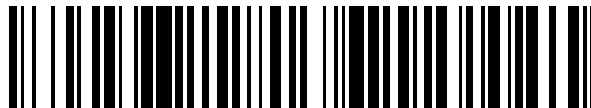


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 717 756**

51 Int. Cl.:

C12Q 1/68 (2008.01)

C12N 15/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **29.09.2014 PCT/EP2014/070769**

87 Fecha y número de publicación internacional: **02.04.2015 WO15044412**

96 Fecha de presentación y número de la solicitud europea: **29.09.2014 E 14781126 (9)**

97 Fecha y número de publicación de la concesión europea: **09.01.2019 EP 3052645**

54 Título: **Moléculas adaptadoras de ADN para la preparación de genotecas de ADN y método para su producción y uso**

30 Prioridad:

30.09.2013 EP 13186776

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

25.06.2019

73 Titular/es:

**QIAGEN GMBH (100.0%)
Qiagen Str. 1
40724 Hilden, DE**

72 Inventor/es:

**HAHN, PETER;
AZZAWI, ALEXANDER y
GRÜNEFELD, PETER**

74 Agente/Representante:

ELZABURU, S.L.P

ES 2 717 756 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

Moléculas adaptadoras de ADN para la preparación de genotecas de ADN y método para su producción y uso

5 La invención se refiere a moléculas adaptadoras de ADN para la preparación de genotecas de ADN y a métodos para su producción y su uso. La invención es útil para la aplicación en biología molecular, en particular para la secuenciación de nueva generación y/o la multiplicación de genotecas (del inglés, "Next Generation Sequencing" y/o "Library Multiplexing").

Estado de la técnica

10 La secuenciación de nueva generación (NGS) es una tecnología moderna para determinar la secuencia de nucleobases de una secuencia de ADN o ARN. Las ventajas de la NGS (velocidad, eficacia económica, exactitud) están dando lugar a una distribución cada vez mayor en todas las instituciones académicas que tienen que ver con el análisis de información genética y su aplicación. Varias compañías mundiales ofrecen soluciones para la NGS, incluyendo Roche, Life Technologies e Illumina.

15 Para la secuenciación de una serie de muestras en una ejecución, se pueden añadir códigos de barras con una secuencia conocida a las muestras definidas y utilizarlos para asignar secuencias a las muestras después de la secuenciación.

Los métodos, composiciones y kits para la secuenciación múltiple (secuenciación simultánea de varias muestras diferentes) se describen en el documento WO 2011156529 A2. Una pluralidad de polinucleótidos diana procedentes de dos o más muestras diferentes, se secuencian en una cámara de reacción y la muestra a partir de la cual se obtiene cada uno de los polinucleótidos diana secuenciados, se identifica a través de códigos de barras.

20 Otro documento que describe métodos de secuenciación con un rendimiento mejorado de las muestras es WO 2008061193 A2. En ese documento, secuencias adaptadoras (denominadas secuencias marcadoras) están ligadas a muestras que a continuación se mezclan y se secuencian.

El documento WO 2013033721 A1 describe métodos para optimizar el diseño de códigos de barras para la secuenciación múltiple de ADN.

25 El documento US 2013059762 A1 proporciona métodos, composiciones, kits, sistemas y aparatos que son útiles para una PCR múltiple de uno o varios ácidos nucleicos presentes en una muestra. En particular, se proporcionan diversos cebadores específicos de la diana que permiten la amplificación selectiva de una o varias secuencias diana. Por lo tanto, los adaptadores se ligan a secuencias diana en una reacción de ligación con extremos romos.

30 Durante la preparación de genotecas de ADN para la secuenciación de nueva generación (NGS) un ADNds de alto peso molecular en la mayoría de los casos se fragmenta mediante métodos físicos o químicos y secuencias adaptadoras específicas de la la plataforma de secuenciación se ligan a las moléculas de ADN generadas. Las secuencias adaptadoras contienen en general sitios de unión para cebadores y podrían contener secuencias de códigos de barras, que permiten secuenciar múltiples genotecas de ADN con códigos de barras en una mezcla (multiplexación) y asignar posteriormente la información de la secuenciación del código de barras respectivo a la muestra original.

35 Antes de que se pueda llevar a cabo esta ligación de las secuencias adaptadoras con los fragmentos, las terminaciones del ADN, que se forman al azar y son imprevisibles durante la fragmentación, se tienen que reparar. Los extremos 3' sobresalientes se escinden y los extremos 5' sobresalientes se rellenan hasta formar una hebra doble mediante una polimerasa. Posteriormente, los extremos 5' de los fragmentos se fosforilan con una cinasa. Después, los adaptadores de ADN no fosforiladas, que tienen extremos romos en un lado para permitir la ligación en ese lugar y que tienen un extremo 3' sobresaliente en el otro lado para evitar una ligación en ese lugar, se ligan a las moléculas de ADN generadas con la ayuda de una ligasa. Dado que la ligación de secuencias adaptadoras no fosforiladas no da lugar a la formación de enlaces fosfodiéster en el extremo 5' del adaptador, dejando de ese modo una muesca en la hebra doble de ADN después de la ligación, el enlace fosfodiéster que falta se tiene que cerrar mediante una enzima, que sea capaz de realizar un desplazamiento de muesca.

40 Las secuencias adaptadoras utilizadas comúnmente no se pueden añadir a la mezcla de reparación de los extremos, ya que las enzimas en la mezcla de reparación de los extremos las modificarían de manera que la direccionalidad de la ligación se perdería y se formarían dímeros y oligómeros de adaptadores.

Objetivo

50 El objetivo de la presente invención es modificar las secuencias adaptadoras de manera que se puedan añadir al ADN fragmentado al comienzo de la preparación de la genoteca. Otro objetivo de la invención es evitar la formación de dímeros y oligómeros de adaptadores.

Invención

La presente invención describe moléculas adaptadoras de ADN, que comprenden una molécula de polinucleótido de hebra doble; en la que

- 5 a. el extremo 5' de la primera hebra se modifica de manera que el primer nucleótido no contiene un grupo hidroxilo libre ni un fosfato libre en la posición 5', de modo que no está disponible un sitio de unión para las cinasas,
- b. el extremo 3' de la primera hebra se modifica de manera que el último nucleótido no contiene un grupo hidroxilo libre en la posición 3', de modo que no se puede producir una ligación,
- c. el extremo 5' de la hebra inversa se modifica de manera que el primer nucleótido no contiene un grupo hidroxilo libre ni un fosfato libre en la posición 5', de modo que no está disponible un sitio de unión para las cinasas, y
- 10 d. el extremo 3' de la hebra inversa presenta un grupo hidroxilo libre (en la posición 3' del último nucleótido).

La primera hebra y la hebra inversa se asocian entre sí mediante apareamiento de bases complementarias, sin ningún extremo sobresaliente (extremos romos). Preferiblemente, los nucleótidos dentro de la hebra doble, distintos de los nucleótidos terminales mencionados anteriormente en los extremos 3' y 5', no están modificados. Sin embargo, no se excluye que la hebra doble contenga nucleótidos con una estabilidad mejorada frente a las nucleasas, en particular nucleótidos con una estructura principal modificada, como los fosforotioatos. Preferiblemente, ambas hebras tienen exactamente la misma longitud (con un número idéntico de nucleótidos), en donde la secuencia de nucleótidos de la hebra inversa es el complemento inverso de la secuencia de nucleótidos de la primera hebra.

Ventajosamente, las moléculas adaptadoras de ADN de acuerdo con la invención se pueden añadir durante la preparación de genotecas de ADN, directamente al ADN fragmentado. Al contrario que en el estado de la técnica, ahí no hay necesidad de inactivar las enzimas de reparación de los extremos antes de la adición de las moléculas adaptadoras de ADN de acuerdo con la invención y antes de la adición de la ligasa. Puesto que las moléculas adaptadoras de acuerdo con la invención tienen extremos romos, no son un sustrato para las polimerasas con actividad polimerasa 5'→3' y exonucleasa 3'→5', como la ADN polimerasa de T4, la ADN polimerasa de T7 o la polimerasa Pfu.

Además, las moléculas adaptadoras de ADN de acuerdo con la invención están diseñadas ventajosamente de manera que la ligación enzimática solo puede tener lugar en una dirección (en el grupo hidroxilo libre del extremo 3' de la hebra inversa).

La expresión "grupo hidroxilo libre" tal y como se utiliza en esta memoria, se refiere a -OH o el -O⁻ desprotonado. La expresión "fosfato libre" se utiliza en esta memoria para describir un fosfato desprotonado (-OPO₃²⁻), monohidrogenofosfato o dihidrogenofosfato.

El término "polimerasa" tal y como se utiliza en esta memoria, se refiere a una polimerasa de ADN dependiente de ADN, preferiblemente una polimerasa con actividad polimerasa 5'→3' y exonucleasa 3'→5' y preferiblemente sin actividad exonucleasa 5'→3'. Las polimerasas preferidas son enzimas, que son capaces de crear extremos romos, como la polimerasa de T4, la ADN polimerasa de T7 o la polimerasa Pfu.

El término "cinasa", tal y como se utiliza en esta memoria, se refiere a una polinucleótido 5'-hidroxil-cinasa que cataliza la adición de un fosfato a un extremo 5' hidroxilo libre de una molécula de polinucleótido. Las cinasas preferidas son la cinasa de T4 o la cinasa de T7.

Una "molécula de polinucleótido" tal y como se utiliza en esta memoria, es un biopolímero compuesto de 13 o más monómeros de nucleótidos unidos covalentemente en una cadena.

La molécula adaptadora de ADN de acuerdo con la invención contiene preferiblemente de forma adicional sitios de unión para cebadores de la amplificación y secuenciación y, preferiblemente, una secuencia de código de barras.

La secuencia de código de barras tiene preferiblemente una longitud de tres a 20, más preferiblemente de cuatro a ocho pares de bases. Preferiblemente, la secuencia de código de barras está situada cerca del extremo 5' de la primera hebra. La molécula adaptadora de ADN tiene preferiblemente una longitud de 20 a 90, más preferiblemente de 30 a 70, lo más preferido de 40 a 60 pares de bases.

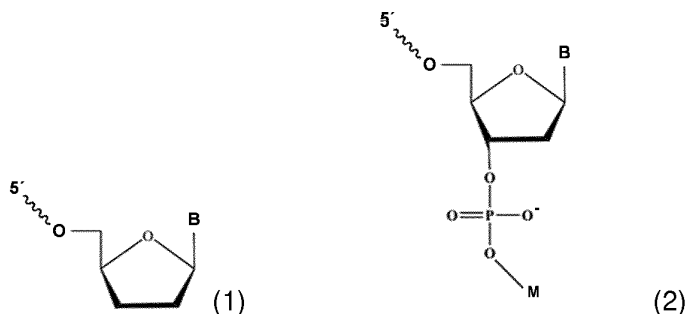
Con la expresión, el primer y/o el último nucleótido no contiene un grupo hidroxilo libre, se entiende que la molécula adaptadora de ADN de acuerdo con la invención se modifica de una manera que el hidroxilo o el fosfato que normalmente está presente en los extremos 5' de ambas hebras, y el hidroxilo que normalmente está presente en el extremo 3' de la hebra inversa, se reemplazan por otros grupos, denominados en esta memoria "grupos terminadores", que dan lugar a un nucleótido modificado que no es un sustrato para las enzimas respectivas mencionadas anteriormente. Preferiblemente, esos grupos terminadores se seleccionan independientemente a partir de hidrógeno, alquilo sustituido o no sustituido, alcoxi amino y otros terminadores de cadena conocidos. No se excluye que esos grupos terminadores pueden contener por sí mismos grupos hidroxilo libres o incluso un fosfato

libre, ya que estos no serán sustratos para las enzimas.

La molécula adaptadora de ADN de acuerdo con la invención se modifica preferiblemente de modo que los extremos 5' de ambas hebras y el extremo 3' de la primera hebra, contienen terminadores de cadena.

- 5 Los terminadores de cadena para el extremo 3' así como para el extremo 5', son bien conocidos por una persona experta en la técnica. En el extremo 5' de la hebra inversa y el extremo 3' de la primera hebra, se puede elegir cualquier terminador de cadena posible para bloquear el 3'-OH libre o el 5'-OH libre, respectivamente.

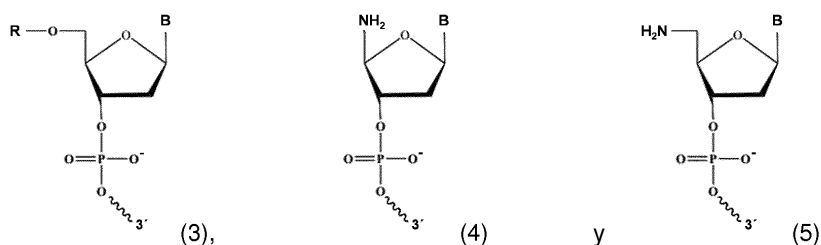
Ejemplos preferidos de terminadores de cadena para el extremo 3' de la primera hebra son 2'3'-didesoxinucleósidos (fórmula 1) o desoxinucleótidos modificados en 3', preferiblemente tal y como se representan en la fórmula 2:



- 10 con B = nucleobase, se selecciona preferiblemente a partir de Adenina (A), Guanina (G), Citosina (C) y Timina (T), y M se selecciona a partir de NH₂-L, en donde L se selecciona preferiblemente a partir de alquilo C1 a C6 lineal o ramificado, preferiblemente alquilo C3 o C4,

en donde 5' ~~~~~O es el enlace covalente con la secuencia de nucleótidos en desarrollo de la primera hebra de la molécula adaptadora de ADN en dirección 3' → 5'.

- 15 Ejemplos preferidos de terminadores de cadena para el extremo 5' son desoxinucleótidos eterificados (fórmula 3), 4'-amino-desoxinucleótidos (fórmula 4) y 5'-amino-desoxinucleótidos (fórmula 5), tal como:

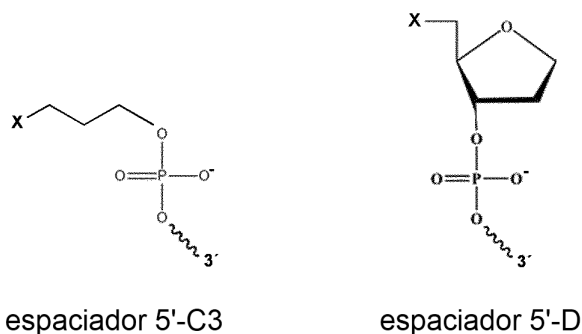


con B = nucleobase seleccionada como anteriormente y preferiblemente R = alquilo C1 a C3, más preferiblemente CH₃,

- 20 en donde O~~~~3' es el enlace covalente a la secuencia de nucleótidos en desarrollo de la primera hebra o la hebra inversa de la molécula adaptadora de ADN en dirección 5' → 3'.

- 25 Preferiblemente, los extremos 5' de ambas hebras se modifican independientemente uno de otro mediante 5'-OMe-desoxinucleótidos, la más preferida 5'-OMe desoxitimidina (5'-OMeT), modificaciones terminales del espaciador 5'-C3 y/o del espaciador 5'-D. Esta modificación da como resultado de hecho, que la fosforilación mediante cinasas de polinucleótidos, especialmente a través de la polinucleótido cinasa de T4, no se puede llevar a cabo en el extremo 5' modificado. Ventajosamente, los extremos 5' de ambas hebras permanecen sin fosforilar para evitar la formación de dímeros y oligómeros de adaptadores.

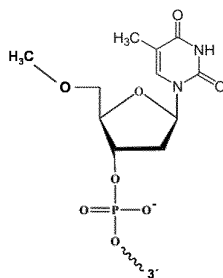
Preferiblemente, un espaciador 5'-C3 o un espaciador 5'-D se diseña del modo siguiente



en donde X se selecciona a partir de -H, -OH, -OR_x y halógeno, preferiblemente X es -OH u -OR_x, más preferiblemente X es -OH,

5 en donde R_x es un residuo de alquilo C1 a C3 opcionalmente sustituido y/o ramificado, más preferiblemente R_x es CH₃,

en donde $\text{O} \sim \sim \sim 3'$ es el enlace covalente con la secuencia de nucleótidos en desarrollo de la primera hebra o de la hebra inversa de la molécula adaptadora de ADN en dirección 5' → 3'.



10 5'-OMe desoxitimilato (monofosfato de 5'-O-metil desoxitimidina, al que también se hace referencia en esta memoria como 5'-OMeT).

Preferiblemente, el extremo 5' de la primera hebra de la molécula adaptadora de ADN se modifica con espaciador 5' o un desoxinucleótido eterificado en 5' (preferiblemente según la fórmula 3), preferiblemente un 5'-Ometil desoxinucleótido. El término espaciador significa un residuo de hidrocarburo preferiblemente con uno a seis átomos de carbono, preferiblemente un grupo alquidilo con 2 a 4 átomos de carbono, lo más preferido C3 lineal (espaciador 5'-C3). La modificación del extremo 5' de la primera hebra de la molécula adaptadora de ADN da lugar de hecho a que la fosforilación, en particular a través de la polinucleótido cinasa de T4, no sea posible. Además, las modificaciones preferidas tienen la menor influencia (impedimento estérico) sobre la conformación de la hebra doble de ADN para permitir una ligación eficaz del 3'-OH libre reactivo de la hebra inversa.

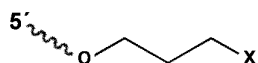
El 5'-OMe desoxinucleótido más preferido es 5'-OMe desoxitimilato (5'-OMeT). También son posibles modificaciones mediante 5'-OMe desoxicitidilato (monofosfato de 5'-O-metil desoxicitidina), 5'-OMe desoxiadenilato (monofosfato de 5'-O-metil desoxiadenosina) o 5'-OMe desoxiguanilato (monofosfato de 5'-O-metil desoxiguanosina). Para la modificación de un 5'-OMe desoxinucleótido (o un 5'-amino-desoxinucleótido), se añade un desoxinucleótido complementario adicional en el extremo 3' de la hebra inversa de la molécula adaptadora de ADN, de modo que el nucleótido modificado en el extremo 5' de la primera hebra puede formar una pareja de bases normal, no estabilizada con ese nucleótido complementario para no interferir en la siguiente reparación de muesca. Si por ejemplo, la modificación del extremo 5' de la primera hebra es 5'-OMeTimidina (o monofosfato de 5'-aminodesoxitimidina), el nucleótido complementario es una desoxiadenina.

Cuando para la preparación de la genoteca se emplea la molécula adaptadora de ADN de acuerdo con la invención, se liga solamente con el extremo 3' de la hebra inversa con un extremo fosforilado en 5' de la primera hebra de un fragmento de ADN. Entre el extremo 5' de la primera hebra del adaptador de ADN y el 3'-OH de la hebra inversa del fragmento de ADN, sigue habiendo una muesca (debido a la modificación del extremo 5' de la primera hebra de la molécula adaptadora de ADN, el extremo 5' no contiene un fosfato libre). La muesca se repara de forma preferible posteriormente mediante una reparación por escisión de nucleótidos (también denominada en esta memoria reparación de muesca) con una polimerasa con actividad exonucleasa 5' → 3' y una actividad polimerasa 5' → 3' (como la ADN polimerasa I) escindiendo y sustituyendo el nucleótido modificado en el extremo 5' (preferiblemente 5'-OMeT) por un nucleótido sin modificar (preferiblemente monofosfato de desoxitimidina). La ligasa que preferiblemente aún

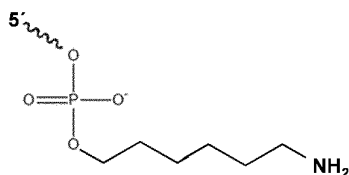
está presente y activa en la muestra, liga a continuación el grupo 3'-hidroxilo libre de la hebra inversa del fragmento de ADN con el extremo 5' reparado de la primera hebra del adaptador de ADN.

El extremo 3' de la primera hebra de la molécula adaptadora de ADN de acuerdo con la invención, se modifica preferiblemente con un espaciador 3'-C3 o un modificador 3'-amino. Este extremo 3' modificado está bloqueado para la ligasa y no puede tener lugar una unión a un ADN fosforilado en 5', preferiblemente moléculas de ADNds.

5



espaciador 3'-C3



modificador 3'-amino,

en donde X se selecciona a partir de -H, -OH, -OR_x y halógeno, preferiblemente X es -OH u -OR_x, más preferiblemente X es -OH,

en donde R_x es un residuo de alquilo C1 a C3 opcionalmente sustituido y/o ramificado, más preferiblemente R_x es CH₃,

10 en donde 5' O es el enlace covalente con la secuencia de nucleótidos en desarrollo de la primera hebra de la molécula adaptadora de ADN en dirección 3'→5'.

También forma parte de la invención un método para la producción de moléculas adaptadoras de ADN, que comprende una molécula de ADN de hebra doble con las siguientes etapas:

15 a. modificación del extremo 5' de la primera hebra de manera que no está disponible ningún sitio de unión para la cinasa, especialmente ningún grupo hidroxilo libre ni ningún fosfato libre,

b. modificación del extremo 3' de la primera hebra de manera que no está disponible ningún grupo hidroxilo libre y no se puede formar ningún enlace con ADNds fosforilado en 5',

c. modificación del extremo 5' de la hebra inversa de manera que no está disponible ningún sitio de unión para una cinasa, especialmente ningún grupo hidroxilo libre ni ningún fosfato libre,

20 con lo que las etapas a-c pueden tener lugar en un orden arbitrario.

Las hebras dobles se obtienen normalmente mediante la asociación de dos hebras sencillas, sintetizadas mediante síntesis de oligonucleótidos convencional. En general, las modificaciones mencionadas anteriormente se introducen mediante el uso de componentes principales ya modificados previamente para la síntesis de oligonucleótidos para los extremos 3' y 5' respectivos. Sin embargo, el término modificación también incluye la posibilidad (menos preferida) de modificar los extremos 3' y 5' respectivos después de la síntesis de oligonucleótidos, por ejemplo, la transformación de una 5'-timidina convencional (más exactamente, monofosfato de desoxitimidina) con un 5'-OH libre en un éter de metilo para obtener 5'-OMeT.

25

Las modificaciones preferidas se han descrito anteriormente.

Otra parte de la invención es un método para la generación de una genoteca de ADN que comprende preferiblemente las siguientes etapas:

30

1. fragmentación de ADN de hebra doble, preferiblemente por métodos físicos o químicos o enzimáticos o una combinación de los mismos,

2. reparación de los extremos del ADN de hebra doble fragmentado, de preferencia con la ayuda de una ADN polimerasa con actividad exonucleasa 3'→5' y polimerasa 5'→3', tal como la ADN polimerasa de T4,

35

3. fosforilación de los extremos 5' del ADN de hebra doble fragmentado, de preferencia con una polinucleótido cinasa, tal como la polinucleótido cinasa de T4,

4. ligación de las moléculas adaptadoras de ADN de acuerdo con la invención, con el ADN de hebra doble, fragmentado, con extremos reparados y fosforilado en 5'; con lo que las moléculas adaptadoras de ADN de acuerdo con la invención se añaden antes de la etapa 2 y preferiblemente no se realiza una inactivación de las enzimas (en particular, las enzimas de reparación de los extremos) antes de la ligación.

40

Las enzimas se añaden en tampones adecuados, conocidos en el estado de la técnica. Las etapas 2 y 3 se pueden realizar de forma paralela, en particular, mediante la adición de una mezcla de polimerasa y cinasa. Ambas enzimas en conjunto también se denominan en este documento "enzimas de reparación de los extremos". La mezcla también se denomina "mezcla de enzimas de reparación de los extremos", respectivamente. La polimerasa utilizada para la

5 reparación de los extremos y la cinasa funcionan bien en el mismo tampón, preferiblemente que comprende magnesio y un agente tamponador, como TRIS-HCl y preferiblemente a un pH entre 7 y 8. Este tampón también se denomina adicionalmente "tampón de reparación de los extremos". Para la reparación de los extremos también se añaden desoxinucleósidos trifosfatos (preferiblemente dATP, dTTP, dGTP y dCTP) o bien al "tampón de reparación de los extremos" o por separado.

10 Las moléculas adaptadoras de ADN de acuerdo con la invención se construyen como se ha explicado anteriormente, preferiblemente incluyendo una o varias características preferidas descritas anteriormente. Preferentemente las moléculas adaptadoras de ADN añadidas antes de la etapa 2, contienen secuencias de códigos de barras. En la etapa 1, varias sondas de ADN individuales se pueden tratar en paralelo en compartimentos individuales de un dispositivo. Preferiblemente, para cada sonda de ADN individual se añade una molécula adaptadora de ADN con

15 diferentes códigos de barras. Para la secuenciación de nueva generación, las muestras se agrupan generalmente después de la ligación de los códigos de barras. Estos códigos de barras permiten la asignación de secuencias a las muestras después de la secuenciación.

En principio, los adaptadores de códigos de barras se pueden añadir en cualquier momento después de la etapa 1 y antes o durante la etapa 4.

20 Ventajosamente, las moléculas adaptadoras de ADN de acuerdo con la invención ya se pueden añadir al ADN fragmentado, antes de que se realice la reparación de los extremos y la fosforilación de los extremos 5'. Esto tiene las ventajas de que solo en la etapa 1 se añaden sondas individualizadas, que son ADN fragmentado y moléculas adaptadoras de ADN de acuerdo con la invención con el código de barras. Esto permite la automatización de las etapas subsiguientes, por ejemplo, con un robot de pipeteado, ya que en las etapas 2 a 4, se añaden

25 preferiblemente las mismas sustancias y volúmenes a todas las muestras.

En los protocolos conocidos en el estado de la técnica, pequeños volúmenes (2 µl o menos) de adaptadores de códigos de barras individualizados se tienen que añadir manualmente a la muestra, directamente antes de la ligación, después de la reparación de los extremos y la inactivación de las enzimas de reparación de los extremos. Este pipeteado de pequeños volúmenes propenso a errores se puede evitar de acuerdo con la invención.

30 Sin embargo, en la invención, los adaptadores de códigos de barras se pueden añadir a cada muestra después de la etapa 1 y las siguientes etapas las puede llevar a cabo completamente un robot de pipeteado. Al robot se pueden proporcionar mezclas universales de reparación de extremos y de ligación, y solo tiene que añadir volúmenes idénticos ajustados previamente, a cada muestra. Esto reduce al mínimo una inexactitud y una pérdida del volumen de reacción. Además, el robot no tiene que cambiar la punta de la pipeta después de cada etapa de pipeteado.

35 En particular, para la preparación de una genoteca de ADN, preferiblemente se liga una segunda molécula adaptadora de ADN al otro extremo del ADN genómico de hebra doble, fragmentado, con extremos reparados y fosforilado en 5'. Esta segunda molécula adaptadora de ADN contiene preferiblemente secuencias de cebador inverso y es preferiblemente una molécula adaptadora universal, que es la misma para cada sonda. La segunda molécula adaptadora de ADN es o bien una molécula adaptadora de ADN de acuerdo con la invención o una

40 molécula adaptadora de ADN común, tal y como se utiliza en el estado de la técnica, que también comprende una molécula de polinucleótidos de hebra doble. La segunda molécula adaptadora de ADN comprende preferiblemente nucleótidos con una estructura principal modificada (como tioatos de fósforo).

Preferiblemente, la segunda molécula adaptadora de ADN es una molécula adaptadora de ADN de acuerdo con la invención (con los extremos 5' de la primera hebra y la hebra inversa y el extremo 3' de la primera hebra modificados como se ha descrito anteriormente). Como molécula adaptadora universal, la segunda molécula adaptadora de ADN no tiene que contener un código de barras, pero contiene preferiblemente secuencias de cebador inverso. Ventajosamente también se puede añadir después de la etapa 1.

45 Si la segunda molécula adaptadora de ADN (o molécula adaptadora universal) es una molécula adaptadora de ADN de acuerdo con la invención, esto tiene la ventaja de que no es necesaria una inactivación de las enzimas, en particular de enzimas de reparación de los extremos, antes de realizar la ligación.

50 Si la segunda molécula adaptadora de ADN (o molécula adaptadora universal) es una molécula adaptadora de ADN común, conocida en el estado de la técnica, se tiene que realizar una inactivación de las enzimas de reparación de los extremos antes de la ligación (etapa 4), la cual se realiza preferiblemente mediante incubación a temperatura elevada, que depende de la enzima usada, preferiblemente una temperatura superior a 60°C. En este caso, la

55 segunda molécula adaptadora de ADN se añade después de la inactivación, de preferencia con la ligasa en la mezcla de ligación.

Como las etapas 1 a 3 se pueden realizar independientemente en otro laboratorio, la invención también comprende un método para la generación de una genoteca de ADN que comprende la etapa de ligación de moléculas

adaptadoras de ADN de acuerdo con la invención, a ADN de hebra doble, fragmentado, con los extremos reparados y fosforilado en 5'; con lo que preferiblemente no se realiza una inactivación de las enzimas, en particular de las enzimas de reparación de los extremos, antes de la ligación.

5 La ligación es una ligación de extremos romos, que se lleva a cabo preferiblemente en condiciones bien conocidas en la técnica (por ejemplo, 2 h a TA o 16°C durante la noche). La enzima preferida usada para la ligación es ADN ligasa de T4.

10 La etapa de ligación, tal y como se ha explicado anteriormente, viene acompañada preferiblemente por una reparación de muesca (etapa 5) mediante la adición de una enzima de reparación de muesca, preferiblemente una ADN polimerasa dependiente de ADN con actividad exonucleasa 5'→3' y actividad polimerasa 5'→3'. La enzima de reparación de muesca es preferiblemente termoestable, como por ejemplo, Taq polimerasa. Al calentar preferiblemente la mezcla de reacción, preferiblemente hasta temperaturas superiores a 60°C, la ligasa se inactiva mientras que la enzima de reparación de los extremos se activa. Las enzimas utilizadas para la etapa 4 y la etapa 5 se pueden añadir simultáneamente como una mezcla - "mezcla de enzimas de ligación y reparación de muesca", que contiene la enzima de reparación de muesca y la ligasa. El agente tamponador utilizado para la etapa 4 y la etapa 5 "como tampón de ligación" puede ser, ventajosamente, el mismo que en el "tampón de reparación de los extremos", tal como TRIS-HCl y preferiblemente con un pH entre 7 y 8. El "tampón de ligación" también contiene preferiblemente magnesio. Para la reparación de muesca, también se añaden desoxinucleósidos trifosfatos (preferiblemente dATP, dTTP, dGTP y dCTP) o bien al "tampón de ligación" o por separado.

20 El ADN utilizado para la generación de la genoteca de ADN es preferiblemente ADN genómico. Una de las características importantes de la genoteca de ADN genómico reside en que la genoteca de ADN genómico conserva sustancialmente el número de copias de un conjunto de genes o secuencias en el genoma original (el ADN genómico) y la relación de abundancia del conjunto de genes o secuencias en el ADN genómico.

Otro objeto de la invención es un kit para la generación de una genoteca de ADN, preferiblemente para llevar a cabo el método para generar una genoteca de ADN descrita anteriormente, que comprende:

- 25 i. moléculas adaptadoras de ADN de acuerdo con la invención,
- ii. preferiblemente para la reparación de los extremos: ADN polimerasa con actividad exonucleasa 3'→5' y polimerasa 5'→3' así como desoxinucleósidos trifosfatos (preferiblemente dATP, dTTP, dGTP y dCTP),
- iii. preferiblemente para la fosforilación en 5': una polinucleótido cinasa,
- iv. preferiblemente para la ligación: una ligasa,
- 30 v. preferiblemente una segunda molécula adaptadora de ADN (preferiblemente un adaptador universal) como se ha descrito anteriormente,
- vi. preferiblemente para la reparación de muesca: una polimerasa con actividad exonucleasa 5'→3', y
- vii. preferiblemente un tampón (como se ha descrito anteriormente que comprende preferiblemente magnesio y ATP).

35 Por tanto, el kit comprende al menos moléculas adaptadoras de ADN de acuerdo con la invención y lo más preferiblemente una ligasa.

Las diferentes enzimas se seleccionan preferiblemente como se ha descrito anteriormente.

40 Los componentes ii e iii se proporcionan preferiblemente en forma de "mezcla de enzimas para la reparación de extremos" como se ha definido anteriormente, preferiblemente junto con un "tampón para la reparación de extremos" como se ha definido anteriormente o incluso mezclados con ese tampón.

Los componentes iv y v se proporcionan preferiblemente en forma de "mezcla de enzimas para la ligación y la reparación de muesca" mezcla de enzimas para la reparación de extremos" preferiblemente junto con un "tampón de ligación" como se ha definido anteriormente o incluso mezclados con ese tampón.

45 Si no se proporcionan en los tampones, el kit también puede contener desoxinucleósidos trifosfatos (preferiblemente dATP, dTTP, dGTP y dCTP) en forma de "mezcla de dNTPs" como un componente separado.

50 También forma parte de la invención el uso de las moléculas adaptadoras de ADN descritas anteriormente o el kit descrito anteriormente, para la generación de una genoteca de ADN, en particular para el uso en el campo de la secuenciación de nueva generación y/o la multiplicación de genotecas. Las moléculas adaptadoras de ADN de acuerdo con la invención son particularmente adecuadas para el uso con soluciones automatizadas, ya que permiten protocolos muy simples y sólidos.

Las modificaciones de adaptadores de ADN de acuerdo con la invención permiten y facilitan una multiplicación de

las genotecas, especialmente en protocolos automatizados, ya que los adaptadores de códigos de barras ya se pueden añadir manualmente al ADN al comienzo de la preparación de la genoteca (al ADN fragmentado, antes de que se lleve a cabo la reparación de los extremos y la fosforilación de los extremos 5') y, posteriormente, la preparación de la genoteca se puede llevar a cabo en un robot de pipeteado. Una característica ventajosa adicional es que no es necesaria una inactivación térmica de las enzimas de reparación de los extremos: los fragmentos de la genoteca con los extremos reparados no se modifican adicionalmente con las enzimas que están presentes en la mezcla de reparación de los extremos y las moléculas adaptadoras de ADN de acuerdo la invención tampoco son un sustrato para esas enzimas. Una inactivación de esas enzimas antes de la ligación, previamente a la adición de la ADN ligasa, no es necesaria y la ligación se puede llevar a cabo en presencia de enzimas activas de reparación de los extremos. Esto permite unos protocolos para la preparación de genotecas automatizadas, acortados drásticamente sin tener que realizar un calentamiento y un enfriamiento de las muestras que requiere tiempo.

La invención se describe adicionalmente a través de las siguientes figuras y ejemplos, sin estar limitada a los mismos.

La Fig. 1 muestra un adaptador de código de barras conocido en el estado de la técnica, que está optimizado para una ligación con extremos romos, pero no se puede añadir durante o antes de la reparación de los extremos (y no sin la inactivación de las enzimas de reparación de los extremos). En la parte izquierda se muestra el sitio de la ligación con extremos romos. Para evitar una ligación en el sitio 3', la primera hebra (SEQ ID No. 9) tiene un extremo 3' sobresaliente que comprende los nucleósidos modificados A*C* con fosforotioato (para evitar una degradación del extremo sobresaliente). La hebra complementaria inversa (SEQ ID No. 10) es más corta. Los extremos 5' de ambas hebras no están fosforilados (tienen un 5'OH libre) para reducir la formación de dímeros del adaptador. La secuencia del código de barras está subrayada.

La Fig. 2 muestra un ejemplo de un adaptador de código de barras de acuerdo con la invención. De nuevo, en el lado izquierdo se muestra el sitio de ligación con extremos romos. Para evitar la ligación en el sitio 3', la primera hebra (SEQ ID No. 11) está modificada por M1. El extremo 5' de la primera hebra está modificado por M2 para evitar la fosforilación y para reducir la formación de dímeros y oligómeros de adaptadores. El extremo 5' de la hebra inversa (SEQ ID No. 12) está modificado por M3 para reducir la formación de dímeros y oligómeros de adaptadores. La secuencia del código de barras está subrayada. El extremo 3' de la hebra inversa tiene un 3'-OH libre.

La Fig. 3 muestra un adaptador de código de barras modificado de acuerdo con la invención, en donde la modificación en 5' de la primera hebra (SEQ ID No. 13) es 5'-OMeT. Para la generación de extremos romos, la hebra inversa (SEQ ID No. 14) tiene una A complementaria en el extremo 3' que tiene un grupo 3'-OH libre.

La Fig. 4 muestra la comparación de los valores Ct de genotecas de ADN generadas utilizando el protocolo convencional y el protocolo nuevo (o modificado) para la preparación de una genoteca con moléculas adaptadoras de ADN sin modificar y modificadas, en donde 1A es la referencia. Los identificadores 1A a 3E se corresponden con el de la tabla 1 (primera columna).

Las Figs. 5 a 11 muestran el análisis por tamaño del fragmento de genotecas de ADN usando adaptadores con códigos de barras modificados de acuerdo con la invención (Figs. 6 a 11) o adaptadores de códigos de barras sin modificar (Fig. 5), ya sea después de la preparación de una genoteca convencional (B) o después de la preparación de una genoteca con el protocolo modificado de acuerdo con la invención (A). Las genotecas de ADN se habilitan utilizando un chip de alta sensibilidad de Agilent en el intervalo de tamaño entre el marcador de menor tamaño y el de mayor tamaño (35-10380 pb). Los identificadores 1A a 3E se corresponden con los de la tabla 1. El primer pico (40 pb) se corresponde con moléculas adaptadoras no unidas. El segundo pico en la Fig. 5B y las Figs. 6 a 11, se corresponde con la genoteca de ADN. El pico con más de 10.000 pb se corresponde con el ADN genómico restante sin cortar. En la Fig. 5A el patrón de picos múltiples se corresponde con oligómeros de adaptadores, que no están presentes en las otras figuras.

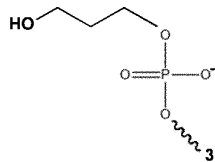
Ejemplo 1: Comparación de moléculas adaptadoras de ADN modificadas de forma diferente en un protocolo convencional y en un protocolo modificado para la preparación de una genoteca

Para someter a ensayo diferentes combinaciones de modificaciones de adaptadores de ADN, se sintetizaron las siguientes hebras sencillas y se aplicaron con diferentes combinaciones:

Tabla 1:

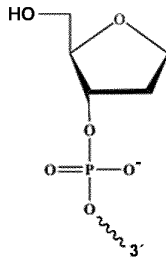
Identificador	Hebra inversa (5'→3')	
1 B2_Index1_unmod	TGTGACTTCAATTTACTATGTAGCAAAGGATACTCCGACGCGGCCGCA GCATCACGA	SEQ ID No.1
2 B2_Index1_C3	/5SpC3/TGTGACTTCAATTTACTATGTAGCAAAGGATACTCCGACGCGG CCGACGCATCACGA	SEQ ID No. 2
3 B2_Index1_spacer	/5dSp/TGTGACTTCAATTTACTATGTAGCAAAGGATACTCCGACGCGGC CGCAGCATCACGA	SEQ ID No. 3
	Primera hebra	
A B2_Com_Index1_unmod	TCGTGATGCTGCGGCCGCGTCGGAGTATCCTTTGCTACATAGTAAATT GAAGTCACA	SEQ ID No.4
B B2_Comp_Index1_C3_Amino	/5SpC3/TCGTGATGCTGCGGCCGCGTCGGAGTATCCTTTGCTACATAG TAAATTGAAGTCACA /3AmMO/	SEQ ID No. 5
C B2_Comp_Index1_C3_C3	/5SpC3/TCGTGATGCTGCGGCCGCGTCGGAGTATCCTTTGCTACATAG TAAATTGAAGTCACA /3SpC3/	SEQ ID No. 6
D B2_Comp_Index1_OMeT_Amino	5'- OMeT/TCGTGATGCTGCGGCCGCGTCGGAGTATCCTTTGCTACATAGT AAATTGAAGTCACA/3 AmMO/	SEQ ID No. 7
E B2_Comp_Index1_OMeT_C3	5'- OMeT/TCGTGATGCTGCGGCCGCGTCGGAGTATCCTTTGCTACATAGT AAATTGAAGTCACA/3 SpC3/	SEQ ID No. 8

El término "/5SpC3/" se refiere a un espaciador 5'-C3 como se ha descrito anteriormente, tal como se define en el presente documento, el espaciador 5'-C3 es un terminador 1-hidroxipropil-3-fosfatidilo con la siguiente fórmula:

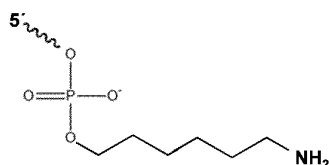


5

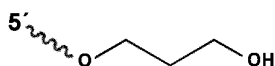
El término "/5dSp/" se refiere a 1,2-didesoxirribosil-3-fosfato como un espaciador 5'-D con la siguiente fórmula:



En la presente memoria, el término "/3AmMO/" se refiere a 6-aminohexil-1-fosfato como un modificador 3'-amino con la siguiente fórmula:

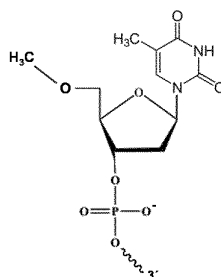


En la presente memoria, "/3SpC3/" es 1-oxi-3-propanol como un espaciador 3'-C3 con la siguiente fórmula:



5

En la presente memoria, "OMeT/" es el monofosfato de 5'-O-metil desoxitimidina como un modificador del extremo 5' con la siguiente estructura química:



10 Se emplean diferentes combinaciones de hebras adaptadoras modificadas, en paralelo con secuencias adaptadoras no modificadas (1A) en un protocolo convencional para la preparación de una genoteca de ADN (QIAGEN GeneRead® Library Prep (L) Handbook 03/2013). Para ello, 1 µg de ADN genómico (ADNg) procedente de *E. coli* para cada preparación, que se ha fragmentado antes hasta una longitud de 150 pares de bases, se aplica en una reacción de reparación de los extremos. Posteriormente, las enzimas que están siendo inactivadas térmicamente y las diferentes combinaciones de adaptadores modificados junto con el adaptador universal X sin modificar (primera hebra X_fwd: 5'-GTAAAACGACGGCCAGT-3', SEQ ID No. 15; hebra inversa: X_rev: 5'-ACTGGCCGTCGTTTAC*T-T-3', SEQ ID No. 16; * = fósforo tioato), se aplican en la reacción de ligación.

15

En el presente ejemplo se usan las siguientes combinaciones de adaptadores de acuerdo con la tabla 1: 1A (adaptador sin modificar - estado de la técnica), 2B, 2C, 2D, 2E, 3B, 3E.

20

De forma paralela al protocolo convencional, se lleva a cabo un protocolo modificado para la preparación de una genoteca modificada con las mismas combinaciones de moléculas adaptadoras. En contraposición al protocolo convencional, el adaptador del código de barras B2_Index_1 se pipetea junto con el ADNg en la reacción de reparación de los extremos. Después de terminar la reacción de reparación de los extremos solo se añade el adaptador universal X a la ligación.

25

En la siguiente tabla se ilustran los dos protocolos para la preparación de una genoteca ("protocolo nuevo" y "protocolo convencional"):

Tabla 2:

	Protocolo convencional (volumen en µl) estado de la técnica	Protocolo nuevo (volumen en µl) de acuerdo con la invención
1. Reparación de los extremos		
ADN fragmentado	18	18
H ₂ O	2,5	0,9
Adaptador de código de barras B2_Index_1	0,0	1,6

	Protocolo convencional (volumen en µl) estado de la técnica	Protocolo nuevo (volumen en µl) de acuerdo con la invención
Tampón de reparación de los extremos que contiene dNTPs, 10X	2,5	2,5
Mezcla de enzimas para la reparación de los extremos	2	2
Volumen total de reacción:	25	25
2. Ligación del adaptador		
ADN reparado en los extremos en mezcla de reacción	25	25
Tampón de ligación, 2X	40	40
Adaptador universal X	1,6	1,6
Adaptador de código de barras B2_Index_1	1,6	0,0
Mezcla de ligación y reparación de muesca	4	4
Mezcla de dNTPs (10 mM)	1	1
H ₂ O	6,8	8,4
Volumen total:	80	80

Después de la preparación de la reacción de reparación de los extremos, la mezcla se incuba durante 20 minutos a 25°C seguido por una etapa de inactivación enzimática durante 10 minutos a 70°C. A continuación, la reacción de ligación y de reparación de muesca se prepara añadiendo la mezcla de ligación y de reparación de muesca, el tampón de ligación y los dNTPs junto con el adaptador universal al ADN con extremos reparados. Esta mezcla se incuba durante 10 minutos a 25°C seguido de incubación durante 5 minutos a 72°C.

Las genotecas resultantes se purifican con la selección por tamaño GeneRead® (QIAGEN, Hilden, DE, punto de corte 100 pb). Después de la purificación, las genotecas resultantes se cuantifican a través de una PCR cuantitativa en tiempo real (qPCR) con la ayuda de cebadores localizados en los adaptadores y se analiza la distribución por tamaño de los fragmentos de las muestras usando electroforesis en gel capilar.

La comparación de los valores Ct de las genotecas de ADN generadas mediante el uso del protocolo convencional no muestra ninguna diferencia significativa entre la molécula adaptadora sin modificar 1A y las diferentes modificaciones de los adaptadores (excepto el adaptador 2C). Este hecho se ilustra en la Figura 3. Se muestra que las modificaciones de los adaptadores de acuerdo con la invención no influyen en la funcionalidad de la molécula adaptadora.

La comparación de los valores Ct de las genotecas de ADN generadas mediante el protocolo convencional, en comparación con aquellos en los que se empleaba el protocolo nuevo, está mostrando que los valores Ct para las moléculas adaptadoras modificadas después del protocolo nuevo son comparables o ligeramente mejores que después del protocolo convencional. La molécula de adaptador sin modificar proporciona valores Ct significativamente peores cuando se utiliza el protocolo nuevo, en comparación con el protocolo convencional, lo que muestra que el rendimiento en condiciones de uso de las moléculas adaptadoras modificadas de acuerdo con la invención, es independientemente del protocolo para la preparación de la genoteca aplicado, comparativamente alto.

La distribución por tamaño de fragmento de las genotecas de ADN resultantes se compara con Agilent Bioanalyzer High Sensitivity DNA Chips. Las distribuciones por tamaño de fragmento de las genotecas de ADN resultantes con adaptadores de ADN sin modificar (estado de la técnica) y los adaptadores de ADN modificados de acuerdo con la invención para el protocolo convencional, son comparables. Los tamaños de los fragmentos tienen por término medio aproximadamente 240 pb (150 pb del fragmento de la genoteca después de la fragmentación + 90 pb de las moléculas adaptadoras). Lo mismo es válido para las genotecas que son el resultado del uso de moléculas adaptadoras de ADN modificadas de acuerdo con la invención con el protocolo nuevo.

Únicamente en las genotecas que son el resultado del uso de moléculas adaptadoras sin modificar con el protocolo nuevo, son visibles picos claros en la región de tamaños de fragmento inferiores en el electroferograma. Esto muestra (Fig. 5A) que las moléculas adaptadoras sin modificar se dimerizan y oligomerizan cuando se aplican en el protocolo nuevo y, por lo tanto, no se pueden aplicar en el protocolo nuevo.

- 5 Los dímeros y oligómeros de adaptadores falsificarían la cuantificación de las genotecas durante una PCR cuantitativa y por otro lado reducirían la capacidad de lectura cuando se secuencian la genoteca, porque se amplifican y se secuencian como la genoteca de ADN, sin contener ninguna información útil.

Listado de secuencias

- 10 <110> QIAGEN GmbH
 <120> Moléculas adaptadoras de ADN para la preparación de genotecas de ADN y método para su producción y uso
- 15 <130> 00139P0010EPWO
 <150> EP 13 186 776.4
 <151> 2013-09-30
- 20 <160> 16
 <170> PatentIn versión 3.5
- 25 <210> 1
 <211> 57
 <212> ADN
 <213> artificial
- 30 <220>
 <223> Adaptador universal de ADN de cadena inversa
- <400> 1
 tgtgacttca atttactatg tagcaaagga tactccgacg cggccgcagc atcacga 57
- 35 <210> 2
 <211> 58
 <212> ADN
 <213> artificial
- 40 <220>
 <223> Adaptador universal de ADN de cadena inversa
- <220>
 <221> base_modificada
- 45 <222> (1)..(1)
 <223> n es 1-hidroxiopropil-3-fosfatidilo
- <400> 2
 ntgtgacttc aatttactat gttagcaaagg atactccgac gcggccgcag catcacga 58
- 50 <210> 3
 <211> 58
 <212> ADN
 <213> artificial
- 55 <220>
 <223> Adaptador universal de ADN de cadena inversa
- <220>
 <221> base_modificada
- 60 <222> (1)..(1)
 <223> n es 1,2-didesoxirribosil-3-fosfato
- <400> 3

ES 2 717 756 T3

ntgtgacttc aatttactat gttagcaagg atactccgac gcggccgcag catcacga 58

<210> 4
 <211> 57
 5 <212> ADN
 <213> artificial

<220>
 <223> Adaptador universal de ADN de cadena directa
 10

<400> 4
 tcgtgatgct gcggccgcgt cggagatcc ttgctacat agtaaattga agtcaca 57

<210> 5
 <211> 59
 15 <212> ADN
 <213> artificial

<220>
 <223> Adaptador universal de ADN de cadena directa
 20

<220>
 <221> base_modificada
 <222> (1)..(1)
 25 <223> n es 1-hidroxiopropil-3-fosfatidilo

<220>
 <221> base_modificada
 <222> (59)..(59)
 30 <223> n es 6-aminohexil-1-fosfato

<400> 5
 ntcgtgatgc tgcggccgcg tcggagtatc cttgctaca tagtaaattg aagtcacan 59

<210> 6
 <211> 59
 35 <212> ADN
 <213> artificial

<220>
 <223> Adaptador universal de ADN de cadena directa
 40

<220>
 <221> base_modificada
 <222> (1)..(1)
 45 <223> n es 1-hidroxiopropil-3-fosfatidilo

<220>
 <221> base_modificada
 <222> (59)..(59)
 50 <223> n es 1-oxi-3-propanol

<400> 6
 ntcgtgatgc tgcggccgcg tcggagtatc cttgctaca tagtaaattg aagtcacan 59

<210> 7
 <211> 59
 55 <212> ADN
 <213> artificial

<220>
 <223> Adaptador universal de ADN de cadena directa
 60

<220>
 <221> base_modificada
 <222> (1)..(1)
 65

ES 2 717 756 T3

<223> n es 5'-O-metil desoxitimidina monofosfato
 <220>
 <221> base_modificada
 5 <222> (59)..(59)
 <223> n es 6-aminohehexil-1-fosfato
 <400> 7
 10 ntcgtgatgc tgcggccgcg tcggagtatc ctttgctaca tagtaaattg aagtcacan 59
 <210> 8
 <211> 59
 <212> ADN
 <213> artificial
 15 <220>
 <223> Adaptador universal de ADN de cadena directa
 <220>
 20 <221> base_modificada
 <222> (1)..(1)
 <223> n es 5'-O-metil desoxitimidina monofosfato
 <220>
 25 <221> base_modificada
 <222> (59)..(59)
 <223> n es 1-oxi-3-propanol
 <400> 8
 30 ntcgtgatgc tgcggccgcg tcggagtatc ctttgctaca tagtaaattg aagtcacan 59
 <210> 9
 <211> 56
 <212> ADN
 35 <213> artificial
 <220>
 <223> Adaptador universal de ADN de cadena directa
 <220>
 40 <221> base_modificada
 <222> (54)..(54)
 <223> n es el nucleósido modificado A con fosforotioato
 <220>
 45 <221> base_modificada
 <222> (55)..(55)
 <223> n es el nucleósido modificado C con fosforotioato
 <400> 9
 50 cgtgatgctg cggccgctgc ggagtatcct ttgctacata gtaaattgaa gtcnna 56
 <210> 10
 <211> 52
 55 <212> ADN
 <213> artificial
 <220>
 <223> Adaptador universal de ADN de cadena inversa
 60 <400> 10
 actcaattt actatgtagc aaaggatact cgcagcggc cgcagcatca cg 52
 <210> 11
 65 <211> 58
 <212> ADN

ES 2 717 756 T3

<213> artificial
 <220>
 <223> Adaptador universal de ADN de cadena directa
 5
 <220>
 <221> base_modificada
 <222> (1)..(1)
 <223> n es desoxinucleótido eterificado o 4'-amino-desoxinucleótido o 5'-amino-desoxinucleótido o 5'-C3-Espaciador
 10 o 5'-D-Espaciador
 <220>
 <221> base_modificada
 <222> (58)..(58)
 15 <223> n es 3'-Amino-Modificador o 3'-C3-Espaciador
 <400> 11
 ncgatgatgct gcggccgcgt cggagtatcc ttgctacat agtaaattga agtcacan 58
 20 <210> 12
 <211> 57
 <212> ADN
 <213> artificial
 25 <220>
 <223> Adaptador universal de ADN de cadena inversa
 <220>
 <221> base_modificada
 30 <222> (1)..(1)
 <223> n es desoxinucleótido eterificado o 4'-amino-desoxinucleótido o 5'-amino-desoxinucleótido o 5'-C3-Espaciador
 o 5'-D-Espaciador
 <400> 12
 35 ntgtgacttc aatttactat gtgcaaaagg atactccgac gcggccgcag catcacg 57
 <210> 13
 <211> 58
 <212> ADN
 40 <213> artificial
 <220>
 <223> Adaptador universal de ADN de cadena directa
 45 <220>
 <221> base_modificada
 <222> (1)..(1)
 <223> n es 5'-O-metil desoxitimidina monofosfato
 50 <220>
 <221> base_modificada
 <222> (58)..(58)
 <223> n es desoxinucleótido eterificado o 4'-amino-desoxinucleótido o 5'-amino-desoxinucleótido o 5'-C3-Espaciador
 o 5'-D-Espaciador
 55 <400> 13
 ncgatgatgct gcggccgcgt cggagtatcc ttgctacat agtaaattga agtcacan 58
 <210> 14
 60 <211> 58
 <212> ADN
 <213> artificial
 <220>
 65 <223> Adaptador universal de ADN de cadena inversa

ES 2 717 756 T3

<220>
<221> base_modificada
<222> (1)..(1)
<223> n es desoxinucleótido eterificado o 4'-amino-desoxinucleótido o 5'-amino-desoxinucleótido o 5'-C3-Espaciador
o 5'-D-Espaciador

5

<400> 14
ntgtgacttc aatttactat gtagcaaagg atactccgac gcggccgcag catcacga 58

10 <210> 15
<211> 17
<212> ADN
<213> artificial

15 <220>
<223> Adaptador universal de ADN de cadena directa

<400> 15
gtaaaacgac ggccagt 17

20 <210> 16
<211> 19
<212> ADN
<213> artificial

25 <220>
<223> Adaptador universal de ADN de cadena inversa

30 <220>
<221> n
<222> (17)..(17)
<223> n es 2'-Desoxicitosina-5'-O-monofosfotioato

35 <220>
<221> n
<222> (18)..(18)
<223> n es 2'-Desoxicitosina-5'-O-monofosfotioato

40 <400> 16
actggccgtc gttttannt 19

REIVINDICACIONES

1. Molécula adaptadora de ADN, que comprende una molécula de hebra doble con extremos romos, en la que
 - a. el extremo 5' de la primera hebra se modifica de manera que el primer nucleótido no contiene un grupo hidroxilo libre ni un fosfato libre en la posición 5',
 - 5 b. el extremo 3' de la primera hebra se modifica de manera que el último nucleótido no contiene un grupo hidroxilo libre en la posición 3',
 - c. el extremo 5' de la hebra inversa se modifica de manera que el primer nucleótido no contiene un grupo hidroxilo libre ni un fosfato libre en la posición 5',
 - d. el extremo 3' de la hebra inversa presenta un grupo hidroxilo libre.
- 10 2. Molécula adaptadora de ADN según la reivindicación 1, que contiene una secuencia de código de barras, así como sitios de unión para cebadores de la amplificación y la secuenciación.
3. Molécula adaptadora de ADN según la reivindicación 1 o 2, caracterizada por que tiene una longitud de 20 a 90, preferiblemente de 40 a 70, más preferiblemente de 40 a 60 pares de bases.
- 15 4. Molécula adaptadora de ADN según una de las reivindicaciones 1 a 3, caracterizada por que los grupos hidroxilo en los extremos 5' de ambas hebras y en el extremo 3' de la primera hebra están esterificados o eterificados.
5. Molécula adaptadora de ADN según una de las reivindicaciones 1 a 4, caracterizada por que el extremo 5' de la hebra inversa se modifica mediante la fijación de un espaciador 5'-C3 o un espaciador 5'-D.
6. Molécula adaptadora de ADN según una de las reivindicaciones 1 a 5, caracterizada por que el extremo 5' de la primera hebra se modifica mediante la fijación de un espaciador 5'-C3 o un 5'-O-metil desoxinucleótido.
- 20 7. Molécula adaptadora de ADN según una de las reivindicaciones 1 a 6, caracterizada por que el extremo 3' de la primera hebra se modifica mediante un espaciador 3'-C3 o un modificador 3'-amino.
8. Método para la producción de moléculas adaptadoras de ADN, que comprende una molécula de hebra doble con extremos romos, con las etapas siguientes:
 - a. modificación del extremo 5' de la primera hebra de manera que no está disponible ningún grupo hidroxilo libre y ningún fosfato libre,
 - 25 b. modificación del extremo 3' de la primera hebra de manera que no está disponible ningún grupo hidroxilo libre,
 - c. modificación del extremo 5' de la hebra inversa de manera que no está disponible ningún grupo hidroxilo libre y ningún fosfato libre.
- 30 9. Método para la producción de moléculas adaptadoras de ADN según la reivindicación 8, en donde la modificación del extremo 5' de la primera hebra se consigue mediante la fijación de monofosfato de 5'-O-metil desoxitimidina o un espaciador 5'-C3.
10. Método para la producción de moléculas adaptadoras de ADN según la reivindicación 8 o 9, en el que la modificación del extremo 3' de la primera hebra se consigue mediante la fijación de un espaciador 3'-C3 o modificadores 3'-amino.
- 35 11. Método para la producción de moléculas adaptadoras de ADN según una de las reivindicaciones 8 a 10, en el que la modificación del extremo 5' de la hebra inversa se consigue mediante la fijación de un espaciador 5'-C3 o un espaciador 5'-D.
12. Método para la generación de una genoteca de ADN, que comprende las etapas
 1. fragmentación de ADN de hebra doble, preferiblemente por métodos físicos o químicos o enzimáticos o una combinación de los mismos,
 - 40 2. reparación de los extremos del ADN de hebra doble fragmentado, de preferencia con la ayuda de una ADN polimerasa con actividad exonucleasa 3'→5' y polimerasa 5'→3', tal como la ADN polimerasa de T4,
 3. fosforilación de los extremos 5' del ADN de hebra doble fragmentado, de preferencia con una polinucleótido cinasa, tal como la polinucleótido cinasa de T4,
 - 45 4. ligación de las moléculas adaptadoras de ADN con extremos romos según una cualquiera de las reivindicaciones 1-7 con el ADN de hebra doble, fragmentado, con extremos reparados y fosforilado en 5'; con lo que las moléculas adaptadoras de ADN según una cualquiera de las reivindicaciones 1-7 se añaden antes de la

etapa 2.

- 5 13. Método para la generación de una genoteca de ADN, que comprende la etapa de ligación de moléculas adaptadoras de ADN con extremos romos según una de las reivindicaciones 1 a 7 con ADN de hebra doble, fragmentado, con extremos reparados y fosforilado en 5', con lo que no se realiza una inactivación de las enzimas antes de la ligación.
14. Kit que comprende
- moléculas adaptadoras de ADN con extremos romos según una de las reivindicaciones 1 a 7,
 - preferiblemente una ADN polimerasa con actividad exonucleasa 3'→5' y polimerasa 5'→3', así como desoxinucleósidos trifosfatos,
- 10
- preferiblemente una polinucleótido cinasa,
 - preferiblemente una ligasa,
 - una polimerasa con actividad exonucleasa 5'→3', y
 - preferiblemente un tampón.
15. Uso de moléculas adaptadoras de ADN con extremos romos según una de las reivindicaciones 1 a 7 o de un kit según la reivindicación 14, para la generación de una genoteca de ADN, en particular en el campo de la secuenciación de nueva generación y/o la multiplicación de genotecas.

5' - CGTGATGCTGCGGCCGCTCGGAGTATCCTTTGCTACATAGTAAATTGAAGTCA*C*A-3'
 3' - GCACTACGACGCCGGCGCAGCCTCATAGGAAACGATGTATCATTAACTTCA-5'

Estado de la técnica

Fig. 1

5' - **M2**CGTGATGCTGCGGCCGCTCGGAGTATCCTTTGCTACATAGTAAATTGAAGTCACAM1-3'
 3' - GCACTACGACGCCGGCGCAGCCTCATAGGAAACGATGTATCATTAACTTCAGTGT**M3**-5'

Fig. 2

5' - **OmeT**CGTGATGCTGCGGCCGCTCGGAGTATCCTTTGCTACATAGTAAATTGAAGTCACAM1-3'
 3' - **AG**CACTACGACGCCGGCGCAGCCTCATAGGAAACGATGTATCATTAACTTCAGTGT**M3**-5'

Fig. 3

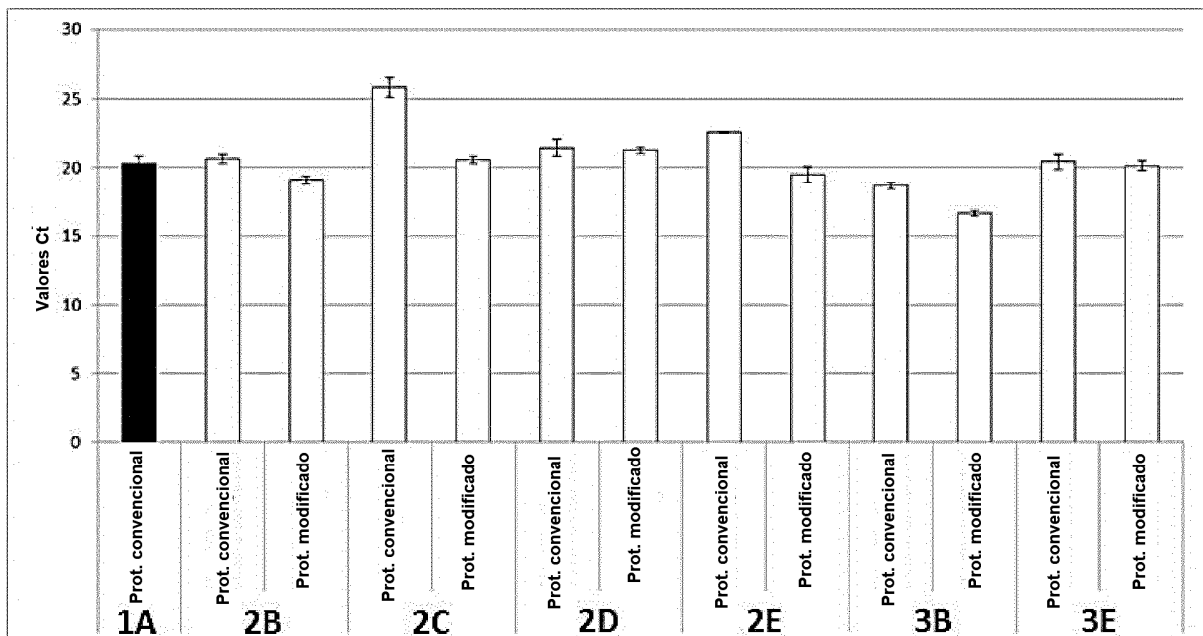


Fig. 4

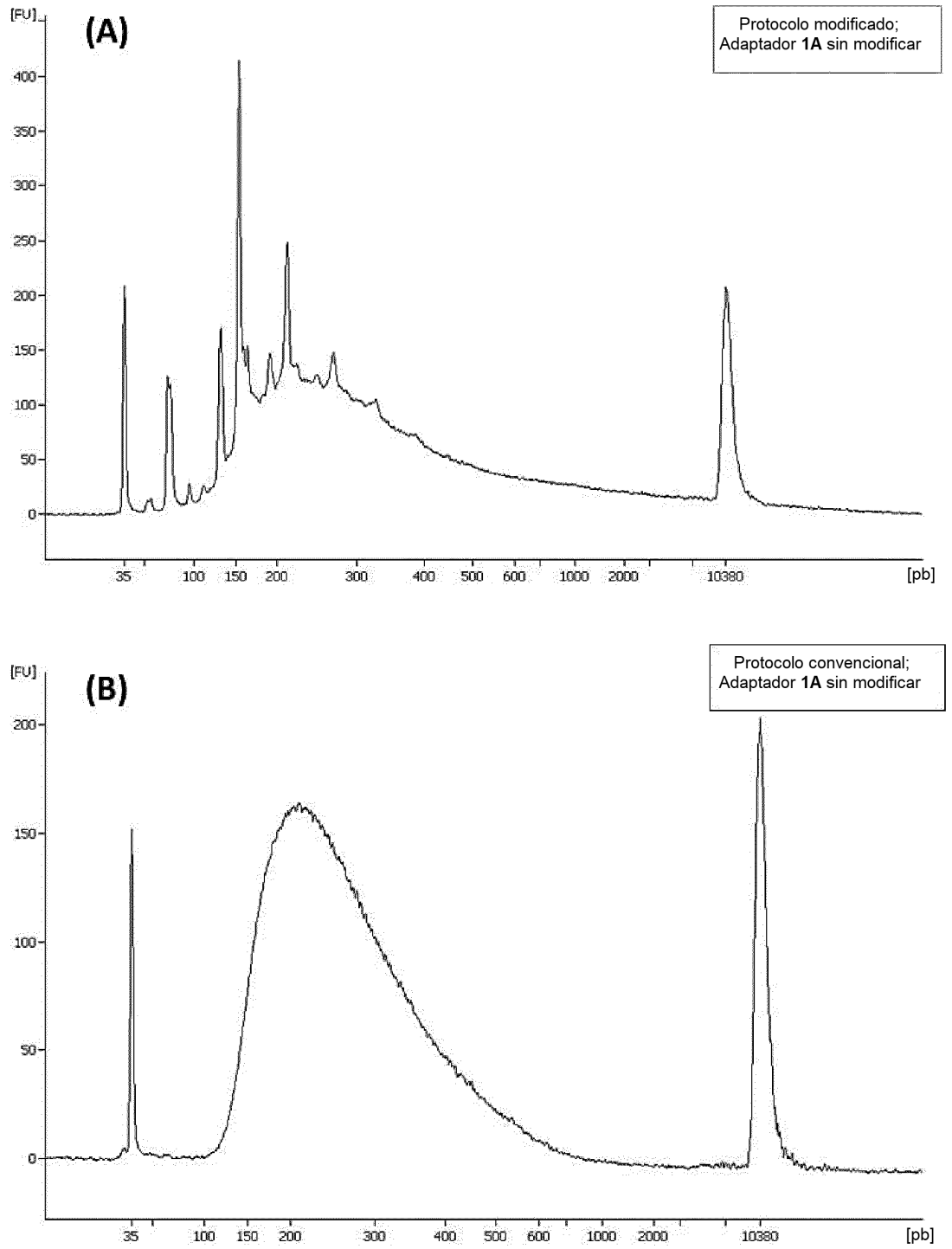


Fig. 5

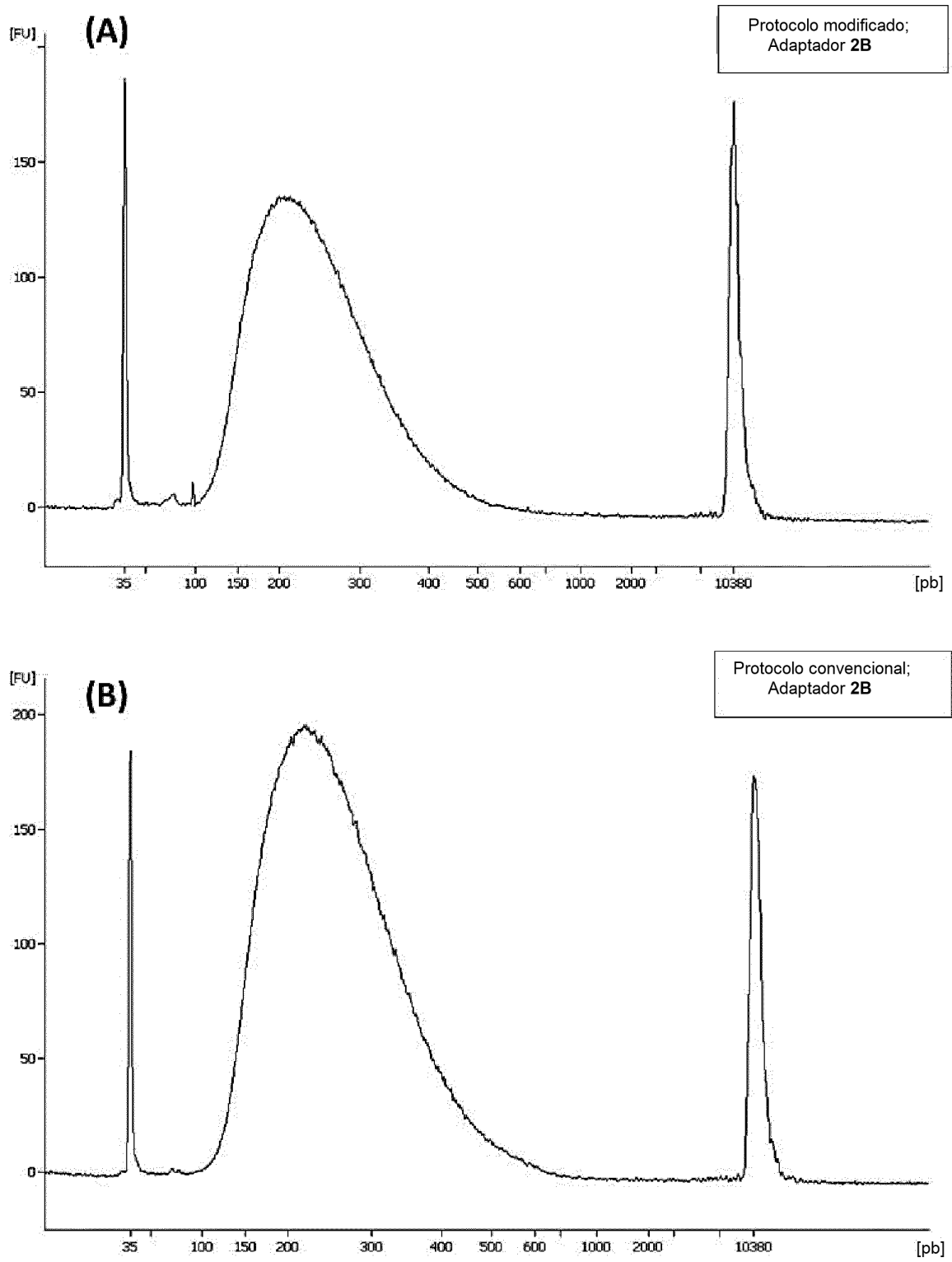


Fig. 6

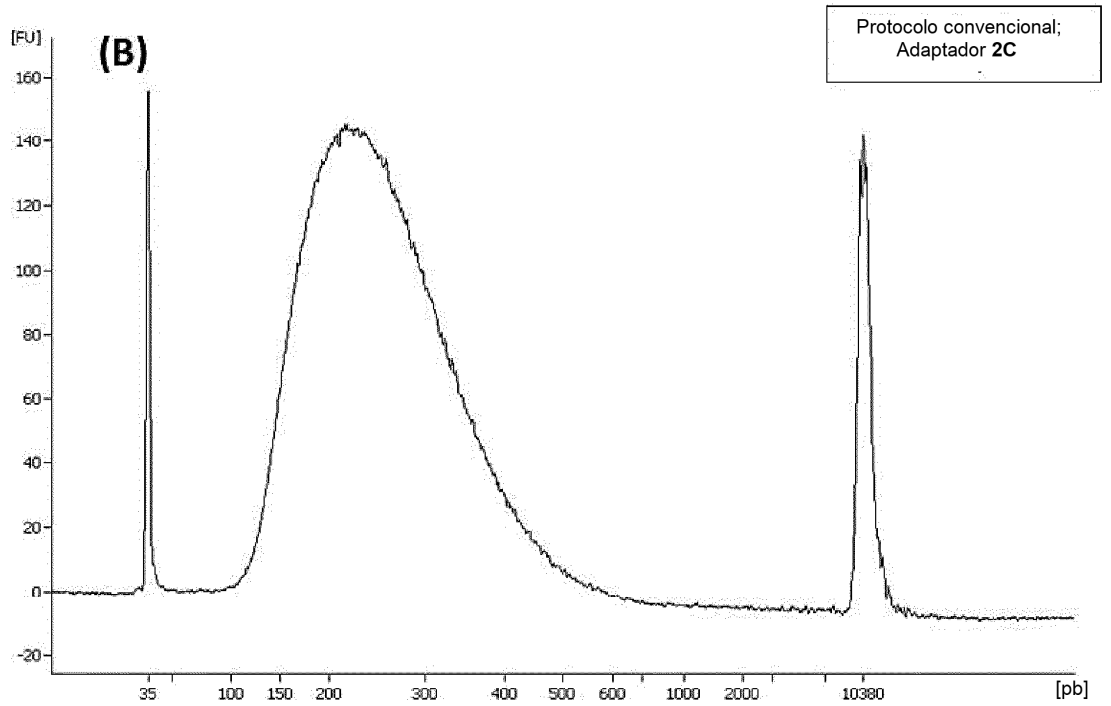
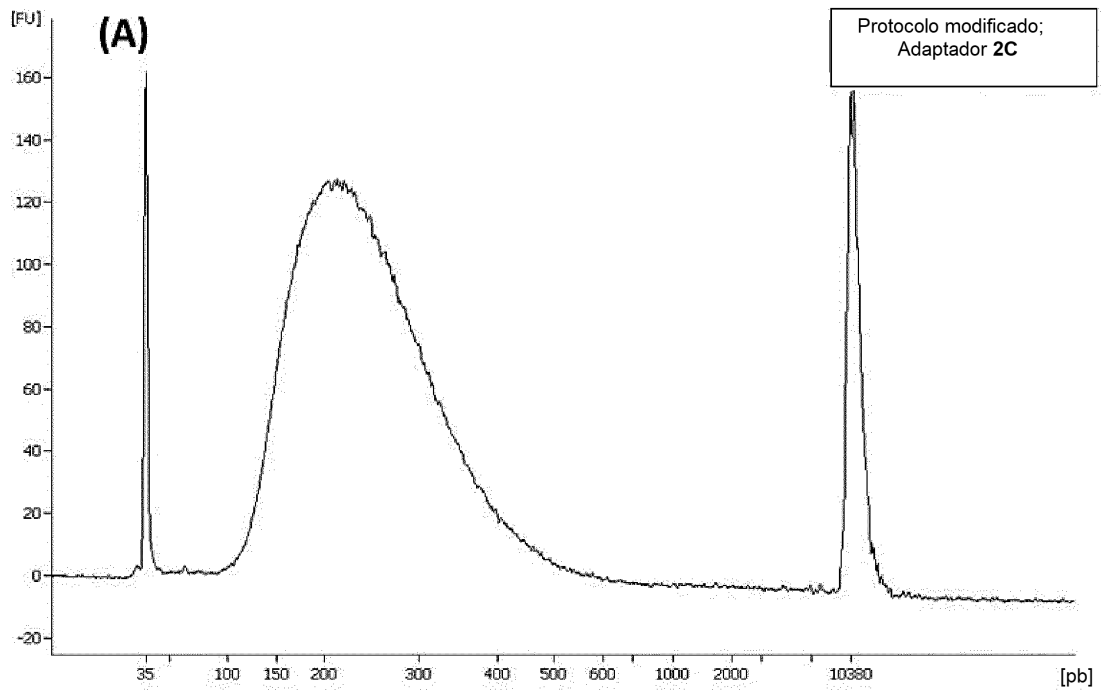


Fig. 7

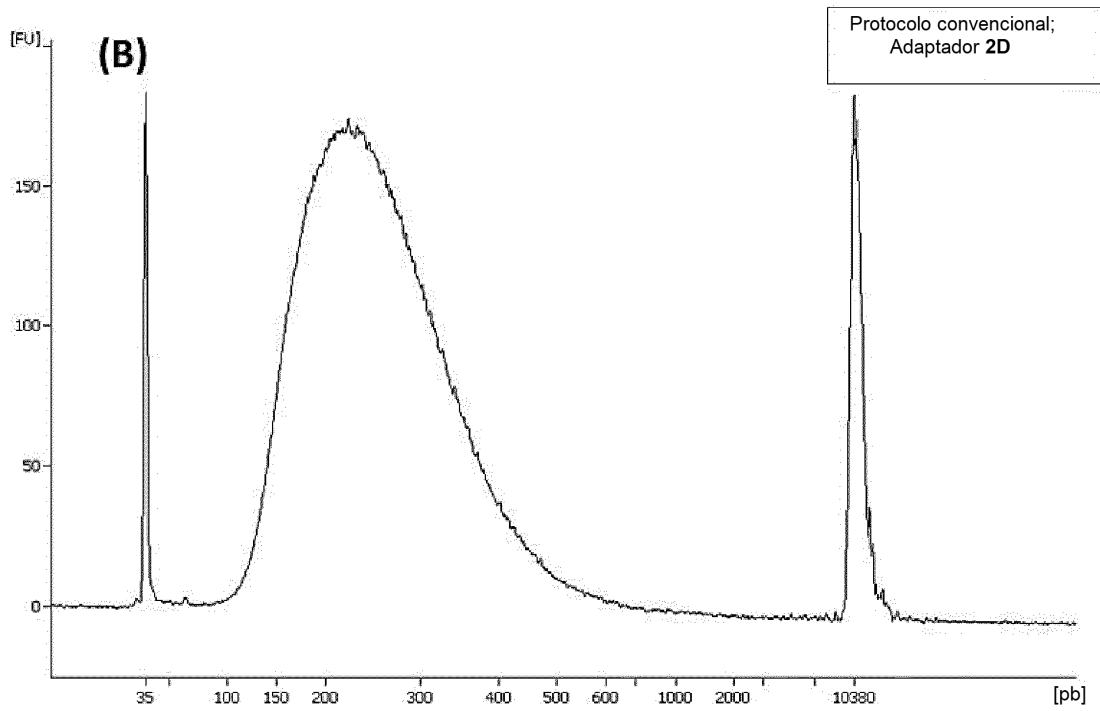
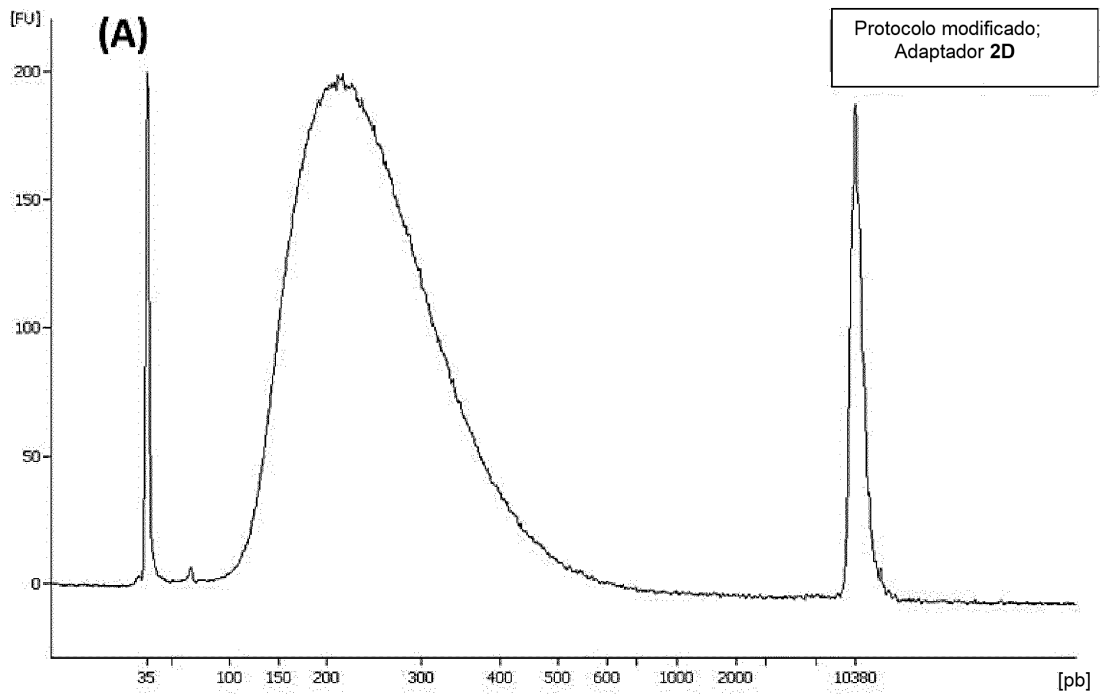


Fig. 8

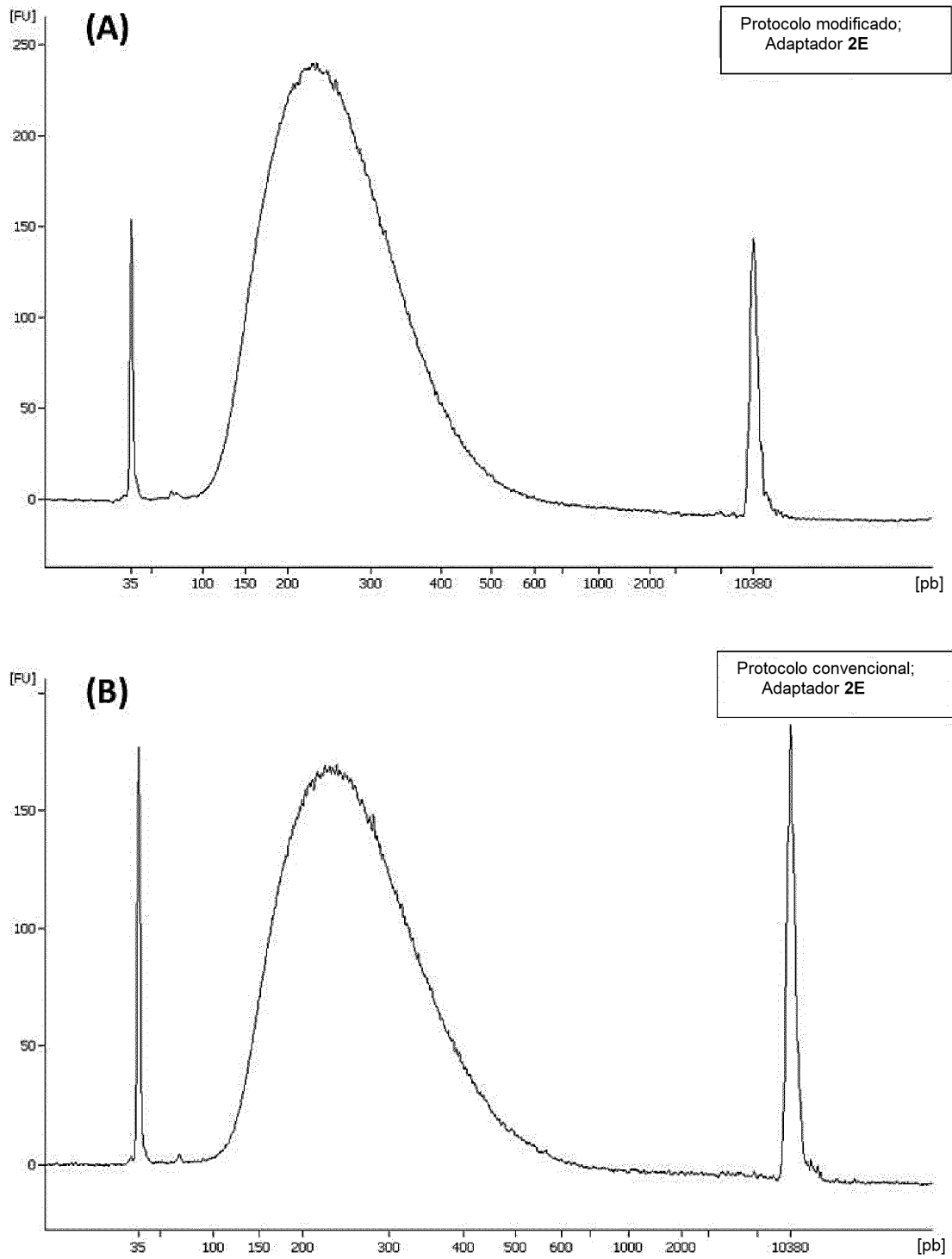


Fig. 9

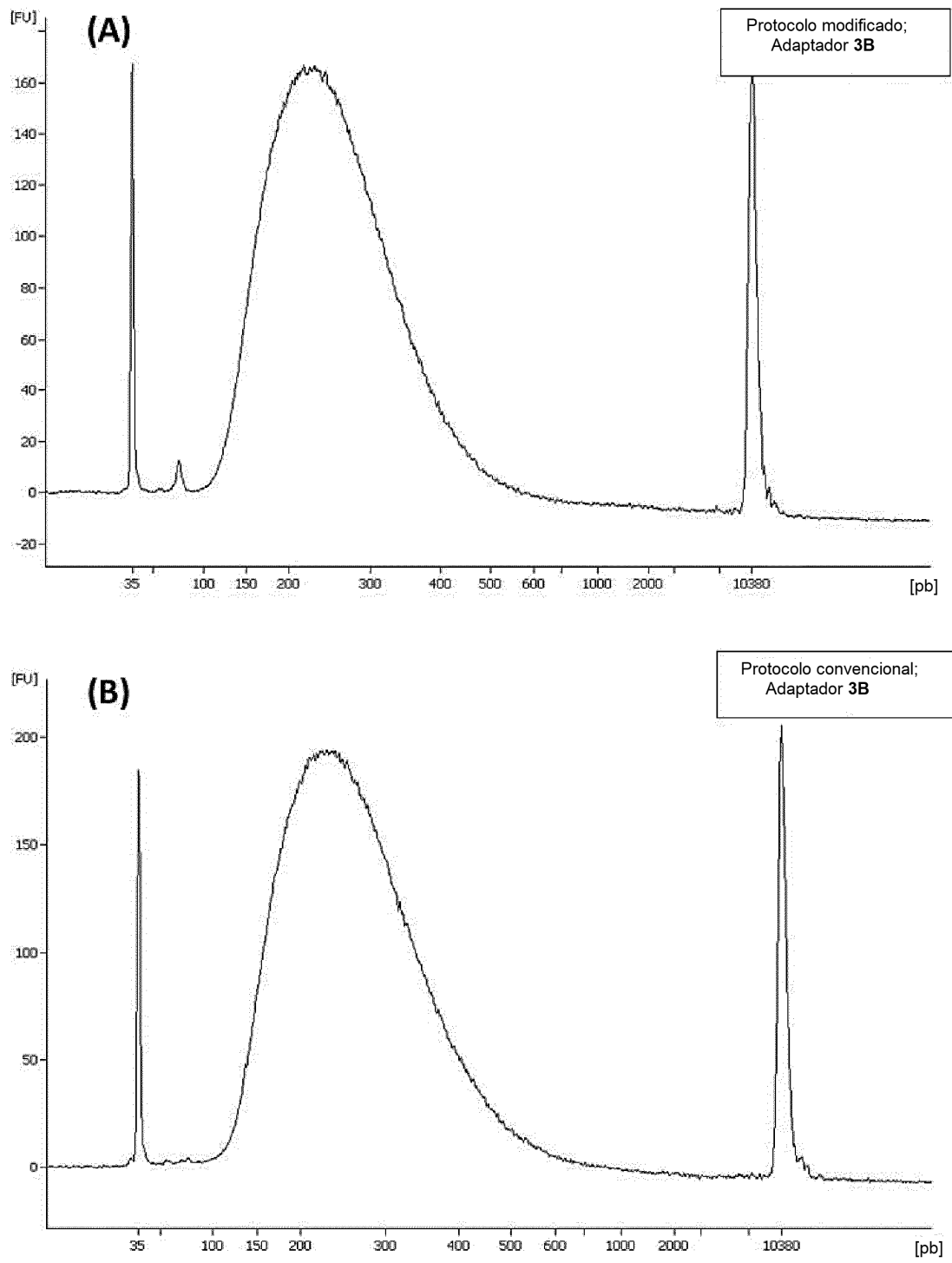


Fig. 10

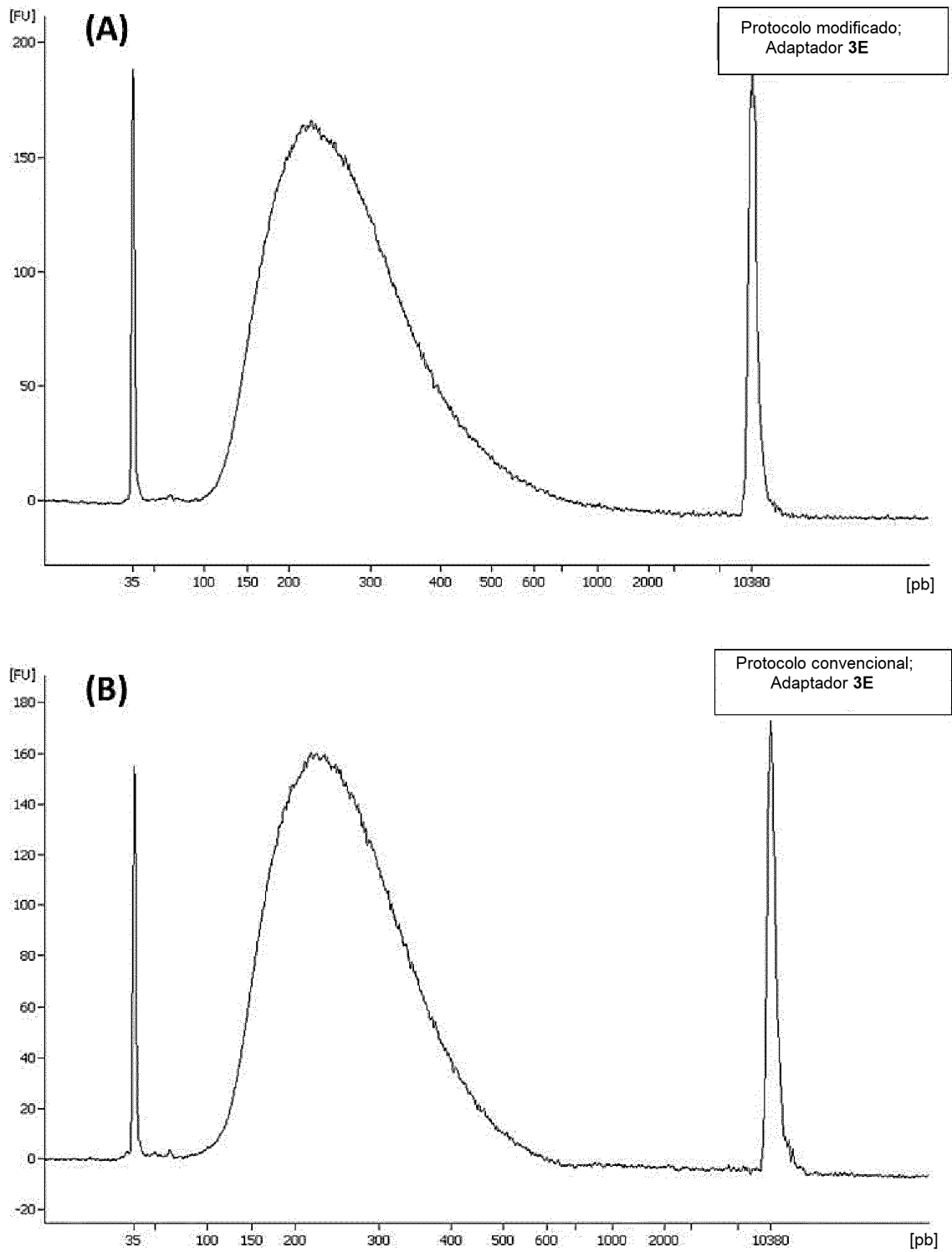


Fig. 11