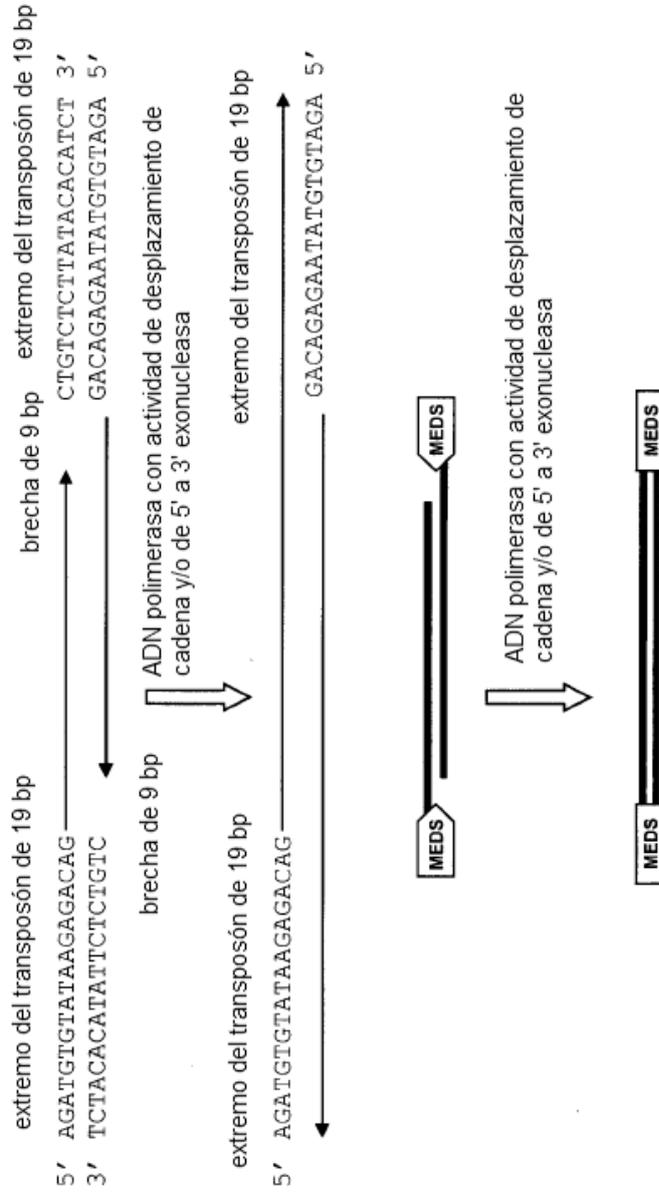


**FIG. 6**

Fragmentación de ADN  
 objetivo y rotulación en 5' a  
 55 °C y a 37 °C en  
 presencia de  
 dimetilformamida

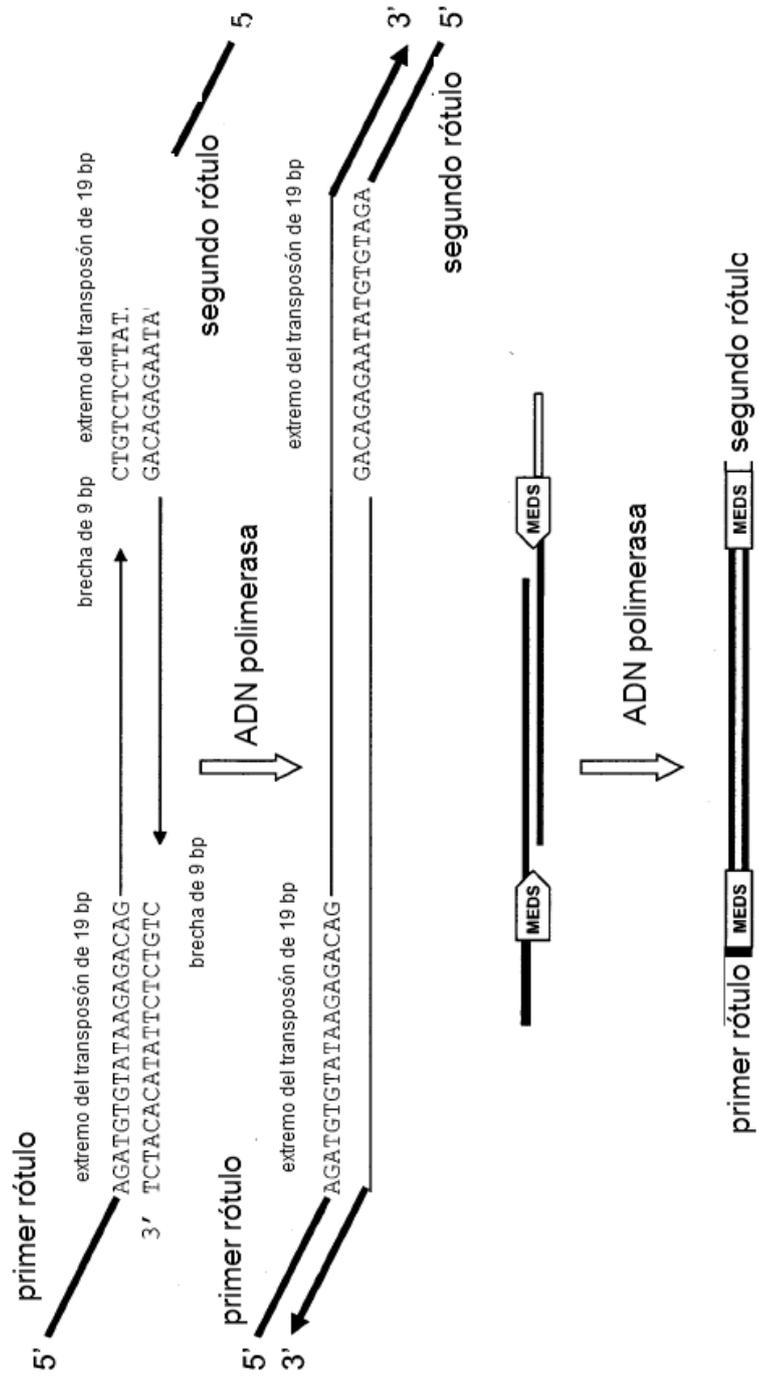
Uso de una ADN polimerasa para rotulación en 3' de fragmentos de ADN rotulados en 5' (5' que comprende o consiste en extremo del transposón transferido EZ-Tn5<sup>TM</sup> de 19 bases)

**FIG. 7**



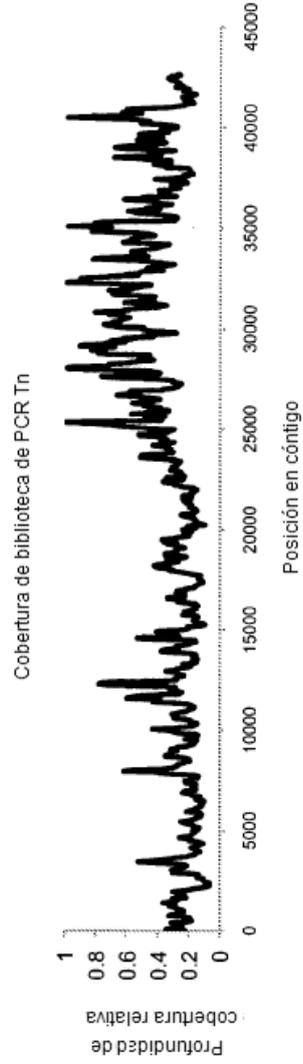
**FIG. 8** Uso de una ADN polimerasa para rotulación en 3' de fragmentos de ADN rotulados en 5'

(con extremos de transposón transferidos EZ-Tn5TM que tienen dos porciones 5' diferentes, cada uno insertado en una cadena de ADN objetivo diferente)



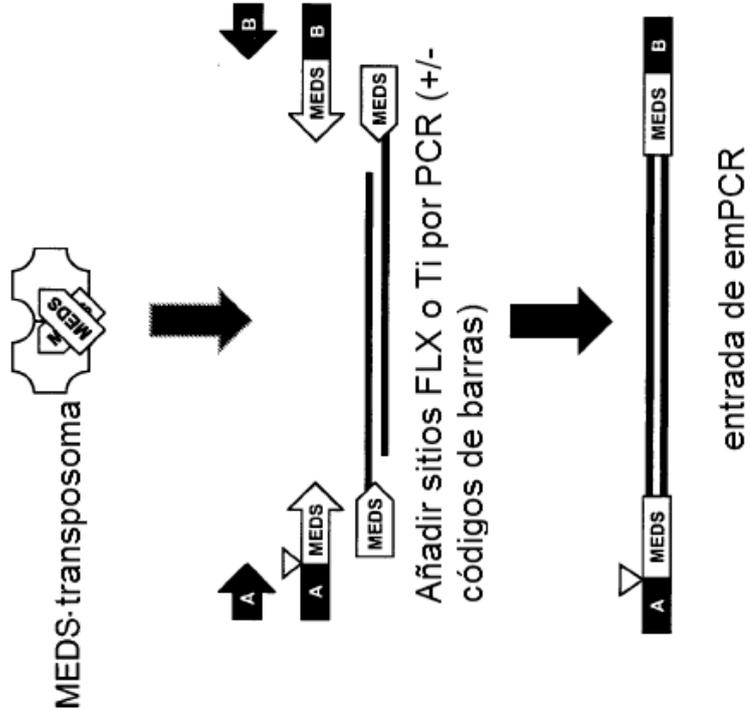
**FIG. 9** Secuenciación profunda de una biblioteca de fragmentos de ADN

	Nebulización	Biblioteca no PCR Tn	Biblioteca PCR Tn
Cantidad de lecturas	19303	26503	24138
Cantidad de bases	4.85 x 10 <sup>6</sup>	5.92 x 10 <sup>6</sup>	5.18 x 10 <sup>6</sup>
% de error	0.48%	0.36%	0.52%
Proporción reunida	0.9758	0.9001	0.8453
Proporción de bases Q40 Plus	0.999	0.998	0.998
Cobertura promedio	110 (20% CV)	109 (37% CV)	147 (30% CV)



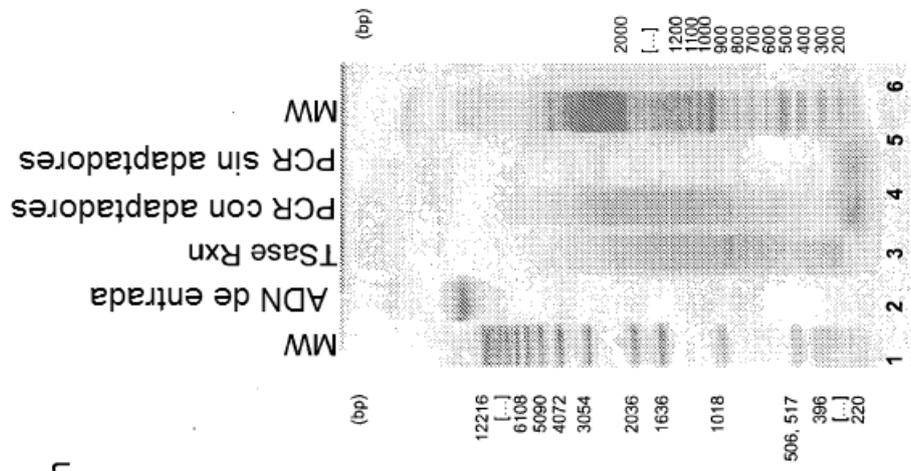
Preparación de bibliotecas de  
secuenciación compatible con Roche/454  
FLX con código de barras

**FIG. 10**

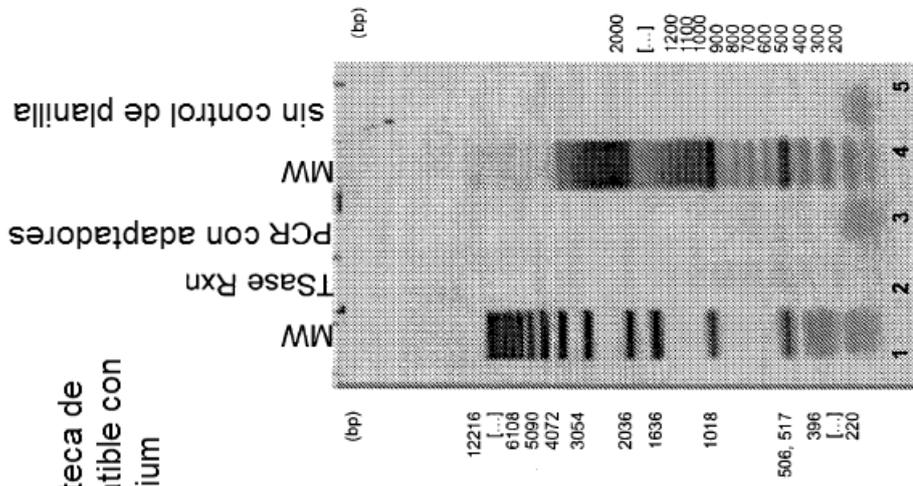


Preparación de bibliotecas de  
secuenciación compatibles con  
Roche/454 FLX con código de  
barras

**FIG. 11**

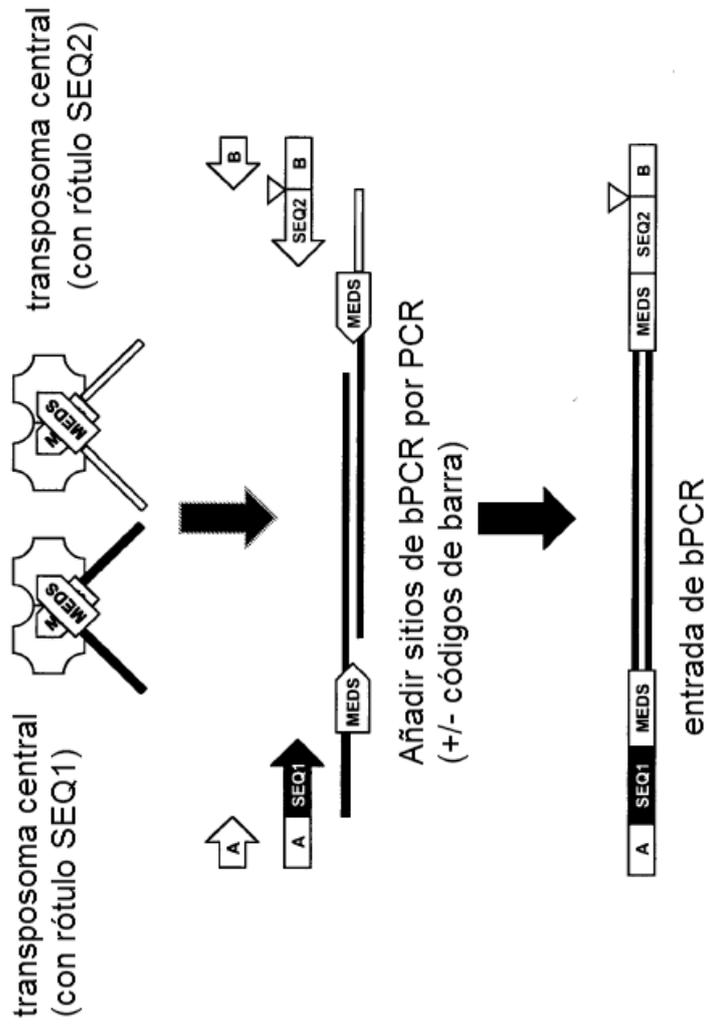


**FIG. 12**  
 Preparación de biblioteca de  
 secuenciación compatible con  
 Roche/454 FLX Titanium



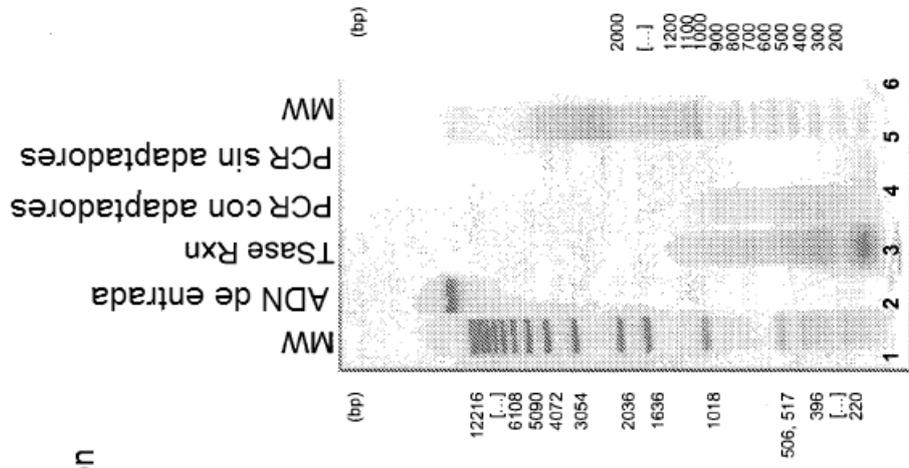
Preparación de bibliotecas de secuenciación compatibles con Illumina GAI con código de barras

FIG. 13



Biblioteca de secuenciación compatible con Illumina GAI con código de barras

FIG. 14

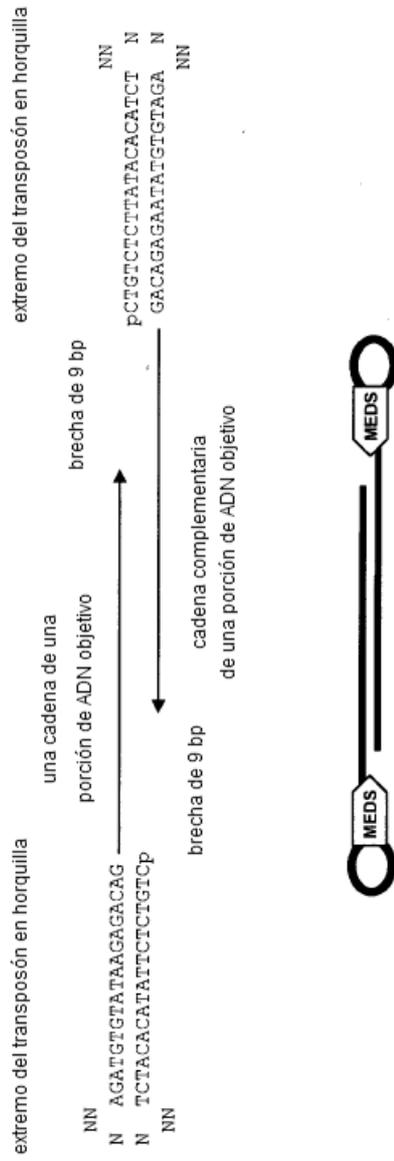


Comparación de los métodos de la técnica anterior con las formas de realización de la presente invención

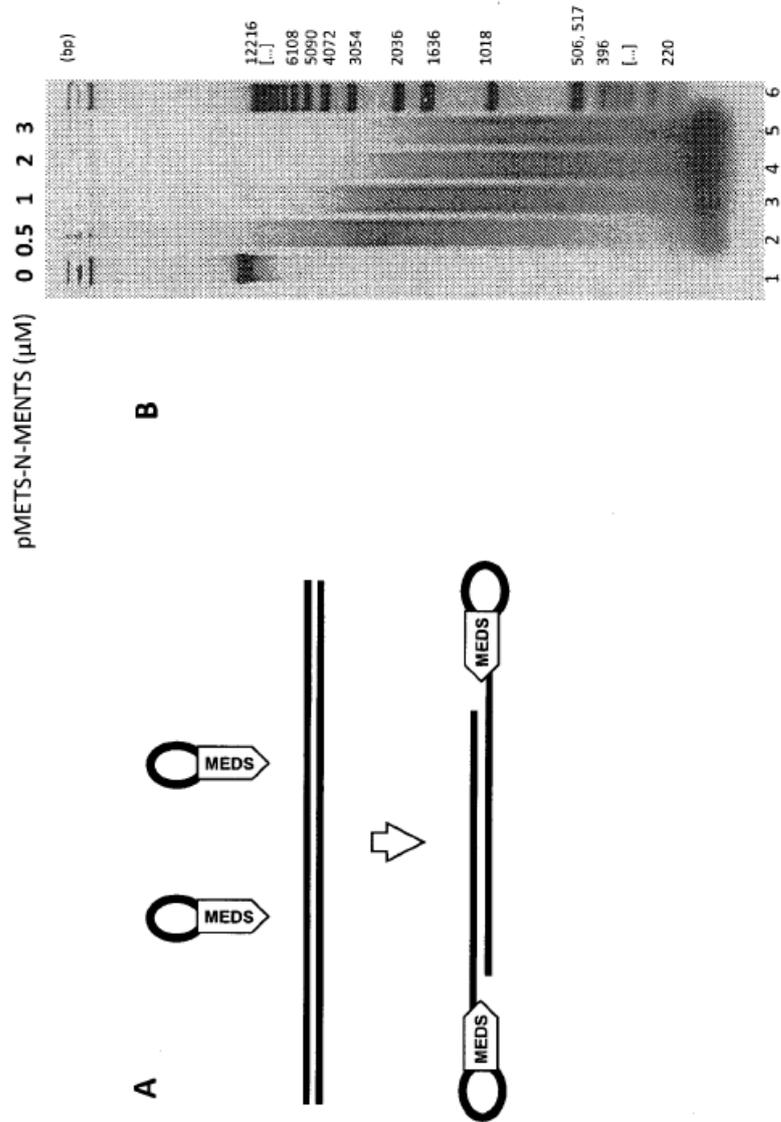
**FIG. 15**

Etapa de procesamiento	Prep biblioteca estándar (~1-5 µg de entrada)	Biblioteca de transposomas Prep (50 ng de entrada)
Fragmentación	✓ (15-30 min)	
Colección	✓ (15 min)	
Concentración	✓ (15 min)	
Selección de tamaño	✓ (60 min)	Añadir transposomas (5 min)
Reparación final	✓ (60 min)	
Limpieza	✓ (15 min)	
Cola A	+/- (30 min)	
Ligación del adaptador	✓ (60 min)	
Limpieza	✓ (15 min)	✓ (15 min)
Enriquecimiento de la biblioteca	Variable (~60 min)	PCR (~60 min)
Tiempo total	~6 horas	<2 horas

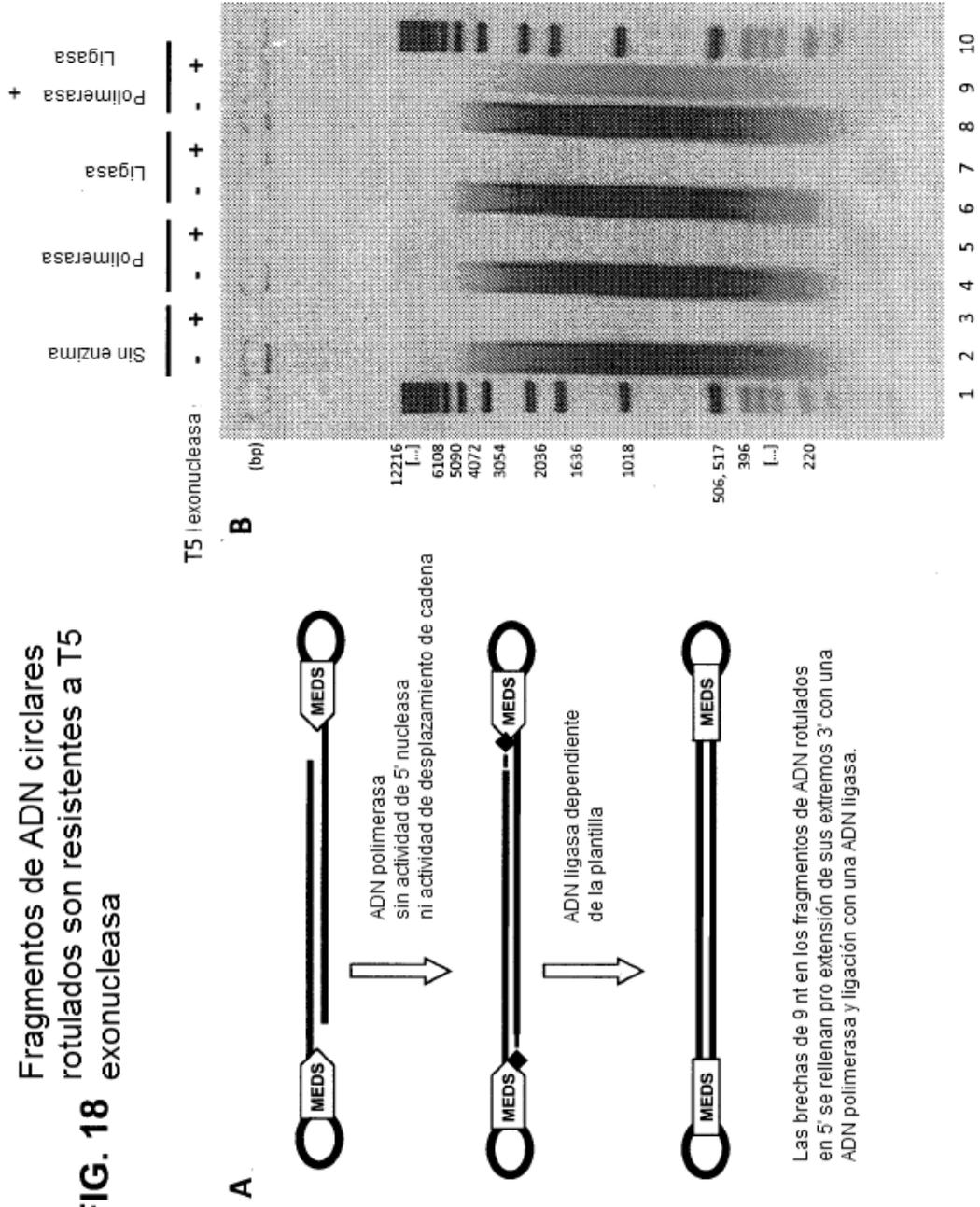
**FIG. 16** Rotulación del extremo 5' y fragmentación de dsADN por EZ-Tn5TM transposasa y composición de extremo del transposón en horquilla



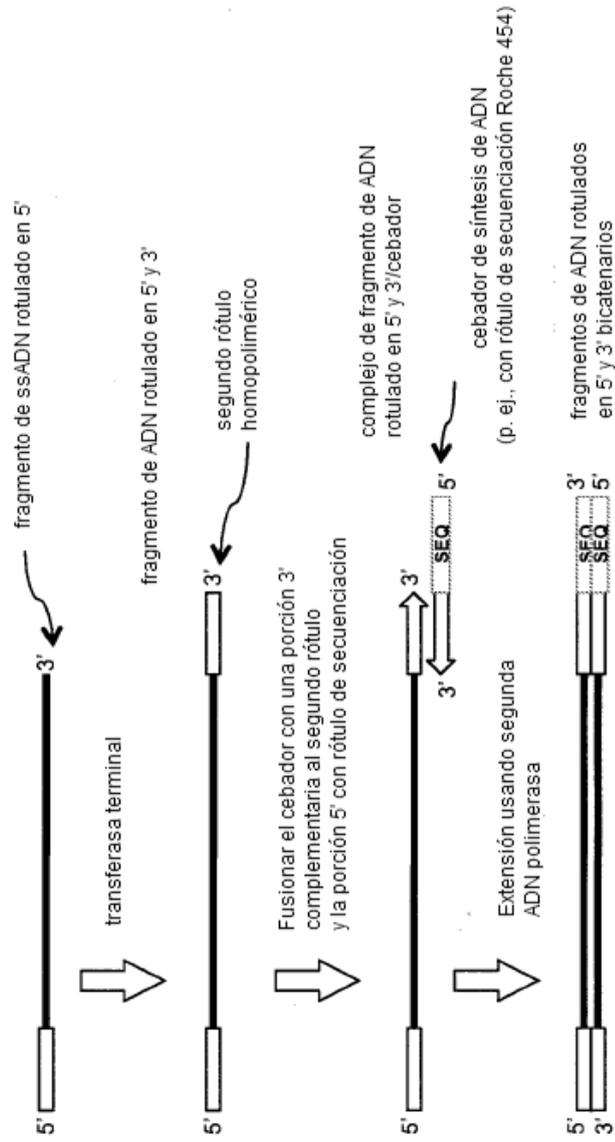
**FIG. 17** Rotulación en 5' y fragmentación de dsADN por una transposasa y una composición de extremo del transposón en horquilla



**FIG. 18** Fragmentos de ADN circulares rotulados son resistentes a T5 exonucleasa

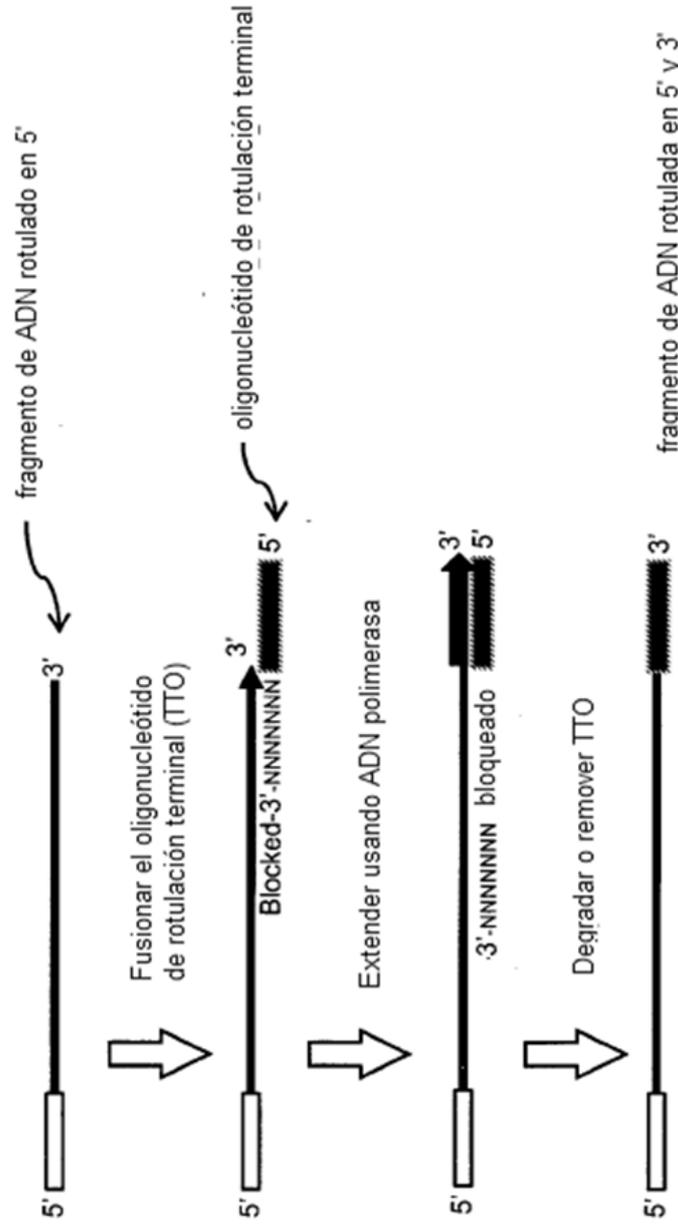


**FIG. 19** Unión de un rótulo de secuenciación con fragmentos de ADN rotulados en 5' u 3' que tienen un segundo rótulo usando una segunda ADN polimerasa



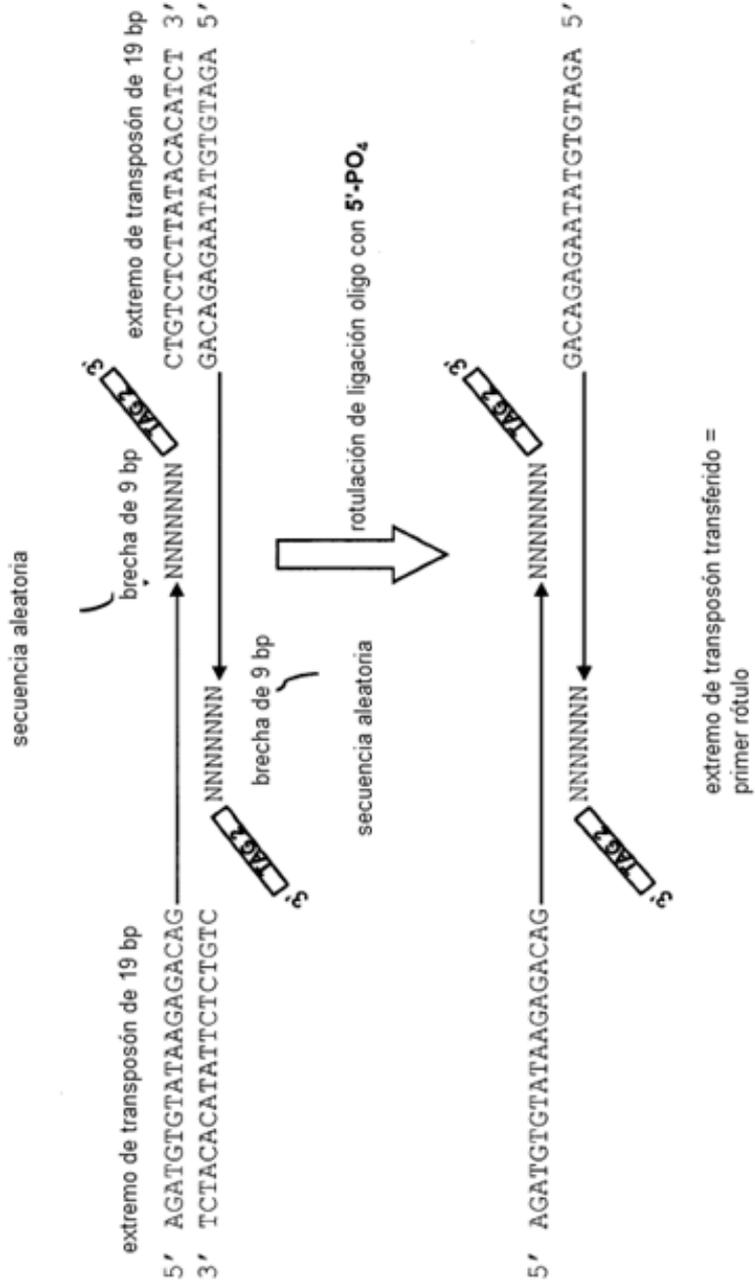
Por ejemplo, si el primer rótulo comprende un rótulo de secuenciación Roche/454A, el segundo rótulo puede comprender el rótulo de secuenciación Roche/454B.

**FIG. 20** Uso de un oligonucleótido de rotulación terminal para generar fragmentos de ADN rotulados en 5' y 3'



Los fragmentos de ADN rotulados en 5' pueden tener un primer rótulo que comprende o consiste en un oligonucleótido de extremo de transposón transferido que exhibe una secuencia de rótulo arbitraria en su porción 5'. Por ejemplo, si el primer rótulo comprende un rótulo de secuenciación Roche 454A, el segundo rótulo puede comprender el rótulo de secuenciación 454B.

**FIG. 21** Uso de una ADN ligasa y una rotulación de ligación oligo para generar fragmentos de ADN rotulados en 5' y 3'





**FIG. 23** Análisis de PCR de ADN circular rotulado

