

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 732 253**

51 Int. Cl.:

C12Q 1/68 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **31.10.2006 E 11004071 (4)**

97 Fecha y número de publicación de la concesión europea: **03.04.2019 EP 2423325**

54 Título: **Método de preparación de bibliotecas de polinucleótidos plantilla**

30 Prioridad:

01.11.2005 GB 0522310
14.06.2006 US 486953

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
21.11.2019

73 Titular/es:

ILLUMINA CAMBRIDGE LIMITED (100.0%)
Illumina Centre, 19 Granta Park, Great Abington
Cambridge, Cambridgeshire CB21 6DF, GB

72 Inventor/es:

GORMLEY, NIALL, ANTHONY;
SMITH, GEOFFREY, PAUL;
BENTLEY, DAVID y
RIGATTI, ROBERTO

74 Agente/Representante:

CONTRERAS PÉREZ, Yahel

ES 2 732 253 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

Método de preparación de bibliotecas de polinucleótidos plantilla

5 **Campo de la invención**

La invención se refiere a un método de preparación de una biblioteca de polinucleótidos plantilla y también al uso de la biblioteca de plantillas en métodos de amplificación de ácidos nucleicos en fase sólida. En particular, la invención se refiere a un método de preparación de una biblioteca de polinucleótidos plantilla que tienen secuencias comunes en sus extremos 5' y en sus extremos 3'.

Antecedentes de la invención

El desarrollo de fármacos de biología molecular y farmacológicos actual hace un uso intensivo del análisis de ácidos nucleicos. Las áreas más conflictivas son las de secuenciación del genoma completo, detección de polimorfismo de un solo nucleótido, identificación sistemática y control de la expresión de genes, que normalmente requieren análisis de grandes cantidades de ácido nucleico.

Un área de la tecnología que revolucionó el estudio de los ácidos nucleicos fue el desarrollo de técnicas de amplificación de ácido nucleico, tales como la reacción en cadena de la polimerasa (PCR). Las reacciones de amplificación, tales como PCR, pueden permitir al usuario amplificar específica y selectivamente un ácido nucleico diana particular de interés a partir de una mezcla compleja de ácidos nucleicos. Sin embargo, también existe una necesidad continua de técnicas de amplificación de ácido nucleico que permitan la amplificación simultánea de mezclas complejas de plantillas de diversa secuencia, tales como fragmentos de ADN genómico (p. ej., amplificación del "genoma completo") o bibliotecas de ADNc, en una única reacción de amplificación.

La amplificación por PCR no puede producirse en ausencia de hibridación de cebadores de amplificación directos e inversos para secuencias de unión a cebador en el plantilla a amplificar en las condiciones de las etapas de hibridación de la reacción de PCR, es decir, si no hay una complementariedad suficiente entre cebadores y el plantilla. Por lo tanto, se requiere cierto conocimiento previo de la secuencia del plantilla antes de poder llevar a cabo una reacción de PCR para amplificar un plantilla *específico*, a menos que se usen cebadores aleatorios con una pérdida consecuente de especificidad. El usuario generalmente ha de conocer de antemano la secuencia de al menos los sitios de unión a cebador en el plantilla para que puedan diseñarse los cebadores apropiados, aunque la secuencia restante del plantilla puede ser desconocida. La necesidad de conocimiento previo de la secuencia del plantilla aumenta la complejidad y el coste de la amplificación por PCR de mezclas complejas de plantillas, tales como fragmentos de ADN genómico.

Los documentos WO 98/44151 y WO 00/18957 describen métodos de formación de series de polinucleótidos basadas en amplificación de ácidos nucleicos en "fase sólida", que es análoga a una reacción en cadena de la polimerasa en la que los productos de amplificación están inmovilizados sobre un soporte sólido para formar series comprendidas por agregados de ácido nucleico o "colonias". Cada agregado o colonia de dicha serie se forma a partir de una pluralidad de cadenas de polinucleótidos idénticas inmovilizadas y una pluralidad de cadenas de polinucleótidos complementarias idénticas inmovilizadas. Las series así formadas se denominan generalmente en la presente memoria "series agregadas" y sus características generales se entenderán por referencia a los documentos WO 98/44151 o WO 00/18957.

Como se ha indicado anteriormente, los métodos de amplificación en fase sólida de los documentos WO 98/44151 y WO 00/18957 son esencialmente una forma de la reacción en cadena de la polimerasa realizada sobre un soporte sólido. Como cualquier reacción de PCR, estos métodos requieren el uso de cebadores de amplificación directos e inversos (que pueden ser idénticos o diferentes) capaces de hibridar con un plantilla a amplificar. En los métodos de los documentos WO 98/44151 y WO 00/18957 ambos cebadores se inmovilizan sobre el soporte sólido en el extremo 5'. Se conocen otras formas de amplificación en fase sólida en las que solo un cebador está inmovilizado y el otro está presente en solución libre (Mitra, R. D y Church, G. M., Nucleic Acids Research, 1999, Vol.27, N.º 24).

En común con todas las técnicas de PCR, la amplificación por PCR en fase sólida requiere el uso de cebadores de amplificación directos e inversos que incluyen secuencias de nucleótidos "específicas de plantilla" que son capaces de hibridar con secuencias en el plantilla a amplificar, o su complemento, en condiciones de las etapas de hibridación de la reacción de PCR. Las secuencias en el plantilla con el que los cebadores hibridan en condiciones de la reacción de PCR pueden referirse en la presente memoria como secuencias "de unión a cebador".

Determinadas realizaciones de los métodos descritos en los documentos WO 98/44151 y WO 00/18957 utilizan cebadores "universales" para amplificar plantillas que comprenden una parte plantilla variable que se desea amplificar flanqueada en 5' y 3' por secuencias de unión a cebador común o "universal". Los cebadores directos e inversos "universales" incluyen secuencias capaces de hibridar con las secuencias de unión a cebador "universal" en la construcción plantilla. La parte plantilla variable puede en sí misma ser de secuencia conocida, desconocida o

- parcialmente desconocida. Este enfoque tiene la ventaja de que no es necesario diseñar un par específico de cebadores para cada plantilla a amplificar; pueden utilizarse los mismos cebadores para la amplificación de diferentes plantillas siempre que cada plantilla se modifique por la adición de las mismas secuencias de unión a un cebador universal en sus extremos 5' y 3'. La secuencia plantilla variable puede por lo tanto ser cualquier fragmento de ADN de interés. Un enfoque análogo puede usarse para amplificar una mezcla de plantillas, tal como una pluralidad o biblioteca de moléculas de ácido nucleico plantilla (p. ej., fragmentos de ADN genómico), usando un solo par de cebadores directos e inversos universales siempre que cada molécula plantilla en la mezcla se modifique por la adición de las mismas secuencias de unión de cebador universal.
- 10 Dichos enfoques de "cebador universal" para amplificación por PCR, y en particular amplificación por PCR en fase sólida, son ventajosos ya que permiten múltiples moléculas plantilla de secuencia idéntica o diferente, conocida o desconocida, a amplificar en una única reacción de amplificación, que puede llevarse a cabo en un soporte sólido que lleva un solo par de cebadores "universales". La amplificación simultánea de una mezcla de plantillas de diferentes secuencias por PCR requeriría de otra manera una pluralidad de pares de cebadores, siendo cada par complementario a cada plantilla único en la mezcla. La generación de una pluralidad de pares de cebadores para cada plantilla individual no es una opción viable para mezclas complejas de plantillas.

La adición de secuencias de cebador universal sobre los extremos de las plantillas a amplificar por PCR puede conseguirse mediante diversos métodos conocidos por los expertos en la materia. Por ejemplo, un cebador universal que consiste en una secuencia universal en su extremo 5' y una secuencia degenerada en su extremo 3' pueden usarse en una PCR (DOP-PCR, p. ej. PNAS 1996 vol. 93 págs. 14676-14679) para amplificar fragmentos al azar a partir de un plantilla complejo o una mezcla compleja de plantillas. La parte 3' degenerada del cebador hibrida en posiciones al azar sobre ADN y puede extenderse para generar una copia del plantilla que tiene la secuencia universal en su extremo 5'.

Alternativamente, adaptadores que contengan secuencias de cebador universal pueden ligarse en los extremos de las plantillas. Los adaptadores pueden ser mono o bicatenarios. Sin ser bicatenarios, pueden tener extremos solapantes que son complementarios a los extremos salientes de las moléculas plantilla que se han generado con una endonucleasa de restricción. Alternativamente, los adaptadores bicatenarios pueden ser romos, en cuyo caso los plantillas también tienen extremos romos. Los extremos romos de las plantillas pueden haberse formado durante un proceso para cizallar el ADN en fragmentos, o pueden haberse formado mediante una reacción de reparación de extremos, como es conocido por los expertos en la materia.

Pueden usarse un solo adaptador o dos adaptadores diferentes en una reacción de ligamiento con plantillas. Si un plantilla se ha manipulado de tal manera que sus extremos son iguales, es decir, ambos son romos o ambos tienen el mismo saliente, entonces el ligamiento de un solo adaptador compatible generará un plantilla con ese adaptador en ambos extremos. Sin embargo, si se usan dos adaptadores compatibles, adaptador A y adaptador B, entonces se forman tres permutaciones de productos ligados: plantilla con adaptador A en ambos extremos, plantilla con adaptador B en ambos extremos, y plantilla con adaptador A en un extremo y adaptador B en el otro extremo. Este último producto es, en algunas circunstancias, el único producto deseado de la reacción de ligamiento y por consiguiente son necesarias etapas de purificación adicionales después de la reacción de ligamiento para purificarlo de los productos de ligamiento que tengan el mismo adaptador en ambos extremos.

La actual invención presentada en la presente memoria es un método que usa un solo adaptador en una reacción de ligamiento para generar una biblioteca de polinucleótidos plantilla, cada uno de los cuales tiene secuencias de cebador universal comunes, pero diferentes, en sus extremos 5' y 3'. El método puede aplicarse para preparar mezclas simples o complejas de plantillas para amplificación, por ejemplo, una superficie sólida, usando secuencias de cebador, sin conocimiento previo de secuencias plantilla. La invención puede aplicarse para la preparación de plantillas a partir de muestras complejas tales como genomas completos o mezclas de ADNc, así como aplicaciones monoplantilla.

Sumario de la invención

La invención proporciona un método, como se define en las reivindicaciones, de generación de una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3', comprendiendo el método las etapas que consisten en:

- proporcionar uno o más dúplex de polinucleótido diana con extremos romos,
- añadir un solo desoxinucleótido "A" de los dúplex de polinucleótido diana, produciendo así un saliente de una base en 3',
- ligar polinucleótidos adaptadores con apareamiento erróneo idénticos en los dos extremos de cada uno de uno o más dúplex de polinucleótido diana para formar una o más construcciones adaptador-diana, en el que cada adaptador con apareamiento erróneo se forma a partir de dos cadenas polinucleotídicas hibridadas que forman un complejo biomolecular que comprende al menos una región bicatenaria y una región no apareada, y en el que cada adaptador con apareamiento erróneo comprende un saliente de una sola base en 3' que comprende un

nucleótido "T" en el extremo bicatenario, que es complementario al saliente de una sola base en 3' que comprende un nucleótido "A" en los dúplex de polinucleótido diana, realizar una reacción inicial de extensión de cebador en la que un oligonucleótido cebador hibrida con una parte de adaptador de cada una de las construcciones adaptador-diana y se extiende por adición secuencial de nucleótidos para formar productos de extensión complementarios a al menos una cadena de cada una de las construcciones adaptador-diana, en el que los productos de extensión, y opcionalmente productos de amplificación derivados de los mismos, proporcionan en su conjunto una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3'; en el que los dúplex de polinucleótido diana a ligar son una mezcla compleja de fragmentos de ADN genómico que representan un genoma completo o esencialmente completo, y en el que la biblioteca de moléculas polinucleotídicas plantilla generada es representativa del genoma completo o esencialmente completo.

La invención también se refiere al uso de una biblioteca de moléculas polinucleotídicas plantilla preparadas según el método como un plantilla para la amplificación por PCR en fase sólida. Por ende, en una realización particular, la invención proporciona un método de amplificación de ácidos nucleicos en fase sólida de moléculas polinucleotídicas plantilla que comprende las etapas que consisten en:

preparar una biblioteca de moléculas polinucleotídicas de plantilla que tienen secuencias comunes en sus extremos 5' y 3' usando el método según el primer aspecto de la invención y llevar a cabo una reacción de amplificación de ácidos nucleicos en fase sólida, en el que se amplifican dichas moléculas polinucleotídicas plantilla.

La invención también se refiere al uso de una biblioteca de moléculas polinucleotídicas plantilla preparadas según el método de la invención como un plantilla para la amplificación del genoma completo. Por ende, en una realización particular, la invención proporciona un método de amplificación del genoma completo que comprende las etapas que consisten en:

usar el método según el primer aspecto de la invención para preparar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y 3' a partir de una mezcla compleja de fragmentos de genoma completo y llevar a cabo una reacción de amplificación de genoma completo en el que se amplifican dichas moléculas polinucleotídicas plantilla.

La invención también proporciona un kit para su uso en la preparación de una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y 3' en el que la secuencia común en el extremo 5' de cada plantilla individual en la biblioteca no es idéntica y no complementaria en su totalidad a la secuencia común en el extremo 3' de dicho plantilla, comprendiendo el kit polinucleótidos adaptadores con apareamiento erróneo y cebadores oligonucleotídicos como se define en las reivindicaciones capaces de hibridar con los polinucleótidos adaptadores con apareamiento erróneo.

Breve descripción de los dibujos

La Figura 1 ilustra diversos ejemplos de adaptadores con apareamiento erróneo bifurcados para su uso en el método de la invención, específicamente representan diferentes estructuras de extremo romo o saliente permisibles en el extremo "ligable" del adaptador. La Figura 1(e) ilustra esquemáticamente los componentes de secuencia de dos cadenas parcialmente complementarias (denominadas oligo A y oligo B) que forman el adaptador bifurcado cuando están hibridadas. El extremo 5' del oligo B es complementario (COMP) a una parte de la secuencia SEQ CEBADOR en el oligo A. El oligo A incluye un solo saliente de nucleótido "T" en el extremo 3'. El extremo 5' del oligo A está fosforilado. P representa un grupo fosfato terminal; X e Y representan funcionalidades de captura de superficie.

La Figura 2 ilustra una realización del método de la invención basado en el uso de los adaptadores bifurcados ilustrados en la Figura 1. La Figura 2(a) representa las etapas de fragmentación de una muestra compleja tal como ADN genómico para generar una pluralidad de fragmentos dúplex diana, ligamiento de fragmentos dúplex diana con adaptadores con apareamiento erróneo (bifurcados) para generar construcciones adaptador-plantilla y eliminar adaptadores no unidos. El adaptador bifurcado incluye un grupo biotina en el extremo 5' que no está ligado al fragmento diana para facilitar la captura en fase sólida de las construcciones adaptador-diana, p. ej., sobre perlas magnéticas de estreptavidina. La Figura 2(b) representa una reacción inicial de extensión con cebador en la que los cebadores hibridan con regiones de adaptador con apareamiento erróneo en cada cadena de una construcción adaptador-diana y se extienden para generar productos de extensión complementarios a cada cadena de la construcción adaptador-diana. Por simplicidad, el ligamiento y las etapas de extensión con cebador se ilustran para una sola construcción adaptador-diana.

La Figura 3 ilustra una realización alternativa de la invención en la que las construcciones adaptador-diana se someten a rondas múltiples de hibridación con cebador y extensión para generar copias múltiples monocatenarias de cada construcción adaptador-diana. Por simplicidad, las etapas de extensión con cebador se ilustran para una sola construcción adaptador-diana.

La Figura 4 ilustra otra realización adicional de la invención en la que las construcciones adaptador-diana se

someten a amplificación por PCR para generar múltiples copias bicatenarias de cada construcción adaptador-diana. Por simplicidad, la amplificación por PCR se ilustra para una sola construcción adaptador-diana.

5 La Figura 5 ilustra una realización de la invención, que representa etapas de fragmentación de una muestra compleja tal como ADN genómico para generar una pluralidad de fragmentos diana, ligamiento de los fragmentos diana con adaptadores con apareamiento erróneo (bifurcados) para generar construcciones adaptador-plantilla y posteriormente eliminar los adaptadores no unidos, en el que los adaptadores no incluyen un grupo biotina en el extremo 5'. Las construcciones adaptador-diana resultantes pueden someterse a amplificación por PCR para generar múltiples copias bicatenarias de cada construcción adaptador-diana. Por simplicidad, las etapas de ligamiento se ilustran para una sola construcción adaptador-diana.

15 La Figura 6 ilustra ejemplos adicionales de adaptadores con apareamiento erróneo bifurcados para su uso en el método de la invención, representando de nuevo los formatos romos y salientes en el extremo "ligable" permisibles del adaptador. La Figura 6(e) ilustra esquemáticamente las secuencias componentes presentes en las dos cadenas (denominadas oligo C y oligo B) que forman el adaptador cuando están hibridadas. P representa un grupo fosfato terminal; X e Y representan funcionalidades de captura de superficie.

20 La Figura 7 ilustra otra realización adicional de la invención basada en el uso de los adaptadores bifurcados ilustrados en la Figura 6. La Figura 7(a) representa etapas de fragmentación y ligamiento esencialmente similares a las ilustradas en la Figura 5. La Figura 7(b) representa una amplificación por PCR posterior usando cebadores PCR "con cola" que tienen secuencias terminales 3' complementarias a una secuencia en el extremo 5' del adaptador, e ilustra esquemáticamente la composición de secuencias de los productos de amplificación bicatenarios formados en la reacción de PCR. Por simplicidad, las etapas de ligamiento y amplificación por PCR se ilustran para una sola construcción adaptador-diana.

25 La Figura 8 ilustra realizaciones alternativas de adaptadores con apareamiento erróneo para su uso en el método de la invención. P representa un grupo fosfato terminal; X e Y representan funcionalidades de captura de superficie; X y Z representan modificaciones para evitar el ligamiento.

30 La Figura 9a ilustra otra realización adicional de la invención basada en el uso de adaptadores alternativos ilustrados en la Figura 8. La Figura 9(a) representa la fragmentación, el ligamiento y la retirada posterior de adaptadores no unidos. La Figura 9(b) representa la hibridación de cebadores de amplificación idénticos con una región bicatenaria del adaptador en cada cadena de la construcción adaptador-diana. Las construcciones adaptador-diana pueden amplificarse por PCR usando esta única especie de cebador. Por simplicidad, las etapas de ligamiento e hibridación con cebador se ilustran para una sola construcción adaptador-diana.

Descripción detallada de la invención

40 La invención proporciona un método, como se define en las reivindicaciones, de generación de una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y 3'. En este contexto, el término "común" se interpreta que significa común a todas las plantillas en la biblioteca. Como se explica adicionalmente con detalle a continuación, todos los plantillas en la biblioteca contendrán regiones de secuencia común en (o próximas a) sus extremos 5' y 3', en la que la secuencia común en el extremo 5' de cada plantilla individual en la biblioteca no es idéntica y no es complementaria por completo a una secuencia común en el extremo 45 3' de dicho plantilla.

50 El término "biblioteca" se refiere simplemente a una colección o pluralidad de moléculas plantilla que comparten secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3'. El uso del término "biblioteca" para referirse a una colección o pluralidad de moléculas plantilla no debe interpretarse en modo alguno que los plantillas que constituyen la biblioteca se deriven de una fuente particular, o que la "biblioteca" tenga una composición particular. A modo de ejemplo, el uso del término "biblioteca" no debe interpretarse en modo alguno que los plantillas individuales en la biblioteca hayan de ser de diferente secuencia de nucleótidos o que los plantillas estén relacionados en términos de secuencia y/o fuente.

55 La presente divulgación abarca la formación de bibliotecas "monoplantilla", que comprenden múltiples copias de un solo tipo de molécula plantilla, teniendo cada secuencia común en sus extremos 5' y 3', así como bibliotecas "complejas" en las que muchas, si no todas, de las moléculas plantilla individuales comprenden diferentes secuencias diana (como se define más adelante), aunque todas comparten secuencias comunes en sus extremos 5' y 3'. Dichas bibliotecas plantilla complejas pueden prepararse usando el método de la invención que se inicia a partir de una mezcla compleja de polinucleótidos diana tal como (pero no se limita a) fragmentos de ADN genómico aleatorios. La divulgación también se extiende a bibliotecas "complejas" formadas por una mezcla entre sí de 60 diversas bibliotecas "monoplantilla" individuales, cada una de las cuales se ha preparado por separado usando el método de la invención comenzando a partir de un solo tipo de molécula diana (es decir, un monoplantilla). En aspectos de la divulgación más del 50 %, o más del 60 %, o más del 70 %, o más del 80 %, o más del 90 %, o más 65 del 95 % de los plantillas polinucleotídicos individuales en una biblioteca compleja puede comprender diferentes

secuencias diana, aunque todas las plantillas en una biblioteca determinada compartirán secuencia común en sus extremos 5' y una secuencia común en sus extremos 3'.

El uso del término "plantilla" para referirse a moléculas polinucleotídicas individuales en la biblioteca indica simplemente que una o dos cadenas de los polinucleótidos en la biblioteca tienen la capacidad de actuar como plantillas para la polimerización de ácidos nucleicos dependiente de plantilla catalizada por una polimerasa. El uso de este término no debe considerarse como limitante del alcance de la invención con respecto a bibliotecas de polinucleótidos que se usan en realidad como plantillas en una reacción de polimerización catalizada por enzimas posterior.

La biblioteca se forma ligando moléculas polinucleotídicas con adaptador idénticas ("adaptadores con apareamiento erróneo", cuyas características generales se refieren a continuación) en los extremos 5' y 3' de uno o más dúplex polinucleotídicos diana (que pueden ser de secuencia conocida, parcialmente conocida o desconocida) para formar construcciones adaptador-diana y después llevar a cabo una reacción inicial de extensión con cebador en la que se forman productos de extensión complementarios a ambas cadenas de cada construcción adaptador-diana individual. Los productos de extensión con cebador resultantes y opcionalmente copias amplificadas de los mismos, proporcionan en su conjunto una biblioteca de polinucleótidos plantilla.

Cada cadena de cada molécula plantilla en la biblioteca formada en la reacción de extensión con cebador tendrá por lo tanto la siguiente estructura, cuando se observa como una sola cadena:

5'-[secuencia común I]-[secuencia diana]-[secuencia común II]-3'

En la que "secuencia común I" representa una secuencia derivada de copiar una primera cadena del adaptador con apareamiento erróneo y es común a todas las moléculas plantilla en la biblioteca generada en la reacción inicial de extensión con cebador;

"diana" representa una secuencia derivada de una cadena del dúplex polinucleotídico diana y puede ser diferente en moléculas plantilla individuales diferentes en la biblioteca; y

"secuencia común II" representa una secuencia derivada de copiar una segunda cadena de adaptador con apareamiento erróneo y es también común a todas las moléculas plantilla en la biblioteca generada en la reacción inicial de extensión con cebador.

Ya que la "secuencia común I" y la "secuencia común II" son comunes a todas las cadenas plantilla en la biblioteca, pueden incluir secuencias de unión a cebador "universal", permitiendo que todas las plantillas en la biblioteca se amplifiquen finalmente en un procedimiento de PCR en fase sólida usando cebadores universales.

Esta es una característica clave de la invención, sin embargo, las secuencias comunes terminales 5' y 3' denominadas "secuencia común I" y "secuencia común II" no son completamente complementarias entre sí, lo que significa que cada cadena plantilla individual puede contener diferentes secuencias de cebador universal (y no complementarias) en sus extremos 5' y 3'.

Esto resulta generalmente ventajoso para bibliotecas complejas de plantillas a amplificar por PCR (p. ej. amplificación del genoma completo) tanto si se realiza en solución o en un soporte sólido, incluir regiones de secuencia "diferente" en sus extremos 5' y 3', que sin embargo son comunes a todas las moléculas plantilla en la biblioteca, especialmente si los productos de amplificación se secuencian finalmente. Por ejemplo, la presencia de una secuencia única común en un extremo solo de cada plantilla en la biblioteca puede proporcionar un sitio de unión para un cebador de secuenciación, lo que permite que una cadena de cada plantilla en la forma amplificada de la biblioteca se secuencie en una sola reacción de secuenciación usando un solo tipo de cebador de secuenciación.

Normalmente, "secuencia común I" y "secuencia común II" consistirá en no más de 100, o no más de 50 o no más de 40 nucleótidos consecutivos en los extremos 5' y 3', respectivamente, de cada cadena de cada polinucleótido plantilla. La longitud exacta de las dos secuencias puede o no ser idéntica. Las secuencias de nucleótidos de la "secuencia común I" y "secuencia común II" en los polinucleótidos plantilla se determinará en parte por las secuencias de las cadenas adaptadoras ligadas a los polinucleótidos diana y en parte por la secuencia del cebador usado en la reacción inicial de extensión con cebador y cualquier ronda posterior de amplificación de ácidos nucleicos.

En realizaciones en las que el producto inicial de extensión con cebador se somete a amplificación adicional por PCR convencional, los productos de la reacción de amplificación serán polinucleótidos bicatenarios, cuya cadena tiene la estructura:

5'-[secuencia común I]-[secuencia diana]-[secuencia común II]-3'

Se apreciará que "secuencia común II" en los productos de amplificación pueden diferir algo con respecto a la "secuencia común II" presente en los productos de la reacción inicial de extensión con cebador, ya que esta última se determinará en parte por la secuencia del cebador de PCR usado para cebar la síntesis de una cadena polinucleotídica complementaria al producto inicial de extensión con cebador, mientras que la última se determinará exclusivamente copiando las secuencias adaptadoras en los extremos 3' de las construcciones adaptador-plantilla

- en la extensión con cebador inicial. Sin embargo, dado que el cebador de PCR se diseña para hibridar con una secuencia en los productos de extensión iniciales que es complementaria al adaptador 3', las dos formas de "secuencia común II" contendrán una secuencia idéntica, al menos en el extremo 3'. Puede incluirse una secuencia adicional en el extremo 5' de la "secuencia común II" en los productos amplificados, por ejemplo mediante el uso de
- 5 cebadores de PCR "con cola" como se describe con detalle más adelante. En otras realizaciones, las secuencias comunes presentes en los productos de amplificación pueden en realidad ser más cortas que las secuencias comunes que se incluyen en los adaptadores originalmente ligados a la diana.
- Las secuencias de nucleótidos exactas de las regiones comunes de las moléculas plantilla en la biblioteca no son
- 10 generalmente significativas para la invención y el usuario puede seleccionarlas. Las secuencias comunes han de comprender al menos secuencias "de unión a cebador" que permiten la hibridación específica de cebadores de amplificación cuando se usan los plantillas en una reacción de amplificación en fase sólida. Las secuencias de unión a cebador se determinan de este modo por la secuencia de los cebadores que va a usarse finalmente para la amplificación en fase sólida. La secuencia de estos cebadores, a su vez, se selecciona ventajosamente para evitar o
- 15 minimizar la unión de los cebadores a las partes diana de los plantillas en la biblioteca en las condiciones de la reacción de amplificación, pero no está particularmente limitado de otro modo. A modo de ejemplo, si las partes diana de los plantillas derivan de ADN genómico humano, entonces las secuencias de los cebadores a usar en la amplificación en fase sólida deben seleccionarse idealmente para minimizar la unión no específica a cualquier secuencia genómica humana.
- 20 Los polinucleótidos adaptadores usados en el método de la invención se refieren en la presente memoria como adaptadores "con apareamiento erróneo" ya que, como se explicará en detalle en la presente memoria, es esencial que los adaptadores incluyan una región de apareamiento erróneo de secuencia, es decir, no han de formarse por hibridación de cadenas polinucleotídicas completamente complementarias.
- 25 Los adaptadores con apareamiento erróneo para su uso en la invención se forman hibridando dos cadenas polinucleotídicas parcialmente complementarias para proporcionar, cuando las dos cadenas están hibridadas, al menos una región bicatenaria y al menos una región no apareada.
- 30 La "región bicatenaria" del adaptador es una región bicatenaria corta, que comprende normalmente 5 o más pares de bases consecutivos, formados por hibridación de las dos cadenas polinucleotídicas parcialmente complementarias. Este término se refiere simplemente a una región bicatenaria de ácido nucleico en la que las dos cadenas están hibridadas y no implica ninguna conformación estructural particular.
- 35 Generalmente resulta ventajoso que la región bicatenaria sea lo más corta posible sin pérdida de función. Por "función", en este contexto, se refiere a que la región bicatenaria forma un dúplex estable en condiciones de reacción convencionales para una reacción de ligamiento de ácidos nucleicos catalizada por enzimas, que conocerá bien el experto lector (p. ej. incubación a una temperatura en el intervalo de 4 °C a 25 °C en un tampón de ligamiento apropiado para la enzima), de modo que las dos cadenas que forman el adaptador permanecen parcialmente
- 40 hibridadas durante el ligamiento del adaptador a una molécula diana. No es absolutamente necesario que la región bicatenaria sea estable en condiciones normalmente usadas en las etapas de hibridación de la extensión con cebador o reacciones de PCR.
- Puesto que se ligan adaptadores idénticos en ambos extremos de cada molécula plantilla, la secuencia diana en
- 45 cada construcción adaptador-diana estará flanqueada por secuencias complementarias derivadas de la región bicatenaria de los adaptadores. Cuanto más larga sea la región bicatenaria, y por ende las secuencias complementarias derivadas de la misma en las construcciones adaptador-diana, mayor será la posibilidad de que la construcción adaptador-diana sea capaz de volverse a plegar y formar pares de bases en sí en estas regiones de autocomplementariedad interna en las condiciones de hibridación usadas en la extensión con cebador y/o PCR.
- 50 Generalmente se prefiere que la región bicatenaria tenga 20 o menos, 15 o menos, o 10 o menos pares de bases de longitud para reducir este efecto. La estabilidad de la región bicatenaria puede aumentarse, y por ende su longitud posiblemente reducirse, mediante la inclusión de nucleótidos no naturales que muestran formaciones de pares de bases más fuertes que los pares de bases de Watson-Crick convencionales.
- 55 Resulta preferente, pero no absolutamente esencial, que las dos cadenas del adaptador tengan una complementariedad del 100 % en la región bicatenaria. Se apreciará que puedan tolerarse uno o más apareamientos erróneos de nucleótidos en la región bicatenaria, siempre que las dos cadenas sean capaces de formar un dúplex estable en condiciones de ligamiento convencionales.
- 60 Los adaptadores para su uso en la invención incluirán generalmente una región bicatenaria adyacente al extremo "ligable" del adaptador, es decir, el extremo que está unido a un polinucleótido diana en la reacción de ligamiento. El extremo ligable comprende un saliente en 3' de una sola base que comprende un nucleótido "T" para facilitar/promover el ligamiento. El nucleótido terminal 5' en el extremo ligable del adaptador debe estar fosforilado para permitir la unión fosfodiéster a un grupo hidroxilo 3' en el polinucleótido diana.
- 65

La expresión "región no apareada" se refiere a una región del adaptador en el que las secuencias de las dos cadenas polinucleotídicas que forman el adaptador exhiben un grado de no complementariedad de tal manera que las dos cadenas no son capaces de hibridar entre sí en condiciones de hibridación convencionales para una extensión con cebador o reacción de PCR. Las dos cadenas en la región no apareada pueden exhibir cierto grado de hibridación en condiciones de reacción convencionales para una reacción de ligamiento catalizado por enzimas, siempre que las dos cadenas vuelvan a una forma monocatenaria en condiciones de hibridación.

Las condiciones encontradas durante las etapas de hibridación de una reacción de PCR son generalmente conocidas por los expertos en la materia, aunque las condiciones de hibridación exactas variarán de reacción a reacción (véase Sambrook *et al.*, 2001, *Molecular Cloning, A Laboratory Manual*, 3ª Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, NY; *Current Protocols*, eds Ausubel *et al.*). Normalmente dichas condiciones pueden comprender, pero no se limitan a, (después de una etapa de desnaturalización a una temperatura de aproximadamente 94 °C durante aproximadamente 1 minuto) exposición a una temperatura en el intervalo de 40 °C a 72 °C (preferentemente 50-68 °C) durante un periodo de aproximadamente 1 minuto en un tampón de reacción de PCR convencional.

Pueden usarse diferentes condiciones de hibridación para una sola reacción de extensión con cebador que no forma parte de una reacción de PCR (véase de nuevo Sambrook *et al.*, 2001, *Molecular Cloning, A Laboratory Manual*, 3ª Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, NY; *Current Protocols*, eds Ausubel *et al.*). Las condiciones para la hibridación con cebador en una sola extensión con cebador incluyen, por ejemplo, exposición a una temperatura en el intervalo de 30 a 37 °C en un tampón de extensión con cebador convencional. Se apreciará que diferentes enzimas, y por ende diferentes tampones de reacción, puedan usarse para una sola reacción de extensión con cebador en contraste con una reacción de PCR. No es un requisito el uso de una polimerasa termoestable para una sola reacción de extensión con cebador.

Debe entenderse que la "región no apareada" se proporciona por diferentes partes de las mismas dos cadenas polinucleotídicas que forman la(s) región(es) bicatenaria(s). Sin embargo, las partes de las dos cadenas que forman la región no apareada no están hibridadas en condiciones en las que otras partes de las mismas dos cadenas hibridan para formar una o más regiones bicatenarias. Para evitar cualquier tipo de dudas, debe entenderse que un saliente de una sola base o monocatenario en el extremo 5' o 3' de un dúplex polinucleotídico no constituye una "región no apareada" en el contexto de la presente invención.

Las partes de las dos cadenas que forman la región bicatenaria comprenden normalmente al menos 10, o al menos 15 o al menos 20 nucleótidos consecutivos en cada cadena. El límite inferior sobre la longitud de la región no apareada se determinará normalmente por la función, por ejemplo, la necesidad de proporcionar una secuencia adecuada para la unión de un cebador para la extensión con cebador, PCR y/o secuenciación. Teóricamente no existe un límite superior sobre la longitud de la región no apareada, excepto que en general sea ventajoso minimizar la longitud completa del adaptador, por ejemplo a fin de facilitar la separación de adaptadores no unidos de construcciones adaptador-diana después de la etapa de ligamiento. Por lo tanto, se prefiere que la región no apareada sea inferior a 50, o inferior a 40, o inferior a 30 o inferior a 25 nucleótidos consecutivos de longitud en cada cadena.

La longitud total de las dos cadenas que forman el adaptador se encontrará normalmente en el intervalo de 25 a 100 nucleótidos, más normalmente de 30 a 55 nucleótidos.

Las partes de las dos cadenas que forman la región no apareada deben preferentemente tener una longitud similar, aunque no es absolutamente esencial, siempre que la longitud de cada parte sea suficiente para cumplir la función deseada (p. ej., unión a cebador). Los inventores han demostrado mediante experimentos que las partes de las dos cadenas que forman la región no apareada pueden diferir hasta en 25 nucleótidos sin afectar negativamente a la función del adaptador.

Lo más preferentemente, las partes de las dos cadenas polinucleotídicas que forman la región no apareada se encontrarán completamente con apareamiento erróneo, o un 100 % de no complementariedad. Sin embargo, los lectores expertos apreciarán que algunas "coincidencias" de secuencia, es decir, un menor grado de no complementariedad puede tolerarse en esta región sin afectar la función en gran medida. Como se ha indicado anteriormente, el grado de apareamiento erróneo o no complementariedad de secuencia ha de ser tal que las dos cadenas en la región no apareada permanezcan en forma monocatenaria en condiciones de hibridación como se ha definido anteriormente.

La secuencia de nucleótidos exacta de los adaptadores no es generalmente esencial para la invención y puede seleccionarse por el usuario de tal manera que los elementos de secuencia deseados se incluyan finalmente en las secuencias comunes de la biblioteca de plantillas derivados de los adaptadores, por ejemplo para proporcionar sitios de unión para conjuntos particulares de cebadores de amplificación universales y/o cebadores de secuenciación. Pueden incluirse elementos de secuencia adicionales, por ejemplo para proporcionar sitios de unión para cebadores de secuenciación que finalmente se usarán en la secuenciación de moléculas plantilla en la biblioteca, o productos

derivados de amplificación de la biblioteca plantilla, por ejemplo en un soporte sólido. Los adaptadores pueden incluir además secuencias "etiqueta", que pueden usarse para etiquetar o marcar moléculas plantilla derivadas de una fuente particular. Las características y el uso generales de tales secuencias etiqueta se describen en la solicitud en trámite del solicitante publicada como WO 05/068656.

5

Si bien la secuencia de nucleótidos exacta del adaptador es generalmente no limitante para la invención, las secuencias de las cadenas individuales en la región no apareada debe ser de tal manera que ninguna cadena individual exhiba ninguna autocomplementariedad interna que pudiera conducir a una autohibridación, formación de estructuras de horquilla, etc. en condiciones de hibridación convencionales. La autohibridación de una cadena en la

10

región no apareada debe impedirse ya que esto puede evitar o reducir la unión específica de un cebador de amplificación a esta cadena.

Los adaptadores con apareamiento erróneo se forman preferentemente a partir de dos cadenas de ADN, pero pueden incluir mezclas de nucleótidos naturales y no naturales (p. ej., uno o más ribonucleótidos) unidos mediante una mezcla de enlaces de cadena principal fosfodiéster y no fosfodiéster. Otras modificaciones no nucleotídicas pueden incluirse tales como, por ejemplo, restos de biotina, grupos bloqueantes y restos de captura para la unión a una superficie sólida, como se analiza con detalle a continuación.

15

El uno o más "dúplex de polinucleótido diana" a los que se ligan los adaptadores puede ser cualquier molécula de polinucleótido que se desea amplificar mediante PCR en fase sólida, generalmente con una vista a la secuenciación. Los dúplex de polinucleótido diana pueden originarse en forma de ADN bicatenario (p. ej. fragmentos de ADN genómico) o pueden haberse originado en forma monocatenaria, como ADN y haberse convertido a ADNbc antes del ligamiento. La secuencia exacta de las moléculas diana no es generalmente significativa para la invención, y puede ser conocida o desconocida. Las moléculas de ADN modificadas que incluyen enlaces de cadena principal no natural y/o nucleótidos no naturales pueden servir como diana, siempre que las modificaciones no excluyan el ligamiento con adaptador y/o la copia en una reacción de extensión con cebador.

20

Aunque en teoría el método puede aplicarse a un único dúplex diana (es decir, una molécula bicatenaria individual), se usa una mezcla o una pluralidad de dúplex de polinucleótido diana. El método de la invención se aplica a una mezcla de diferentes moléculas diana que difieren entre sí con respecto a la secuencia de nucleótidos sobre la totalidad o parte de su longitud, p. ej., una mezcla compleja de plantillas. El método puede aplicarse a una pluralidad de moléculas diana derivadas de una fuente común, por ejemplo, una biblioteca de fragmentos de ADN genómico derivada de un individuo particular. En una realización preferida, los polinucleótidos diana comprenderán fragmentos aleatorios de ADN genómico humano. Los fragmentos se derivan de un genoma completo y de un individuo o

25

diversos individuos. Las moléculas diana de ADN pueden tratarse química o enzimáticamente antes de o posteriormente del ligamiento de las secuencias adaptadoras. Las técnicas para la fragmentación de ADN genómico incluyen, por ejemplo, digestión enzimática o cizallamiento mecánico.

El "ligamiento" de adaptadores a extremos 5' y 3' de cada polinucleótido diana implica la unión de dos cadenas polinucleotídicas del adaptador a un polinucleótido diana bicatenario de tal manera que se forman enlaces covalentes entre ambas cadenas de las dos moléculas bicatenarias. En este contexto "unión" significa un enlace covalente de dos cadenas polinucleotídicas que previamente no estaban unidas covalentemente. Preferentemente, dicha "unión" tendrá lugar por formación de un enlace fosfodiéster entre las dos cadenas polinucleotídicas pero pueden usarse otros medios de enlace covalente (p. ej., enlaces de cadena principal no fosfodiéster). Sin embargo,

30

35

es un requisito esencial que los enlaces covalentes formados en las reacciones de ligamiento permitan una lectura a través de una polimerasa, de tal manera que la construcción resultante pueda copiarse en una reacción de extensión con cebador usando cebadores que se unen a secuencias en las regiones de la construcción adaptador-diana que se derivan de las moléculas adaptadoras.

Las reacciones de ligamiento se catalizarán preferentemente por enzimas. La naturaleza de la enzima ligasa usada para el ligamiento enzimático no está particularmente limitada. También pueden usarse técnicas de ligamiento no enzimático (p. ej. ligamiento químico), siempre que el ligamiento no enzimático conduzca a la formación de un enlace covalente que permita la lectura a través de una polimerasa, de tal manera que la construcción resultante pueda copiarse en una reacción de extensión con cebador.

40

Los productos deseados de la reacción de ligamiento son construcciones adaptador-diana en las que los adaptadores idénticos están ligados en ambos extremos de cada polinucleótido diana, dada la estructura adaptador-diana-adaptador. Las condiciones de la reacción de ligamiento deben por lo tanto optimizarse para maximizar la formación de este producto, en preferencia para dianas que tienen un adaptador solo en un extremo.

45

Los productos de la reacción de ligamiento pueden someterse a etapas de purificación con el fin eliminar las moléculas adaptadoras no unidas antes de procesarse adicionalmente las construcciones adaptador-diana. Puede usarse cualquier técnica adecuada para eliminar el exceso de adaptadores no unidos, cuyos ejemplos preferidos se detallan a continuación.

50

55

60

65

Las construcciones adaptador-diana formadas en la reacción de ligamiento se someten después a una reacción inicial de extensión con cebador en la que un oligonucleótido cebador hibrida con una parte adaptadora de cada una de las construcciones adaptador-diana y se extiende por adición secuencial de nucleótidos en el extremo 3' hidroxilo libre del cebador para formar productos de extensión complementarios a al menos una cadena en cada una de las construcciones adaptador-diana.

El término reacción "inicial" de extensión con cebador se refiere a una reacción de extensión con cebador en la que los cebadores hibridan directamente con las construcciones adaptador-diana, en oposición a cualquiera de las cadenas complementarias formadas por extensión con cebador usando la construcción adaptador-diana como un plantilla o copias amplificadas de la construcción adaptador-diana. Es una característica clave del método de la invención que la reacción inicial de extensión con cebador se lleve a cabo usando un cebador "universal" que se une *específicamente* a una secuencia cognada en una parte adaptadora de la construcción adaptador-diana, y no se lleve a cabo usando un cebador específico de diana o una mezcla de cebadores aleatorios. El uso de un cebador específico de adaptador para la reacción inicial de extensión con cebador es clave para la formación de una biblioteca de plantillas que tienen secuencia común en el extremo 5' y secuencias comunes en el extremo 3'.

Los cebadores usados para la reacción inicial de extensión con cebador tendrán la capacidad de hibridar con cada cadena individual de construcciones adaptador-diana que tengan adaptadores ligados en ambos extremos y pueden extenderse para obtener dos productos de extensión con cebador distintos, uno complementario a cada cadena de la construcción. Por ello, en la realización más preferida, la reacción inicial de extensión con cebador dará como resultado la formación de productos de extensión con cebador complementarios a cada cadena de cada adaptador-diana.

En una realización preferida, el cebador usado en la reacción inicial de extensión con cebador hibridará con una secuencia de unión a cebador (en una cadena) en la región no apareada del adaptador.

El término "hibridación", como se usa en este contexto, se refiere a una unión/hibridación específica de secuencia del cebador con una secuencia de unión a cebador en una región adaptadora de la construcción adaptador-diana en las condiciones a usar para la etapa de hibridación con cebador de la reacción inicial de extensión con cebador.

Los productos de la reacción de extensión con cebador pueden someterse a condiciones desnaturalizantes convencionales a fin de separar los productos de extensión de las cadenas de las construcciones adaptador-diana. Opcionalmente, las cadenas de las construcciones adaptador-diana pueden eliminarse en esta fase. Los productos de extensión (con o sin las cadenas originales de las construcciones adaptador-diana) forman en su conjunto una biblioteca de polinucleótidos plantilla que puede usarse como plantillas para la PCR en fase sólida.

Si se desea, la reacción inicial de extensión con cebador puede repetirse una o más veces, a través de rondas de hibridación con cebador, extensión y desnaturalización, con el fin de formar múltiples copias de los mismos productos de extensión complementarios a las construcciones adaptador-diana.

En otras realizaciones, los productos de extensión iniciales pueden amplificarse mediante PCR en fase en solución convencional, como se describe con detalle a continuación. Los productos de dicha amplificación por PCR adicional pueden recogerse para formar una biblioteca de plantillas que comprenden "productos de amplificación derivados de" los productos de extensión con cebador iniciales. En una realización preferida, ambos cebadores usados para la amplificación por PCR adicional hibridarán con diferentes secuencias de unión a cebador sobre cadenas opuestas en la región no apareada del adaptador. Otras realizaciones pueden, sin embargo, basarse en el uso de un solo tipo de cebador de amplificación que hibrida con una secuencia de unión a cebador, en la región bicatenaria del adaptador. En realizaciones del método basado en amplificación por PCR, la reacción "inicial" de extensión con cebador se produce en el primer ciclo de PCR.

La inclusión de la etapa de extensión con cebador inicial (y opcionalmente rondas adicionales de amplificación por PCR) para formar copias complementarias de las construcciones adaptador-diana (antes de genoma completo o PCR en fase sólida) resulta ventajosa, por diversas razones. En primer lugar, la inclusión de la etapa de extensión con cebador y posterior amplificación por PCR, actúa como una etapa de enriquecimiento para seleccionar construcciones adaptador-diana, con adaptadores ligados en ambos extremos. Solo construcciones diana con adaptadores ligados en ambos extremos proporcionan plantillas eficaces para el genoma completo o PCR en fase sólida usando cebadores comunes o universales específicos para secuencias de unión a cebador en los adaptadores, por ello es ventajoso producir una biblioteca de plantillas que comprenda solo dianas ligadas doblemente antes de la amplificación en fase sólida o amplificación del genoma completo.

En segundo lugar, la inclusión de la etapa de extensión con cebador inicial y posterior amplificación por PCR, permite que la longitud de las secuencias comunes en los extremos 5' y 3' de la diana aumente antes de la PCR en fase sólida o de genoma completo. Como se ha indicado anteriormente, es generalmente ventajoso que la longitud de las moléculas adaptadoras sea lo más corta posible, para maximizar la eficacia del ligamiento y eliminación posterior de los adaptadores no unidos. Sin embargo, para los fines del PCR del genoma completo o en fase sólida

puede ser ventajoso tener secuencias más largas comunes o secuencias "universales" en los extremos 5' y 3' de los plantillas a amplificar. La inclusión de las etapas de extensión con cebador (y posterior amplificación) significa que la longitud de las secuencias comunes en un extremo (o en ambos extremos) de los polinucleótidos en la biblioteca de plantillas puede aumentarse después del ligamiento por inclusión de secuencias adicionales en los extremos 5' de los cebadores usados para la extensión con cebador (y posterior amplificación). El uso de dichos cebadores "con cola" se describe en detalle a continuación.

Diversas realizaciones específicas no limitantes del método de la invención se describirán ahora con más detalle con referencia a los dibujos adjuntos. A menos que se especifique otra cosa las características descritas como preferidas en relación con una realización específica de la invención se aplica *mutatis mutandis* con respecto a otras realizaciones específicas de la invención. Ejemplos de adaptadores no abarcados por las reivindicaciones son para referencia.

La Figura 1 ilustra diversas realizaciones de un tipo particular de adaptador con apareamiento erróneo para su uso en el método de la invención. El adaptador se forma hibridando dos oligonucleótidos monocatenarios, referidos en la presente memoria como "oligo A" y "oligo B". El oligo A y el oligo B pueden prepararse mediante técnicas de síntesis de oligonucleótidos automatizadas convencionales en el uso rutinario en la técnica. Los oligonucleótidos son parcialmente complementarios de tal manera que el extremo 3' del oligo A es complementario al extremo 5' del oligo B. El extremo 5' del oligo A y el extremo 3' del oligo B no son complementarios entre sí. Cuando las dos cadenas hibridan, la estructura resultante es bicatenaria en un extremo (la región bicatenaria) y monocatenaria en el otro extremo (la región no apareada) y se refiere en la presente memoria como un "adaptador bifurcado" (Fig. 1a). La región bicatenaria del adaptador bifurcado puede tener extremos romos (Fig. 1b) o puede tener un saliente. En el último caso, el saliente puede ser un saliente 3' (Fig. 1c) o un saliente 5' (Fig. 1d), y puede comprender un solo nucleótido o más de un nucleótido.

El extremo 5' de la parte bicatenaria del adaptador bifurcado está fosforilada, es decir, el extremo 5' del oligo B (Fig. 1a-d). La presencia del grupo fosfato 5' identifica a este como el extremo "ligable" del adaptador. El extremo 5' del oligo A puede estar biotinilado o llevar otra funcionalidad (representada por X) que le permita capturarse en una superficie, tal como una perla. Dichas funcionalidades alternativas distintas de las de la biotina son conocidas por los expertos en la materia. El extremo 3' del oligo B también puede estar biotinilado o llevar otra funcionalidad (representada por Y) que le permita capturarse en una superficie (Fig. 1d).

Los enlaces fosfodiéster que comprenden la cadena principal de los oligonucleótidos puede reemplazarse con enlaces no enzimáticamente escindibles tales como enlaces fosforotioato. Preferentemente solo el último, o el último y penúltimo, enlaces fosfodiéster en ambos extremos 3' y 5' de los oligonucleótidos se sustituirá con enlaces fosforotioato. En la realización más preferida de la invención, el oligo A contiene un grupo biotina en su extremo 5', y el oligo B está fosforilado en su extremo 5' y la parte bicatenaria del dúplex contiene un saliente en 3' de una sola base que comprende un nucleótido "T". El oligo A consiste en dos secuencias: una secuencia en el extremo 5' que es idéntica a la del cebador universal a usar para la amplificación por PCR, referida en la presente memoria como secuencia "CEBADOR 1", y en su extremo 3' una secuencia idéntica a la de un cebador de secuenciación universal, referida en la presente memoria como secuencia "SEQ CEBADOR", más un nucleótido "T" adicional en el extremo 3'. El oligo B también consiste en dos secuencias: una secuencia en su extremo 5' es complementaria a solo parte del extremo 3' de la secuencia SEQ CEBADOR en el Oligo A, excluyendo el saliente "T" del Oligo A, y una secuencia complementaria a la de un cebador de amplificación por PCR universal, referida en la presente memoria como "CEBADOR 2-comp" en su extremo 3' (Fig. 1e).

La Figura 2 ilustra una realización del método de la invención basada en el uso de los adaptadores bifurcados ilustrados en la Figura 1. Puede prepararse una mezcla de moléculas de ADN diana de secuencia diferente mezclando un número, más grande que uno, de moléculas de ADN individuales. En el procedimiento preferido, el ADN genómico está fragmentado en moléculas pequeñas, preferentemente menores de 1000 pares de bases de longitud. La fragmentación del ADN puede conseguirse mediante diversos métodos que incluyen: digestión enzimática, escisión química, sometimiento a ultrasonidos, nebulización o hidrocizalla, preferentemente nebulización.

El ADN fragmentado puede fabricarse con extremos romos mediante diversos métodos conocidos por los expertos en la materia. En el método preferido, los extremos del ADN fragmentado se reparan en el extremo con ADN polimerasa de T4 y Klenow polimerasa, un procedimiento bien conocido por los expertos en la materia, y después se fosforila con una enzima polinucleótido quinasa. Después se añade un solo desoxinucleótido "A" en ambos extremos 3' de las moléculas de ADN usando la enzima Taq polimerasa, produciendo un saliente en 3' de una base que es complementaria al otro saliente "T" de una base en 3' en el extremo bicatenario del adaptador bifurcado.

Después se realiza una reacción de ligamiento entre el adaptador bifurcado y los fragmentos de ADN usando una enzima ligasa adecuada (p. ej., ADN ligasa de T4) que une dos copias del adaptador a cada fragmento de ADN, en cualquier extremo, para formar construcciones adaptador-diana. Los productos de esta reacción pueden purificarse a partir de adaptador no ligado mediante diversos medios incluyendo cromatografía de inclusión por tamaños, preferentemente por electroforesis a través de un desbaste de gel de agarosa seguido por escisión de una parte de

la agarosa que contiene el ADN de mayor tamaño que el tamaño del adaptador (Fig. 2a).

Después de haber eliminado el exceso del adaptador, el ADN diana no ligado permanece además ligado a las construcciones adaptador-diana y este puede retirarse por captura selectiva de solo aquellas moléculas de ADN diana que tienen adaptador unido. La presencia de un grupo biotina en el extremo 5' de Oligo A del adaptador
5 permite que cualquier ADN diana ligado al adaptador se capture sobre una superficie recubierta con estreptavidina, una proteína que une selectiva y estrechamente biotina. La estreptavidina puede recubrirse en una superficie por medios conocidos por los expertos en la materia. En el método preferido, pueden usarse perlas magnéticas disponibles en el comercio que están recubiertas con estreptavidina para capturar construcciones adaptador-diana ligadas. La aplicación de un imán al lado de un tubo que contiene estas perlas las inmoviliza de tal manera que
10 pueden lavarse de las moléculas de ADN diana no ligadas (Fig. 2a).

Un oligonucleótido, referido en la presente memoria como CEBADOR 2, que hibrida con la secuencia "CEBADOR 2-comp" sobre la cadena oligo B de las construcciones adaptador-diana puede usarse en una reacción inicial de extensión con cebador para generar una copia complementaria de la cadena adaptador-diana unida a la perla. El
15 producto de extensión con cebador resultante forma un dúplex bicatenario con su cadena adaptador-diana complementaria unida a la perla y entonces puede aislarse y purificarse de la cadena adaptador-diana de la perla por desnaturalización (Fig. 2b).

Existen diversos métodos convencionales para separar la cadena de un dúplex de ADN por desnaturalización, incluyendo desnaturalización térmica, o preferentemente desnaturalización química en solución de hidróxido de sodio 100 mM. El pH de una solución de ADN monocatenario en hidróxido sódico recogido del sobrenadante de una suspensión de perlas magnéticas puede neutralizarse ajustándose con una solución de ácido apropiada, o preferentemente mediante un intercambio de tampón a través de una columna de cromatografía de exclusión por
20 tamaños preequilibrada en una solución tamponada. La solución resultante contiene una biblioteca de moléculas plantilla de ADN monocatenario las cuales comprenden en este orden: la secuencia 5'CEBADOR 2, ADN diana, el complemento de la secuencia SEQ CEBADOR, después el complemento de la secuencia CEBADOR 1. Esta biblioteca de plantillas puede usarse después para la amplificación por PCR en fase sólida usando los oligonucleótidos CEBADOR 1 y CEBADOR 2 inmovilizados. Sin embargo, se apreciará que la utilidad de la biblioteca no está limitada a PCR en fase sólida sino que se extiende a cualquier tipo de amplificación por PCR,
25 incluyendo amplificación del genoma completo realizado completamente en solución.

La Figura 3 ilustra una realización alternativa de la invención en la que las construcciones adaptador-diana preparadas como se ha descrito anteriormente con referencia a la Figura 2 se someten a rondas múltiples de hibridación con cebador y extensión para generar múltiples copias monocatenarias de cada construcción adaptador-
35 diana. En esta realización de la invención, la reacción inicial de extensión con cebador sobre las moléculas adaptador-plantilla inmovilizadas en perlas con CEBADOR 2 se reemplaza de hecho con una amplificación por PCR asimétrica con el oligonucleótido CEBADOR 2 (Fig. 3), siendo esto equivalente a rondas múltiples de la misma reacción de extensión con cebador. En la presente realización, se generan copias monocatenarias múltiples de las cadenas inmovilizadas en perlas en el sobrenadante de la suspensión con perlas debido al termociclado de PCR,
40 debido a que no es necesaria una etapa de desnaturalización distinta para recuperar las copias complementarias recién sintetizadas de las cadenas adaptador-diana inmovilizadas en perlas; las copias pueden purificarse del sobrenadante mediante métodos convencionales conocidos por los expertos en la materia.

En otra realización de la invención, ilustrada en la Figura 4, la reacción inicial de extensión con cebador de las construcciones adaptador-diana inmovilizadas en perlas con CEBADOR 2 forma parte de una amplificación por PCR (simétrica) convencional con los oligonucleótidos CEBADOR 2 y CEBADOR 1. En la presente realización, se generan copias bicatenarias múltiples de las cadenas inmovilizadas en perlas en el sobrenadante de la suspensión con perlas debido al termociclado de PCR, por ende no es necesaria una etapa de desnaturalización distinta para recuperar las copias complementarias recién sintetizadas de las cadenas adaptador-diana inmovilizadas en perlas;
50 las copias pueden purificarse del sobrenadante mediante métodos convencionales conocidos por los expertos en la materia.

En otra realización de la invención, ilustrada en la Figura 5, el adaptador bifurcado no contiene un grupo biotina en el extremo 5' de la cadena Oligo A. En la presente realización, el ADN fragmentado puede tener los extremos romos mediante diversos métodos conocidos por los expertos en la materia. En un método preferido, los extremos del fragmentado están reparados en el extremo con una ADN polimerasa de T4 y Klenow polimerasa y después fosforilados con la enzima polinucleótido quinasa. Acto seguido se añade un solo desoxinucleótido "A" en ambos extremos 3' de las moléculas de ADN con la enzima Taq polimerasa, produciendo un saliente 3' de una base que es complementario a un saliente "T" en 3' de una base en el extremo "ligable" bicatenario del adaptador bifurcado.
60 Después se realiza una reacción de ligamiento entre el adaptador bifurcado y los fragmentos de ADN, p. ej., usando la enzima ADN ligasa de T4 que une dos copias del adaptador con cada molécula plantilla de ADN, una en cada extremo.

Los productos de la reacción de ligamiento pueden purificarse del adaptador no ligado mediante diversos medios, incluyendo cromatografía de inclusión por tamaños, preferentemente por electroforesis a través de un desbaste de
65

gel de agarosa seguido por escisión de una parte de la agarosa que contiene el ADN de mayor tamaño que el tamaño del adaptador. Una alícuota del ADN purificado se usa después en una amplificación por PCR con los oligonucleótidos CEBADOR 2 y CEBADOR 1 (Fig. 5). El primer ciclo de PCR implicará una reacción de extensión con cebador inicial con el cebador 2 (no ilustrado). Los cebadores amplifican selectivamente aquellas moléculas de ADN plantilla que tengan adaptadores ligados en ambos extremos. El producto de la reacción es una biblioteca de moléculas plantilla bicatenarias, cada una de las cuales comprende en orden uno o dos de las cadenas híbridas: secuencia 5' CEBADOR 2, ADN diana, el complemento de la secuencia SEQ CEBADOR, y después el complemento de la secuencia CEBADOR 1. Esta biblioteca puede después usarse en una plataforma de PCR en fase sólida que contiene los oligonucleótidos CEBADOR 1 y CEBADOR 2 inmovilizados.

10

La Figura 6 ilustra ejemplos adicionales de adaptadores* con apareamiento erróneo bifurcados para su uso en el método de la invención. Ejemplos de adaptadores no abarcados por las reivindicaciones son para referencia.

En la presente realización, el adaptador bifurcado se forma hibridando dos oligonucleótidos monocatenarios, referidos en la presente memoria como "oligo C" y "oligo B". Los oligonucleótidos son parcialmente complementarios de tal manera que el extremo 3' del oligo C es complementario al extremo 5' del oligo B. El extremo 5' del oligo C y el extremo 3' del oligo B no son complementarios entre sí. Cuando los dos oligos hibridan la estructura resultante es bicatenaria en un extremo (región bicatenaria) y monocatenaria en el otro extremo (región no apareada) (Figura 6a). La región bicatenaria del adaptador bifurcado puede tener extremos romos (Fig. 6d) o puede tener un saliente. En el último caso, el saliente puede ser un saliente 3' (Fig. 6c) o un saliente 5' (Fig. 6b) y puede comprender una sola base o más de una base.

El extremo 5' de la región bicatenaria del adaptador bifurcado está fosforilada, es decir, el extremo 5' del "oligo B" (Fig. 6a-d) para proporcionar un extremo "ligable". El extremo 5' del oligo C puede estar biotinilado o llevar otra funcionalidad (X) que le permita capturarse sobre una superficie, tal como una perla. El extremo 3' del oligo B puede estar biotinilado o llevar otra funcionalidad (Y) que le permita capturarse sobre una superficie (Fig. 6d).

Los enlaces fosfodiéster que comprenden la cadena principal de los oligonucleótidos puede reemplazarse con enlaces no escindibles enzimáticamente, tales como enlaces fosforotioato. Preferentemente solo el último, o el último y penúltimo, enlaces fosfodiéster del extremo 3' y 5' de los oligonucleótido se sustituirá con enlaces fosforotioato. El oligo C consiste en solo una secuencia: una secuencia idéntica a la del cebador de secuenciación universal denotado "SEQ CEBADOR" (o idéntica a una parte del extremo 3' de la secuencia "SEQ CEBADOR"), más un nucleótido "T" adicional en el extremo 3'. El oligo B consiste en dos secuencias: una secuencia en su extremo 5' que es complementaria a solo una parte del extremo 3' de la secuencia SEQ CEBADOR en el Oligo C, excluyendo el saliente "T" del "Oligo C" y una secuencia en su extremo 3' que es complementaria a la de un cebador de amplificación por PCR universal, referido en la presente memoria secuencia "CEBADOR 2-comp", (Fig. 6e).

La Figura 7 ilustra una realización adicional de la invención basada en el uso de adaptadores bifurcados ilustrados en la Figura 6. En la presente realización, las construcciones adaptador-diana se preparan esencialmente como se ha descrito anteriormente con referencia a la Figura 5, excepto que se usan los adaptadores ilustrados en la Figura 6 (Fig. 7a).

Se usa una alícuota de las construcciones adaptador-diana purificadas en una amplificación por PCR en fase en solución convencional con oligonucleótidos cebadores "con cola". Los cebadores con cola son cebadores que solo hibridan mediante su extremo 3' con una secuencia diana, dejando una cola 5' no hibridada. La longitud y secuencia exacta de la cola no hibridada es no limitante, pero normalmente puede comprender de 10 a 50 nucleótidos, más normalmente de 15 a 30 nucleótidos. Cuando se usan amplificaciones por PCR, la ronda de amplificación por PCR inicial (es decir, la primera y segunda reacción extensión con cebador) se basa en la unión de los extremos 3' de los cebadores con cola con las secuencias de unión a cebador cognatas en las regiones adaptadoras de las construcciones adaptador-diana. Las colas no hibridadas 5' derivadas de los cebadores con cola actúan como plantillas en ciclos de PCR posteriores y por lo tanto se copian en productos de PCR bicatenarios resultantes.

En la presente realización, uno o ambos de los cebadores usados en la reacción de amplificación pueden ser cebadores "con cola". En una realización, los cebadores usados se denotan CEBADOR 3 y CEBADOR 4, en el que el CEBADOR 3 consiste en una secuencia con cola 5', y una secuencia 3' que es complementaria a la secuencia "CEBADOR 2-comp" en el adaptador bifurcado; y el CEBADOR 4 consiste en una secuencia con cola 5' y una secuencia 3' que es idéntica al extremo 5' de la secuencia SEQ CEBADOR presente en la región no apareada del adaptador bifurcado. Después de la amplificación por PCR, las secuencias con cola se incorporan en las copias de la construcción de ADN adaptador-diana.

60

En una realización de la invención, las secuencias con cola en CEBADOR 3 y CEBADOR 4 no son secuencias idénticas. La secuencia con cola del CEBADOR 3 y la secuencia con cola del CEBADOR 4 pueden usarse acto seguido para formar la secuencia de cebadores inmovilizados en superficie a usar en una plataforma de amplificación de ADN en fase sólida (Fig. 7b).

65

En otra realización de la invención, las secuencias con cola en el CEBADOR 3 y CEBADOR 4 son secuencias idénticas. Los productos de la PCR en fase solución tendrán por ende la misma secuencia en sus extremos: la secuencia con cola común del CEBADOR 3 y CEBADOR 4. Esta secuencia con cola común puede usarse acto seguido para formar la secuencia de un solo cebador inmovilizado sobre la superficie en una plataforma de
5 amplificación de ADN en fase sólida. La amplificación en superficie de la biblioteca de plantillas puede de este modo realizarse usando un solo cebador de PCR inmovilizado sobre la superficie.

La Figura 8 ilustra realizaciones alternativas de adaptadores con apareamiento erróneo para su uso en el método de la invención. Ejemplos de adaptadores no abarcados por las reivindicaciones son para referencia.

10

Estos adaptadores bifurcados "modificados" pueden diseñarse para permitir la amplificación en fase sólida de plantillas usando un solo cebador unido a superficie. El adaptador se forma hibridando dos oligonucleótidos monocatenarios, referidos en la presente memoria como "oligo D" y "oligo E". Los oligonucleótidos son parcialmente complementarios de tal manera que el extremo 3' del oligo D es complementario al extremo 5' del oligo E, y el
15 extremo 5' del oligo D es complementario al extremo 3' del oligo E, sin embargo, las partes centrales del oligo D y oligo E no son complementarias. Cuando el oligo D y el oligo E hibridan la estructura resultante es bicatenaria en ambos extremos (regiones bicatenarias) y monocatenaria en el centro (región no apareada) y se refiere en la presente memoria como "adaptador bifurcado modificado" (Fig. 8a).

20 Un extremo del adaptador bifurcado modificado se modifica para impedir el ligamiento de una molécula de ADN en su extremo. Dichas modificaciones son conocidas por los expertos en la materia y pueden incluir, por ejemplo, la presencia de un saliente 5' o 3'. El otro extremo "ligable" puede ser un extremo romo (Fig. 8d) o puede tener un saliente. En el último caso, el saliente puede ser un saliente 3' (Fig. 8c) o un saliente 5' (Fig. 8b) y puede comprender una sola base o más de una base. La cadena 5' del extremo ligable está fosforilada, es decir, el extremo
25 5' del oligo E (Figura 8a-d). El extremo 5' del oligo D puede estar biotinilado o llevar otra funcionalidad que permita que se capture sobre una superficie, tal como una perla. El extremo 3' del oligo E puede estar biotinilado o llevar otra funcionalidad que permita que se capture en una superficie (Fig. 8d). Las modificaciones para impedir el ligamiento (Z, W) pueden ser idénticas o diferentes a las funcionalidades de captura de superficie (X, Y).

30 Los enlaces fosfodiéster que comprenden la cadena principal de los oligonucleótidos pueden reemplazarse con enlaces no enzimáticamente escindibles tales como enlaces fosforotioato. Preferentemente solo el último, o el último y el penúltimo, enlaces fosfodiéster en ambos extremos 3' y 5' de los oligonucleótidos se sustituirán con enlaces fosforotioato.

35 En la realización preferida de la invención el oligo E está fosforilado en su extremo 5' y el extremo 3' del oligo D contiene un saliente 3' de una sola base que comprende un nucleótido "T". El oligo D consiste en dos secuencias: una secuencia en su extremo 5' que es idéntica a la de un cebador de amplificación por PCR universal, referido en la presente memoria como secuencia "CEBADOR 5", próxima a una secuencia idéntica a la de un cebador de secuenciación universal referido como secuencia "SEQ CEBADOR" más el nucleótido "T" adicional en el extremo 3'.
40 El Oligo E consiste en tres secuencias: una secuencia en su extremo 5' que es complementaria a solo una parte del extremo 3' de la secuencia SEQ CEBADOR en el Oligo D, excluyendo el saliente "T" del Oligo D, una secuencia central no complementaria a cualquier parte del Oligo D y un extremo 3' que es complementario a la secuencia "CEBADOR 5" del Oligo D (Fig. 8e).

45 La Figura 9 ilustra otra realización adicional de la invención basada en el uso de adaptadores alternativos ilustrados en la Figura 8. En la presente realización, pueden prepararse construcciones adaptador-diana esencialmente como se ha descrito anteriormente en relación con la Figura 5, a excepción de que se usen los adaptadores bifurcados modificados ilustrados en la Figura 8. Se usa una alícuota de las construcciones adaptador-diana purificadas en una amplificación por PCR en fase en solución usando el oligonucleótido CEBADOR 5 para amplificar selectivamente
50 aquellos productos de ligamiento que tengan el adaptador modificado en ambos extremos (Fig. 9b). El producto de la PCR en fase en solución puede purificarse y amplificarse en una plataforma de PCR en fase sólida con un solo cebador inmovilizado, p. ej., CEBADOR 5. La inclusión de la secuencia con apareamiento erróneo en el oligo E garantiza que todos los productos de esta amplificación en fase sólida contendrán secuencias de unión a cebadores de secuenciación comunes solo en una cadena, permitiendo la secuenciación usando un cebador de secuenciación
55 universal que hibrida con esta secuencia común.

Uso de la biblioteca de plantillas

Las bibliotecas de plantillas preparadas según el método de la invención pueden usarse esencialmente en cualquier
60 método de análisis de ácidos nucleicos que requiera amplificación adicional de los plantillas y/o secuenciación de los plantillas o productos de amplificación de los mismos. Usos a modo de ejemplo de bibliotecas plantilla incluyen, pero no se limitan a, proporcionar plantillas para amplificación del genoma completo y asimismo amplificación por PCR en fase sólida (de bibliotecas plantilla monoplantilla o complejas). Un uso particularmente preferido es la amplificación del genoma completo llevada a cabo en un soporte sólido.

65

Amplificación del genoma completo

Las bibliotecas de plantillas preparadas según el método de la invención que comienza a partir de una mezcla compleja de fragmentos de ADN genómico que representan un genoma completo o esencialmente completo proporcionan plantillas adecuadas para amplificación denominada de "genoma completo". La expresión "amplificación de genoma completo" se refiere a una reacción de amplificación de ácidos nucleicos (p. ej. PCR) en la que el plantilla a amplificar comprende una mezcla compleja de fragmentos de ácidos nucleicos representativa de un genoma completo (o esencialmente completo).

10 Amplificación en fase sólida

Una vez formada, la biblioteca de plantillas preparada según los métodos descritos anteriormente puede usarse para la amplificación de ácidos nucleicos en fase sólida.

15 De este modo, en aspectos adicionales, la invención proporciona un método de amplificación de ácidos nucleicos en fase sólida de moléculas polinucleotídicas plantilla que comprende:
preparar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y 3' usando un método según el primer aspecto de la invención descrito en la presente memoria y llevar a cabo una reacción de amplificación de ácidos nucleicos en fase sólida en la que se amplifican dichas moléculas polinucleotídicas plantilla.

La expresión "amplificación en fase sólida", como se usa en la presente memoria, se refiere a cualquier reacción de amplificación de ácidos nucleicos llevada a cabo o en asociación con un soporte sólido de tal manera que la totalidad o una parte de los productos amplificados se inmovilizan sobre el soporte sólido a medida que se forman.
25 En particular, el término abarca la reacción en cadena de la polimerasa en fase sólida (PCR en fase sólida), que es una reacción análoga a la PCR en fase en solución convencional, excepto que uno o ambos cebadores de amplificación directos o inversos están inmovilizados en el soporte sólido.

Si bien la invención abarca métodos de amplificación en "fase sólida" en los que solo un cebador de amplificación se inmoviliza (estando el otro cebador normalmente presente en solución libre), se prefiere que el soporte sólido se proporcione con los cebadores tanto directo como inverso inmovilizados. En la práctica, habrá una "pluralidad" de cebadores directos idénticos y/o una "pluralidad" de cebadores inversos idénticos inmovilizados en el soporte sólido, ya que el proceso de la PCR requiere un exceso de cebadores para mantener la amplificación. Las referencias en la presente memoria a cebadores directos e inversos deben interpretarse por consiguiente por abarcar una "pluralidad" de dichos cebadores a menos que el contexto indique lo contrario.

Como apreciará un experto lector, cualquier reacción por PCR determinada requiere al menos un tipo de cebador directo y al menos un tipo de cebador inverso específico para el plantilla a amplificar. Sin embargo, en ciertas realizaciones, los cebadores directos e inversos pueden comprender partes específicas de plantilla de secuencia idéntica, y pueden tener secuencia y estructura de nucleótidos completamente idénticas (incluyendo cualquier modificación no nucleotídica). En otras palabras, es posible llevar a cabo una amplificación en fase sólida usando solo un tipo de cebador, y tales métodos de solo un cebador están abarcados dentro del alcance de la invención. Otras realizaciones pueden usar cebadores directos e inversos que contienen secuencias específicas plantilla idénticas pero que difieren en otras características estructurales. Por ejemplo un tipo de cebador puede contener una modificación no nucleotídica que no está presente en el otro.

En otras realizaciones de la invención, los cebadores directo e inversos pueden contener partes específicas de plantilla de diferente secuencia.

50 En todas las realizaciones de la invención, los cebadores de amplificación para PCR en fase sólida se inmovilizan preferentemente por unión covalente al soporte sólido en el extremo 5' del cebador o cerca del mismo, dejando la parte específica de plantilla del cebador libre para su hibridación con su plantilla cognato y el grupo hidroxilo 3' libre para su extensión con cebador. Para este fin puede usarse cualquier medio de unión covalente adecuado conocido en la técnica. La química de unión seleccionada dependerá de la naturaleza del soporte sólido, y cualquier derivatización o funcionalización aplicada al ello. El propio cebador puede incluir un resto, que puede ser una modificación química no nucleotídica, para facilitar la unión. En una realización particularmente preferida, el cebador puede incluir un nucleófilo que contiene azufre, tal como fosforotioato o tiosfosfato, en el extremo 5'. En el caso de hidrogeles de poliacrilamida con soporte sólido (como se describe a continuación), este nucleófilo se unirá a un grupo "C" presente en el hidrogel. Los medios más preferidos de unión de cebadores y plantillas a un soporte sólido es por unión fosforotioato 5' a un hidrogel compuesto por acrilamida polimerizada y N-(5-bromoacetamidilpentil)acrilamida (BRAPA).

Se prefiere el uso de la biblioteca de plantillas preparada según la invención para preparar series agregadas de colonias de ácidos nucleicos, análogas a las descritas en los documentos WO 00/18957 y WO 98/44151, por amplificación por PCR en fase sólida. Los términos "agregado" y "colonia" se usan indistintamente en la presente memoria para referirse a un sitio discreto en un soporte sólido compuesto por una pluralidad de cadenas de ácido

nucleico inmovilizadas idénticas y una pluralidad de cadenas de ácido nucleico complementarias inmovilizadas idénticas. La expresión "serie agregada" se refiere a una serie formada por dichos agregados o colonias. En este contexto, el término "serie" no debe entenderse por requerir una disposición ordenada de agregados.

5 Uso en secuenciación/métodos de secuenciación

La invención también abarca métodos de secuenciación de ácidos nucleicos amplificados generados por amplificación de genoma completo o en fase sólida. De este modo, la invención proporciona un método de secuenciación de ácidos nucleicos que comprende amplificar una biblioteca de plantillas de ácidos nucleicos usando
10 amplificación del genoma completo o en fase sólida como se ha descrito anteriormente y llevar a cabo una reacción de secuenciación de ácidos nucleicos para determinar la secuencia de la totalidad o parte de al menos una cadena de ácido nucleico amplificada producida en la reacción de amplificación del genoma completo o en fase sólida.

La secuenciación puede realizarse usando cualquier técnica de "secuenciación por síntesis" adecuada en la que los
15 nucleótidos se añaden sucesivamente al grupo hidroxilo 3' libre, dando como resultado la síntesis de una cadena polinucleotídica en dirección 5' a 3'. La naturaleza del nucleótido añadido se determina preferentemente después de cada adición de nucleótido.

El punto de inicio para la reacción de secuenciación puede proporcionar hibridando un cebador de secuenciación a
20 un producto de la reacción de amplificación de genoma completo o en fase sólida. En ese sentido, uno o ambos adaptadores añadidos durante la formación de la biblioteca de plantillas puede incluir una secuencia de nucleótidos que permita la hibridación de un cebador de secuenciación con productos amplificados derivados por la amplificación del genoma completo o en fase sólida de la biblioteca de plantillas.

Los productos de reacciones de amplificación en fase sólida en la que los cebadores de amplificación directos e
25 inversos se inmovilizan covalentemente en la superficie sólida también se denominan estructuras "puente" formadas por hibridación de pares de cadenas polinucleotídicas inmovilizadas y cadenas complementarias inmovilizadas, uniéndose ambas cadenas a un soporte sólido en el extremo 5'. Las series compuestas por tales estructuras puente proporcionan plantillas ineficaces para la secuenciación de ácidos nucleicos, ya que la hibridación de un cebador de
30 secuenciación convencional con una de las cadenas inmovilizadas no se ve favorecida en comparación con la hibridación de esta cadena con su cadena complementaria inmovilizada en condiciones de hibridación convencionales.

Con el fin de proporcionar plantillas más adecuadas para la secuenciación de ácidos nucleicos se prefiere eliminar
35 esencialmente toda o al menos una parte de una de las cadenas inmovilizadas en la estructura "puente" para generar un plantilla que es al menos parcialmente monocatenario. La parte del plantilla que es monocatenaria estará disponible así para la hibridación con un cebador de secuenciación. El proceso de eliminación de la totalidad o una parte de la cadena inmovilizada en una estructura de ácido nucleico bicatenaria "puente" puede referirse en la presente memoria como "linealización".

Las estructuras de plantilla "puente" pueden linealizarse escindiendo una o ambas cadenas con una endonucleasa
40 de restricción o por escisión de una cadena con una endonucleasa de corte. Otros métodos de escisión pueden usarse como una alternativa a las enzimas de restricción o enzimas de corte, incluyendo, entre otras, escisión química (p. ejemplo escisión de un enlace diol con peryodato), escisión de sitios abásicos por escisión con
45 endonucleasa, o por exposición al calor o a álcali, escisión de ribonucleótidos incorporados en productos de amplificación compuestos por desoxirribonucleótidos, escisión fotoquímica o escisión de un enlace peptídico.

Se apreciará que una etapa de linealización puede no ser esencial si la reacción de amplificación en fase sólida se
50 realiza solo con un cebador inmovilizado covalentemente y el otro en solución libre.

Con el fin de generar un plantilla linealizado adecuado para la secuenciación es necesario eliminar cantidades
"desiguales" de las cadenas complementarias en la estructura puente formada por amplificación para dejar un
55 plantilla linealizado para la secuenciación que es completa o parcialmente monocatenaria. Lo más preferentemente, una cadena de la estructura puente está esencial o completamente eliminada.

Después de la etapa de escisión, independientemente del método usado de escisión, el producto de reacción de
escisión puede someterse a condiciones desnaturizantes con el fin de eliminar la parte(s) de la(s) cadena(s)
60 escindida(s) que no está(n) unida(s) a un soporte sólido. Las condiciones de desnaturización adecuadas resultarán evidentes para el lector experto con referencia a los protocolos de biología molecular convencionales (Sambrook *et al.*, 2001, *Molecular Cloning, A Laboratory Manual*, 3ª Ed, Cold Spring Harbor Laboratory Press, Cold Spring Harbor Laboratory Press, NY; *Current Protocols*, eds Ausubel *et al.*).

La desnaturización (y posterior rehibridación de las cadenas escindidas) da como resultado la producción de un
65 plantilla de secuenciación que es parcial o esencialmente monocatenario. Tras lo cual puede iniciarse una reacción de secuenciación por hibridación de un cebador de secuenciación con la parte monocatenaria del plantilla.

De este modo, la reacción de secuenciación de ácidos nucleicos puede comprender hibridar un cebador de secuenciación con una región monocatenaria de un producto de amplificación linealizado, incorporar secuencialmente uno o más nucleótidos en una cadena polinucleotídica complementaria a la región de la cadena plantilla amplificada a secuenciar, identificar la base presente en uno más de los nucleótidos incorporados y por lo tanto determinar la secuencia de una región de la cadena plantilla.

Un método de secuenciación preferido que puede usarse según la invención se basa en el uso de nucleótidos modificados que pueden actuar como terminadores de cadena. Una vez que el nucleótido modificado se ha incorporado en la cadena polinucleotídica en crecimiento complementaria a la región de un plantilla a secuenciar no hay grupo OH 3' libre disponible para dirigir la extensión de secuencia adicional y por lo tanto la polimerasa no puede añadir otros nucleótidos. Una vez que la naturaleza de la base incorporada en la cadena en crecimiento se ha determinado, el bloque 3' puede eliminarse para permitir la adición del próximo nucleótido sucesivo. Al ordenar los productos derivados usando estos nucleótidos modificados es posible deducir la secuencia de ADN del plantilla de ADN. Tales reacciones pueden realizarse en un único experimento si cada uno de los nucleótidos modificados tiene unida una etiqueta diferente, conocida por corresponderse con la base particular, para facilitar la discriminación entre las bases añadidas en cada etapa de incorporación. Alternativamente, puede llevarse a cabo una reacción distinta que contiene cada uno de los nucleótidos modificados por separado.

Los nucleótidos modificados pueden llevar una etiqueta para facilitar su detección. Preferentemente, es una etiqueta fluorescente. Cada tipo de nucleótido puede llevar una etiqueta fluorescente diferente. Sin embargo, la etiqueta detectable no tiene que ser una etiqueta fluorescente. Puede usarse cualquier etiqueta que permita la detección de un nucleótido incorporado.

Un método para detectar nucleótidos marcados de manera fluorescente comprende usar luz láser de una longitud de onda específica para los nucleótidos marcados, o el uso de fuentes adecuadas de iluminación. La fluorescencia de la etiqueta en el nucleótido puede detectarse mediante una cámara CCD o por cualquier medio de detección adecuado.

La invención no tiene por objeto limitar el uso del método de secuenciación indicado anteriormente, ya que puede utilizarse esencialmente cualquier metodología de secuenciación que se basa en la incorporación sucesiva de nucleótidos en una cadena polinucleotídica. Las técnicas alternativas adecuadas incluyen, por ejemplo, Pyrosequencing™, FISSEQ (secuenciación de fluorescencia *in situ*), MPSS (secuenciación de marca masivamente paralela) y secuenciación por métodos basados en ligamiento.

El polinucleótido diana a secuenciar usando el método de la invención puede ser cualquier polinucleótido que se desee secuenciar. Usando el método de preparación de bibliotecas de plantillas descrito en detalle en la presente memoria es posible preparar bibliotecas de plantillas a partir de un polinucleótido diana monocatenario o bicatenario de secuencia conocida, desconocida o parcialmente conocida. Con el uso de series agregadas preparadas por amplificación en fase sólida es posible secuenciar dianas múltiples de la misma secuencia o diferente en paralelo.

Kits

La invención también se refiere a kits como se define en las reivindicaciones para su uso en la preparación de bibliotecas de polinucleótidos plantilla usando el método del primer aspecto de la invención.

Las realizaciones preferidas de kits comprenden al menos un suministro de un adaptador con apareamiento erróneo como se define en la presente memoria, más un suministro de al menos un cebador de amplificación que es capaz de hibridar con el adaptador con apareamiento erróneo y cebar la síntesis de un producto de extensión, cuyo producto de extensión incluiría cualquier secuencia diana ligada al adaptador cuando el adaptador está en uso.

Las características preferidas de los adaptadores "con apareamiento erróneo" para la inclusión en el kit son como se describe en cualquier parte de la presente memoria con respecto a los otros aspectos de la invención. La estructura y propiedades de cebadores de amplificación serán conocidas por los expertos en la materia. Los cebadores adecuados de secuencia de nucleótidos apropiada para su uso con los adaptadores incluidos en el kit pueden prepararse fácilmente usando equipos de síntesis de ácidos nucleicos automatizados convencionales y reactivos de uso rutinario en la técnica. El kit puede incluir un suministro de un solo tipo de cebador o suministros distintos (o incluso una mezcla) de dos cebadores diferentes, por ejemplo, un par de cebadores de PCR adecuados para amplificación por PCR de plantillas modificados con el adaptador con apareamiento erróneo en fase en solución y/o en un soporte sólido adecuado (es decir, PCR en fase sólida).

En una realización, el kit puede incluir suministros de diferentes pares de cebadores para su uso en PCR en fase en solución y en fase sólida. En este contexto, los pares de cebadores "diferentes" pueden ser de secuencia nucleotídica esencialmente idéntica pero diferir con respecto a alguna otra característica o modificación, tal como por ejemplo restos de captura de superficie, etc. En otras realizaciones, el kit puede incluir un suministro de cebadores para su uso en una reacción inicial de extensión con cebador y un par (o pares) de cebador diferente para

amplificación por PCR en solución y/o en fase sólida.

Los adaptadores y/o cebadores pueden suministrarse en los kits listos para su uso, o más preferentemente como concentrados que requieren dilución antes de su uso, o incluso en una forma liofilizada o seca que requiere 5 reconstrucción antes de su uso. Si se requiere, los kits pueden incluir además un suministro de un diluyente adecuado para dilución o reconstitución de los cebadores. Opcionalmente, los kits pueden comprender además suministros de reactivos, tampones, enzimas, dNTP, etc., para su uso en la realización de amplificación por PCR. Ejemplos adecuados (pero no limitantes) de dichos reactivos se describen en la sección Materiales y métodos de los 10 Ejemplos anexos. Otros componentes que pueden suministrarse opcionalmente en el kit incluyen cebadores de secuenciación "universales" adecuados para secuenciar plantillas preparados usando los cebadores y adaptadores con apareamiento erróneo.

La invención se entenderá además con referencia al siguiente ejemplo experimental no limitante:

15 **Ejemplo**

Esquema experimental

Los siguientes detalles experimentales describen la exposición completa de una realización de la invención como se 20 ha descrito anteriormente. La fuente de ADN usada es ADN de la estirpe celular humana purificada suministrada por Coriell Cell Repositories, Camden, NJ 08103 USA, catálogo n.º NA07055. El ADN se prepara primero para el ligamiento a adaptadores bifurcados por: fragmentación del ADN por nebulización, reparación en los extremos de los extremos de ADN para que tengan extremos romos y fosforilación, después la adición de un solo nucleótido "A" en los extremos 3' de los fragmentos de ADN humano. La reacción de ligamiento se realiza con el ADN fragmentado 25 preparado y los adaptadores preformados hibridando "Oligo A" y "Oligo B" (secuencias dadas a continuación). El producto de la reacción se aísla/purifica del adaptador sin ligar por electroforesis en gel. Finalmente, el producto de la reacción de ligamiento se somete a ciclos de PCR para amplificar selectivamente el producto ligado que contiene el adaptador en ambos extremos de los fragmentos.

30 Materiales y métodos

Nebulización

Materiales:

35

[0126]

- ADN genómico humano (1 mg/ml) Coriell NA07055
- Tampón (glicerol 53,1 ml, agua 42,1 ml, TrisHCl 1 M pH 7,5 3,7 ml, EDTA 0,5 M 1,1 ml)
- 40 • Nebulizador Invitrogen (N.º K7025-05)
- Kit de purificación de PCR de columnas Qiagen (N.º 28104)

Mezcla: 25 µl (5 microgramos) de ADN 725 µl de tampón

45 Procedimiento:

La solución de ADN enfriada se fragmentó en el nebulizador sobre hielo durante 5 a 6 minutos bajo al menos 32 psi de presión. El volumen recuperado (normalmente en algún punto entre 400 y 600 µl) se dividió en 3 alícuotas y se purificó con un kit de purificación de PCR de Qiagen, aunque usando solo una columna, y finalmente se eluyó en 30 50 µl de EB (Qiagen).

Reparación de extremos:

Materiales:

55

- | | |
|--|--------------------------|
| • ADN Polimerasa de T4 | NEB N.º M0203S |
| • 10xNEB 2 tampón | NEB N.º M7002S |
| • 100x BSA | NEB N.º M9001S |
| • mezcla de dNTP (10 mM cada una) | NEB N.º N0447S |
| • fragmento largo Poli I de ADN de <i>E. coli</i> | (Klenow, NEB N.º M0210S) |
| • polinucleótido quinasa de T4 | NEB N.º M0201S |
| • tampón T4 PNK | NEB N.º M0201S |
| • ATP 100 mM | |
| • kit de purificación de PCR de columnas (N.º 28104) | |

ES 2 732 253 T3

Qiagen

La mezcla de reparación de extremos se ensambló de la siguiente manera:

ADN	30 µl
Agua	12 µ
10xNEB2	5 µl
100xBSA	0,5 µl
10 mM dNTP	2 µl
ADN pol T4 (3U/µl)	5 µl
	50 µl total

5 La reacción se incubó durante 15 min a temperatura ambiente, y después se añadió 1 µl de fragmento grande de Pol I de ADN de *E. coli* (Klenow) y la reacción se incubó durante 15 min adicionales a temperatura ambiente. El ADN se purificó de las enzimas, tampón, etc. cargando la mezcla de reacción en una columna Qiagen, finalmente eluyendo en 30 µl de EB. Los extremos 5' del ADN se fosforilaron después usando polinucleótido quinasa de la siguiente manera:

10

ADN	30 µl
Agua	9,5 µl
Tampón 10xPNK	5 µl
ATP 100 mM	0,5 µl
PNK T4(10U/µl)	5 µl
	50 µl total

La reacción se incubó durante 30 minutos a 37 °C, después se inactivó con calor a 65 °C durante 20 min. Acto seguido se purificó el ADN de las enzimas, tampón, etc. cargando la mezcla de reacción en una columna Quiagen, finalmente eluyendo en 30 µl de EB. Se agruparon tres grupos distintos para dar 90 µl en total.

15

Reacción de cola A

Materiales:

- Taq ADN polimerasa
 - 10 x tampón termopol
 - dATP 1 mM
 - kit de purificación de PCR de columna Qiagen
- NEB N.º M0267L
NEB N.º B9004S
Amersham-Pharmacia N.º 272050
(N.º 28004)

20

Se ensambló la siguiente mezcla de reacción:

ADN	30 µl
10x tampón termopol	5 µ
dATP 1 mM	10 µl
Taq pol (5U/µl)	3 µl
	~50 µl total

25 La reacción se incubó durante 30 min a 70 °C, después el ADN se purificó de las enzimas, tampón, etc., cargando la mezcla de reacción en una columna Qiagen MinElute, finalmente eluyendo en 10 µl de EB.

Hibridación del adaptador bifurcado

Materiales:

30

- "Oligo A" y "OligoB"
- Tris 50 mM/NaCl 50 mM pH 7

35

- Máquina de PCR

Oligo A 100 µM	20 µl
Oligo B 100 µM	20 µl
Tris/NaCl	10 µl
	50 µl a 40 µM de dúplex en Tris 10 mM/NaCl 10 mM pH 7,5

ES 2 732 253 T3

Oligo A: 5'ACACTCTTTCCCTACACGACGCTCTTCCGATC_xT (x = enlace fosforotioato) (SEQ ID NO: 1)

Oligo B: 5'Fosfato-GATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG (SEQ ID NO: 2)

5 Las cadenas adaptadoras hibridaron en una máquina de PCR programada de la siguiente manera:

Rampa a 0,5 °C/s hasta 97,5 °C

Mantener a 97,5 °C durante 150 s

10

Después una etapa de 97,5 °C durante 2 s con una caída de la temperatura de 0,1 °C/ciclo durante 775 ciclos

Reacción de ligamiento

15 Materiales:

- Adaptador bifurcado 40 µM
- ADN genómico con cola A
- Quick Ligasa NEB N.º M2200L
- Tampón Quick Ligasa 2x NEB N.º M2200L
- Máquina de PCR
- kit de purificación de PCR de columnas Qiagen (N.º 28104)

La mezcla de reacción se ensambló de la siguiente manera:

ADN	10 µl
2x tampón	25 µl
Adaptador 40 µM	10 µl
Quick Ligasa	5 µl
~50 µl	total

20

La reacción se incubó durante 20 min a temperatura ambiente, después el ADN se purificó de las enzimas, tampón, etc., cargando la mezcla de reacción en una columna Qiagen, finalmente eluyendo en 30 µl de EB.

Purificación en gel

25

Materiales:

- Agarosa Biorad N.º 161-3101
- escalera de 100 pares de bases NEB N.º N3231L
- TAE
- Tampón de carga (Tris 50 mM pH 8, EDTA 40 mM, sacarosa 40 % p/v)
- Bromuro de etidio
- Bandejas y tanque de gel. Unidad de electroforesis

30 Toda la muestra de la reacción de ligamiento purificada se cargó en un carril de un gel de agarosa al 2 % que contenía bromuro de etidio y se ejecutó a 120 V durante 50 min. Después el gel se observó en una caja de "luz blanca" y los fragmentos superiores a 300 pb a al menos 750 pb se escindieron y purificaron con un kit de purificación de un gel de Qiagen, eluyendo en 30 µl de EB. Para secciones de gel grandes se usaron dos columnas de minElute, eluyendo cada una en 15 µl de EB y se agruparon.

Amplificación por PCR

35

Materiales:

- ADN ligado
- CEBADOR 1: AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGA (SEQ ID NO: 3)
- CEBADOR 2: CAAGCAGAAGACGGCATAACGA (SEQ ID NO: 4)
- mezcla 2x Jump Start Red Taq Sigma N.º P0982 PCR
- Máquina de PCR
- Columnas MinElute de Qiagen Qiagen(N.º 28004)

40 El ADN ligado purificado se eluyó 25 veces, después se preparó una mezcla de reacción de PCR de la siguiente manera:

ES 2 732 253 T3

ADN	1 µl
Mezcla 2x Jump Start Red Taq	25 µl
Cebador 1 100 µM	0,5 µl
Cebador 2 100 µM	0,5 µl
Agua	23 µl
	~50 µl total

El termociclado se llevó a cabo en una máquina de PCR en las siguientes condiciones:

- 5 • 2 min a 94 °C
- [45 s a 94 °C, 45 s a 65 °C, 2 min a 70 °C] 16 ciclos
- 5 minutos a 70 °C
- 10 • Mantener a 4 °C

Los productos de PCR se purificaron de las enzimas, tampón, etc., en una columna MinElute de Qiagen eluyendo en 10 µl de EB. La biblioteca de ADN resultante está lista para amplificación en una plataforma de PCR de superficie.

15 Validación de bibliotecas

1 µl de la biblioteca de ADN se clonó en un vector plasmídico y se sembró en placas de agar. Se seleccionaron 19 colonias, se sometieron a miniprep y los insertos clonados se secuenciaron mediante secuenciación Sanger
20 convencional. Los datos de secuencia obtenidos fueron los siguientes:

Clon 1 (SEQ ID NO: 5)

```
TGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGTGG
GGACCGTCCTGTGCATTGTAGGGTGTCAACAGCATCCCTGACCTCCACCTACAAGATGC
CAGTAGCGAATCCCCTCAGCCCTCATCTCCTTGCCATAGTTGTGTCAACCAAATCATCT
CCACACATTGTTAGATGTTTACTGGGAGGCAGACTCACTCCCCTTGAGAACCCTGTAC
TAGAAATATACCAAGAGAATGAGATCGGAAGAGCTCGTATGCCGCTTCTGCTTG
```

25

Clon 2 (SEQ ID NO: 6)

```
AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGT
GGGCTTTGTTCTTTGAGAGGTTGCAGTCAACATGATTCTTTAAGACCAGAACCCTGCACA
CTTCTTGGGCTGTATTTCTTACATTCCTTTTCTATTTTAACCATATCCCATCTTACCTAC
TTCCAGCATAGTGGTCATATTTAATTTTTACAAAACCATTTTGCCACTTGCTGCCAACTA
TGTTCTTTATAAAGCAGACTTTGAGATGGAGGCTAGTGTTTCAGAGGGGATGCTTAGGAGA
ACTTTGGAGATTAATACTTATGGCAGGTAAGGGAAGGAAGCAGGATTAGACAGAAATATT
GAACTGTGATACAAAGTCAGCAAAGACTTTAGTCAATAGATCGGAAGAGCTCGTATGCCG
TCTTCTGCTTG
```

30 Clon 3 (SEQ ID NO: 7)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTT
CGATTCCCTTCAATGATTATTCCATTTCGAGTACATTCGATGATTCCATTTCGATTCTATAT
GATGATGATTGCATTTCGAGTCCGTGTATTATTCCATTCCATTCCATTAGATGATTCCATT
CGAGTCCATTTCGATGATTCTCTTCGATTCCGTTTCGATAATTACGCTTGATTCCGTTTGAT
GTTGATTCCATTTCGAGTCCATTCAATGTTAATTCCATTTCGATTCTAAGCGATGATTCCAT
TCCTTTCCATTAGAAGATGATTCCATTTCGAGACCATTTCGATGATTGCATTCAACTCATTC
GATGACGATTCCATTCAATTCTGTTCAATGATTCCATTAGATTCCATTTGATGATGAGTC
CATTTCGATTCCATTTGATGATGATTCCATGCGATTCCATTAGATGATGACCCCTTTCATT
TCCAAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG

Clon 4 (SEQ ID NO: 8)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTTT
AAATGCTAGGCATATTGTGTACCCACATTGGTTTGTAGCCAGCTCTATGTCATAGGGCC
5 CTTACCCTTTACCTATTTATTGTTAGTATAATGTCCATAAACAAGCCAATGGCTCAGCAT
GAACTGATGCTAAAGAAAGCTCATGCCTGAGTGATAAATTAAGTGACCTCAGCTATTTCT
CTTCAGTGTTGTGAAAGTTATTTTTAACAGTAGGTTTCTGGTAGATTCTCTAACCACTC
GGTATTTACATGGCCCAACTTGGTTAACTCGACTGGTTACGGCAAATGCTGAAGATCGG
AAGAGCTCGTATGCCGTCTTCTGCTTG

Clon 5 (SEQ ID NO: 9)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGAAGCTCTTCCGATCTTA
AGGAAGTTAGGTAGATAATTTTTGTTTAGGCCATATAGCTTTGATTTTCTGATAACAATT
TTATAAACTTAGAAATTTTCATGTAAGATACAGGAATACTGGAAGCAAAAAAAGAAGGT
GCTTTAACCTTAGGGATTGAAAAAATAGTAATTTAGGTTGAAAATGCTGCTTGAAAGTTA
ATGCTGATAGCATTACTACACATGATGATTTTTTCTGGAAGGAAAGCTTTATCTGGGCCT
TCAATTTAGGAATTTTTCTCTTTGGTTTTTAAAAGCTGCCATATTCATTGAGCTTCATG
GGAAAGATGCAAATAACTAAAACAAATGAACAAAACCATGTTGAGGTCAGGAACTTATT
TCAAGAAAGCAAGTTCTAGGTTTTCTTTTAAAGTGACAGTAGAGCCTTAGGCCTCAAACC
10 ATCTACAACCATGTTAACAGTAAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG

Clon 6 (SEQ ID NO: 10)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCAGATCTTC
CTGCCTCAGCCTCCCGAGTAGTTGGGATTGCAGGCATGTGCCACCATGCCCTGCTAATTT
TTGTATTTTTAACTAAAGGAGGGGTTTTGCCATGTTGGCCAGGCTGGTCTTGAATGCCTG
ACCTCAGGTGATCCGCCACCTCAGCCTCCCAGTGCTGGGATTACAGGTGTGAGCCACTG
CGCCAGCTGAGGGTAACTATTTTTAATGTGGCTGATGAATGTAACCTATCCTGTCCCATG
TCTCTGTCCCAGCTGCAGAGCCCTCGTCGAGATCGGAAGAGCTCGTATGCCGTCTTCTG
CTTG

Clon 7 (SEQ ID NO: 11)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTAA
GGTTTAAATTTTCTATATAGCCTGAAAAGTTTGAATGTTTAATTCAAACCTAATTTATTGA
GCAATGGCATTTAAGAAAATGGAAAGATACAAAGGGACTTTCATCAGATGATAAGTGGAT
AAGAGAGAAAATGCAGACAGATGAGCCAGAGTTGTGTAAGCTGGGAGGCTAGCAGGG
CCTTGTAGATAGCCAAGCTGACTGGGGAACAGAGATAAATGGGAGCAGAGATCAGAAAGT
5 TCATCCTTACCCTGCTGCCCGTGGTGAAAGGAGACTTGCAAGATCGGAAGAGCTCGTATG
CCGTCTTCTGCTTG

Clon 8 (SEQ ID NO: 12)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTAA
GGAATCTTTATTTTCTACATTTGAGTTTGGAAACTGAGCTAGCACATCTAAATCCATCT
AATTTTGGTCATTGGTTTTTAACAAGTTCATCTTATTTTTTTAAACATCTGATCTTTATTT
TATAGAATAGACTACACAAAGTCTTTTGGAAAATTAAAATATTTTAACTTCCAACAATTT
TCAGATTTTACTTATAAAAAAATTTAAAATCCTCTACTTTACTCGCATCTTTATTATTTT
TGACTTTCTAGCTACTTAAAGTTAAGGAGGAAATTAACCTCTCTAAGATCGGAAGAGCTC
GTATGCCGTCTTCTGCTTG

Clon 9 (SEQ ID NO: 13)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTCG
CCATAACCACAGCCCAGGCCTCCGTAGCCACAGCACCATAGCCATGGCCACCATTGTAG
TTTCTGTAGTAGCTGCCACACATGGTGTGGTTGTTGAGGTCATCCTTGGGTAGGAAGGA
GTGTAGGTGACTTCAGTATGGACACTTCTCCGCAGAGGGCCTTTTATATGCCTCAGTGAA
TCAGAACATAGCGTGCCCTGCAAAAATATCTCTAAAGGCCTTTCATTGTGCTGAGAAGT
TCTGGCCCTTACGTATCTCTCTGATTTTCATATCCTGCTACTCTCCTCCATTTATCTATAA
TGCTCAAACCTCTGCTGGCTTTTTGTCTTTTAAAATGCAGCAGGTTTCTTCTCACAATAAG
GAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG

Clon 10 (SEQ ID NO: 14)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTAG
AGAGAAGTTATTTAAGACAAGTAAGGTATCAGGTGTGACTCAAATACCACAGAATCTAG
CTATTGTTAGCAATATTAAGTATATTTTCTTAAATGAAGGATTCTCCATTTACCATATGC
CCCTTGATAATTTCCAGAGAATTTAATTTTTTAAAAGGAATTTTCACCAATTAATTTAT
TGTTTTGATCAAAGAGGACCCACTGAACACCTTATTCATTATTAATGTATCATAAAA
CTTAATTATGGAGCTGGGTACAGGGGCTCATGCCTGTAATCCCAGCACTTTGGGAGGCTG
AGGCAGGAGGACTGCTTGAGTCCAGGAGTTTGAGACCAGCCTGGGTAACATGGTGAACC
CTGTCTCTACAAAAAATACAAAAAATTAGCCAGGTGTGGTGAGATCGGAAGAGCTCGTAT
GCCGTCTTCTGCTTG

5

Clon 11 (SEQ ID NO: 15)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTCC
CAGGGAAGCCAAAAGATTGGACACCCCTCTTCAAACATAAATTCCTCCCATAGTTAGT
TTGGCCTATGCTGGGAAATGAACAAGGGTGGCTTTGAGGTTAGAAGCAAATGGAGTCAG
TTAGGTCAGACTTTTTTCACTATCATACTTTTTCTATGTCAGATTTATCTCACTTGTA
TTTTTGCAAGGGTGGTTTCAGAGCCACTAAGCTTGTGGTAATTTTTTACTGCAATAGAAA
ACTAATGCATTAGGTAACCCTCTTTTTTCCCTCTGATTGCTTGCTCTGGGGTAAGCCAG
CTGCCATGTTGAATTTTTCATTTTTGTTACTGAGCAATCAATGGGCTCACTGCCCGACGTG
CATAGAGGCCAATACTGTGGCACTAGTTTTTGAGAAAAGATCGGAAGAGCTCGTATGCCG
TCTTCTGCTTG

10

Clon 12 (SEQ ID NO: 16)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTTT
TCAGATTTTATATGTATTAACCTCAGAACACACACCTCTTATCACACATATTTTTTCATGT
AATTTATCTAAATCTTATAGAAAAGGGTCCATTTGCATTTTCTCTTATTAGACTCCTGAT
TTCAAATAATATATTAATGAGTATTTTTCTGTGCTGTAGTTATTCATTCTTATAGAT
ATGTAACATAATTCCTTTGCAAAGGTAAAATTGAGCTATCTCTTGTTGAGGATTTGTT
GATCTCTGTCTAAAGTTTCAAAAATAAAGAACTTTAAAAGCAAATGTAAATTCCTTTCA
AGTTTTAGTAAATTAATCAAACTTAGTAGCTTAAACAATACAGATTTATTATGTTACA
GTTCTGTAAGACAGAAATCTGACTTGATCACACCATGGTAAAACCAAGATACTGCCAGGG
TTGGTTTTTTCTTGGGGGGGGTCTGTGGGAAGAGTTGTTTCCTTTGGTTTTCCACAGCC
CAGAGGCTGCTTGCAATTCCTTTGATCACTAGATCGGAAGAGCTCGTATGCCGTCTTCTGC
TTG

Clon 13 (SEQ ID NO: 17)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCTTA
TTGTTTAGGGAATAATGACAAGGAAAAAAGTCTGTAGATATTCAGTACAGAGGCACCCA
TCTTTTTAAATTTCTGAAGATTTTTTACTCATGCTTGGTTGAATCCACAGATGCAGAACC
CATAGGTTTCAGAGGGCCAGCTGTGCTTTGAAAATATTAGCTTGTGTTTTTATTAGAAAGA
AAACTCTGAGGCCAGGCACGGTGGCTCACGCCTGTAATCCCAGCACTTTGGGAGGCTGAG
GTGGGCGGATCACAAGGTGAGGAGATCGAGACCATTCTGGCTAACATGGTGAAACCCTGT
CTCTACTAAAAATACAAAAAATTAGCCGGGCGTGGTAGTGAGCACCTAGATCGGAAGAG
5 CTCGTATGCCGTCTTCTGCTTG

Clon 14 (SEQ ID NO: 18)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCTGC
AATTGGTAAAACAGTAAGCAATGAAACAGACACTTCTCAAATATTTCCAAGATGGTACACG
CTTTTCAGTGTGTATGATCCAATAAAGCCATTGGAAGTAGGCTTTAATAGTCAAAAAAGA
CTATTCAGTTAGATAGGA ACTATTTGCCTATAACTATTGGCCAAAAATAGGTTAAAAAAT
TGTTTTAAATTTGTGCTTTACAAAACATGTGGACTTTTTTTAGAAAATGTGTCAAATTTCA
AAAGAAATATAGACATTATGGAAAGGTCAGTTAAGCACAGCCCTAATCCTGAAAACATAA
CTATGAAAGATACTAGCTGT TACTTGTAAACAAAAGGAAAAAAAAGATATTAGTAACCAA
TAATTAGCAAACAATGCCCATATATTTCCTTTTTTTTTTTTTTTTTTTTTTGAGACAGGGGCT
TACTCAGGCTGGATGTGATCATGAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG

Clon 15 (SEQ ID NO: 19)

AATGATACGGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTCCGATCTTT
AGAAGCTCTATTATACTGGAAAAGAGATATGAGACCCTTCCTACTTTAAGAATCAATGAA
GCCGGGTGTGGTGGCTCACGCCTGTACTCCCAGCATTTTGAGAGGCCAAGCTGGGCAGAT
CACCTGAGGTCGGGAGTTTGAGACCAGCCTAGCCAACATGATGAAATCCTGTTTCTACTA
ATAACACAAAAATTAGCCGGGTGTGGTGGCGCACATCTGGAATCCCAGCTACTCCAGAGG
CTGAGGCAGGAAAATTGCTTGAAGCTGGGAAGCAGAAGTTGCAGTGAAGATCGGAAGAGC
TCGTATGCCGTCTTCTGCTTG

Clon 16 (SEQ ID NO: 20)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGG
CTGGACTGAATAGGATAGCCTTAGCTGTAAAATTGGGCTGATCTTCAAATGGACTCATG
CTTGCCGAATGACTCACGCTCCTGTTTACAAATCAGCTCTGTGAAGAAATGCAGAGTGGG
AGGCTCTGCTTGCCAGACGGAGACCTTAGACCTCCAGGGGCGGAGAACGGAGTACTTCCT
CTGGTGCTCGGCTTCCCTTCCTGGGGGCAGATCTCTCAGCTTCTGGTTGGTGGCTCTCAA
AATCCAGACACAAGGTCAGCTGCAGCCAGCGTGGGCCCTGGAGTAGCTCCAGTTATGGGG
CAGCAATGGCCCCCTCTCATTTTGTAGAGCTCACTTTGCCTGTGGATGGTTTTAATCCATC
TGGATAAACTTGAGGCCCATGGGAATACCATATACTATGGTAACCATGTACACTGCTCTA
AAGATGTGGCTGCTGTTGTATAACTTTTTCTTTATTTTTGTCAATTTCTATTTTCCAG
AGTCTTGCATACCCACTATGTCTACTGTGATAGTGAACGTAAAAACATAACAAGATGTTGG
TGTTATCCTCAATCTCAGATCGGAAGAGCTCGTATGCCGTCTTCTGCTTG

Clon 17 (SEQ ID NO: 21)

5

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTCT
GTGAGAAATCAATGTCTGCTGTTTATAAGCCGCGGGCTGTGATATCCTGTGAGAGTGGC
CCCAGTGGATGAAGACAGATGCTCTCAAGGAGCGCAGATGACGCGGGTTCCGAAGGACTC
GGCACCAGCCCGGAGGCCGGCAACATGGGCAAGGGCCTCTCACGGCTGACCTGTTTCC
TCATCAGCACATCAGGACAATAAGAGCTCCCACCTCACAGGTGGTGAAGAGCCAACGTGG
TGAAGAATGAATAAAGCAGCTCGTGGAAAGTGCTGTGCATGAGGCCTGGCAACCGGTCCC
TGCTCTGAGGTCACCTGCCACGGAGCTGCTGACAGGACCATTAAAAACACAATTGTGCAA
GTGCTCACCCACATTCACAGCAGCAGAATCTCCACCAGCCAAGCATGGAGACGATCCTT
GCATCCATAGACATGAACAGATGAGCAAACGTGGTCTATACGGACGATGAAATAGCACT
CAGCCCTAAGAAGAAATAAAATCCCGACAGAGAGATCGGAAGAGCTCGTATGCCGTCTTC
TGCTT

Clon 18 (SEQ ID NO: 22)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGC
TGAGAGGAGAGGAGAGGCTGGCAGGGGGCTGATGCAGGAGGTGATAGGGCTCCCGGTGAT
AAGAGGTGAGAAGAACAGTCTCTGTGTGCCTAGAGAAGAGATTACCAGAAGTCTGCTATC
TGTTTGTTCGCGGATGTCGGACAGGCAGGATCGGTGATGGCAGGTCTTGGGGGAAGGATT
ATCAGGAGCTAAAAGCTGTCTTACCTTGGCTGCTAAGAACTCATCTCGGATCTTCTTAG
AATCCAAATCGGACTTTTCTCCTAGCAGTGGCTACATCCTTAACCTCAAAAATACCCGT
ATTAGCAGATCTACCTCCATGAAATAGACAATTCTTGACAAACTAAGATCGGAAGAGCTC
GTATGCCGTCTTCTGCTTG

10

Clon 19 (SEQ ID NO: 23)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCTGT
TATTTTTTCATACTGATGATTCTAGGGCGGTTCCCTGGACTTGGCTGGGACCAGAACTGTC
TCACAGCTGCCAGCGCTCCCCACCCCATGAGAACCATCGTGCTGAAGATGAAGAGCCAA
AGCCAAACTGACCAGCCTCTGCAACCCCAAGCATCACCGTGTGAGGACAGCGTGATGTGG
AGCCGCGGAGACCCACGCAGTGCTGCGAGGGGCAACCACATGCAGGGGGGCAGAGGTGT
GGGGGAGCAAGCAGATGCCACCGGCATGCCCTAGGAGCCACAGGGAGCATGCGGGCTGGGC
AGTGATGGGAATGAACGTGAACTCCAGTCTGCCCCAAGAAGTCTCCGGCCCCCTCCTTC
TCAATTCAGGGCACAAAGTGGTAACTGCAGCATGCAGAGGAGTCGAGGAGTCTTCCCTCG
TCCCAGGAGCAGCACCTCGGGCACAGTCTCGGTCCCACAGAACAGCCAAGTGTGGGTTGG
TGGTCTAGAGACCTCCGAAGATCCAGTGGGGGAAGGATGGGCAGCAGAGGGTCTACTCTC
TGAAAATAAGGGGAAGGGATTTTCCCTCCCCACTGCCAAGGTCCCAGCTACTGGACGTGG
GTGGAGATCGGAAGAGCTCGTATGCCGTCTTCTGCCTT

5

Estos resultados confirman que el método de preparación de bibliotecas produce una biblioteca de plantillas de ADN "secuenciables" que contienen una mezcla de fragmentos de diferente secuencia. Se descubrió que el ADN inserto de cada uno de los 19 clones secuenciados se alineaba con una secuencia de referencia del genoma humano (alineamiento no mostrado), ilustrando que el método produce una biblioteca de plantillas que refleja realmente la composición de secuencia de los fragmentos diana de partida (es decir, los clones contienen fragmentos genómicos humanos en lugar de una secuencia "redundante").

10

Realizaciones preferidas se describen en los siguientes párrafos:

15

1. Un método de generación de una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3', comprendiendo el método las etapas que consisten en:

20

ligar polinucleótidos adaptadores con apareamiento erróneo idénticos en los dos extremos de cada uno de uno o más dúplex de polinucleótido diana para formar una o más construcciones adaptador-diana, en el que cada adaptador con apareamiento erróneo se forma a partir de dos cadenas polinucleotídicas hibridadas que forman un complejo biomolecular que comprende al menos una región bicatenaria y una región no apareada, y

25

realizar una reacción inicial de extensión de cebador en la que un oligonucleótido cebador hibrida con una parte de adaptador de cada una de las construcciones adaptador-diana y se extiende por adición secuencial de nucleótidos para formar productos de extensión complementarios a al menos una cadena de cada una de las construcciones adaptador-diana,

30

en el que los productos de extensión, y opcionalmente productos de amplificación derivados de los mismos, proporcionan en su conjunto una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3';

35

2. Un método según el párrafo 1 en el que la reacción inicial de extensión de cebador comprende hibridar un oligonucleótido cebador con una parte de adaptador de cada cadena de cada una de las construcciones adaptador-diana y extender los cebadores para formar productos de extensión complementarios a cada cadena de cada una de las construcciones adaptador-diana.

40

3. Un método según el párrafo 1 o el párrafo 2, en el que la reacción de extensión inicial del cebador comprende las etapas que consisten en:

hibridar un primer oligonucleótido cebador con una parte de adaptador de cada una de las construcciones, b) extender los cebadores por adición secuencial de nucleótidos para formar productos de extensión complementarios a al menos una cadena de cada una de las construcciones adaptador-diana, y c) someter los productos obtenidos en la etapa b) a condiciones desnaturalizantes, separando así los productos de extensión de cadenas de las construcciones adaptador-diana.

4. Un método según el párrafo 3, en el que las etapas a) a c) se llevan a cabo una vez y los productos de extensión se recogen para proporcionar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y en sus extremos 3'.
- 5 5. Un método según el párrafo 3, en el que las etapas a) a c) se repiten una o más veces y toda la extensión se recogió para proporcionar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y en sus extremos 3'.
- 10 6. Un método según uno cualquiera de los párrafos 1 a 5, en el que los cebadores usados en la reacción inicial de extensión con cebador hibridan en una región no apareada de los adaptadores.
- 15 7. Un método según uno cualquiera de los párrafos 1 a 3, en el que la reacción de extensión inicial con cebador se lleva a cabo como parte de una reacción en cadena de la polimerasa y los productos de amplificación de la reacción de PCR se recogen para proporcionar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y 3'.
- 20 8. Un método según el párrafo 7, en el que la reacción en cadena de la polimerasa se realiza usando primeros cebadores de oligonucleótido que son capaces de hibridar con partes de adaptador de cada una de las construcciones adaptador-diana y segundos cebadores de oligonucleótido que son capaces de hibridar con una región de las cadenas extendidas producidas por extensión de los primeros cebadores de oligonucleótido, siendo esta región complementaria a una parte de adaptador de las construcciones adaptador-diana.
- 25 9. Un método según el párrafo 8, en el que los primer y segundos cebadores de oligonucleótido tienen diferentes secuencias de nucleótido.
- 30 10. Un método según el párrafo 9, en el que los primeros cebadores de oligonucleótido son capaces de hibridar con la región no apareada en la primera cadena de las construcciones adaptador-diana y los segundos cebadores de oligonucleótido son capaces de hibridar con un producto de extensión complementario a una región no apareada en una segunda cadena de las construcciones adaptador-diana.
- 35 11. Un método según el párrafo 8, en el que los primer y segundos cebadores de oligonucleótido tienen secuencias de nucleótidos esencialmente idénticas.
- 40 12. Un método según el párrafo 11, en el que los primer y segundos cebadores de oligonucleótido son capaces de hibridar con una región bicatenaria en las partes de adaptador de las construcciones de adaptador-diana.
- 45 13. Un método según uno cualquiera de los párrafos 1 a 12, en el que los polinucleótidos adaptadores con apareamiento erróneo son adaptadores bifurcados formados por hibridación de primeras y segundas cadenas polinucleotídicas parcialmente complementarias, en el que una secuencia de 5 o más nucleótidos consecutivos en el extremo 3' de la primera cadena es complementaria a una secuencia de 5 o más nucleótidos consecutivos en el extremo 5' de la segunda cadena de tal manera que una región bicatenaria de 5 o más pares de bases consecutivos se forma hibridando las dos cadenas y en el que una secuencia de al menos 10 nucleótidos consecutivos en el extremo 5' de la primera cadena y una secuencia de al menos 10 nucleótidos consecutivos en el extremo 3' de la segunda cadena no son complementarias de tal manera que la región no apareada de al menos 10 nucleótidos consecutivos en cada cadena permanece como una forma monocatenaria cuando hibrida la región bicatenaria.
- 50 14. Un método según el párrafo 13, en el que la región bicatenaria formada cuando las dos cadenas hibridan tiene una longitud de 5 a 20 pares de bases consecutivos.
- 55 15. Un método según el párrafo 13 o el párrafo 14, en el que la región no apareada consiste en 10 a 50 nucleótidos no apareados consecutivos en cada cadena.
- 60 16. Un método según uno cualquiera de los párrafos 1 a 15, en el que los dúplex de polinucleótido diana son fragmentos de ADN genómico.
- 65 17. Un método según el párrafo 16, en el que los dúplex de polinucleótido diana son fragmentos de un genoma completo.
18. Un método según uno cualquiera de los párrafos 1 a 15, en el que los dúplex de polinucleótido diana son ADNc bicatenarios.
19. Un método de amplificación de ácidos nucleicos en fase sólida de moléculas polinucleotídicas plantilla que comprende las etapas que consisten en:
preparar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos

5' y secuencias comunes en sus extremos 3' usando el método según uno cualquiera de los párrafos 1 a 18, y realizar una reacción de amplificación de ácido nucleico en fase sólida en el que se amplifican dichas moléculas polinucleotídicas plantilla.

- 5 20. Un método según el párrafo 19, en el que dichas moléculas polinucleotídicas plantilla se amplifican por PCR en fase sólida usando cebadores de amplificación directos e inversos.
21. Un método según el párrafo 20, en el que dicha biblioteca de moléculas polinucleotídicas plantilla se prepara usando el método de uno cualquiera de los párrafos 7 a 12 y una PCR en fase sólida se lleva a cabo usando
10 cebadores idénticos a los usados en la preparación de la biblioteca de moléculas polinucleotídicas plantilla.
22. Un método según el párrafo 20, en el que dicha biblioteca de moléculas polinucleotídicas plantilla se prepara usando el método de uno cualquiera de los párrafos 7 a 12 y una PCR en fase sólida se lleva a cabo usando
15 cebadores que no son idénticos a los usados en la preparación de la biblioteca de moléculas polinucleotídicas plantilla.
23. Un método según el párrafo 20, en el que dicha biblioteca de moléculas polinucleotídicas plantilla se prepara usando el método de uno cualquiera de los párrafos 1 a 6 y una PCR en fase sólida se lleva a cabo usando
20 cebadores de amplificación directos capaces de hibridar con una parte de adaptador de los productos de extensión formados en la reacción inicial de extensión con cebador y cebadores inversos que son capaces de hibridar con productos de extensión formados por la extensión de cebadores de amplificación directos hibridados con los productos de extensión formados en la reacción inicial de extensión con cebador.
24. Uso de una biblioteca de moléculas polinucleotídicas plantilla preparada según el método de uno cualquiera
25 de los párrafos 1 a 18 como un plantilla para amplificación por PCR en fase sólida.
25. Uso de una biblioteca de moléculas polinucleotídicas plantilla preparada según el método del párrafo 17 como un plantilla para amplificación del genoma completo.
- 30 26. Un método de amplificación del genoma completo que comprende las etapas que consisten en: preparar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y en sus extremos 3' usando el método según el párrafo 17, y llevar a cabo una reacción de amplificación de ácido nucleico en la que se amplifican dichas moléculas polinucleotídicas plantilla.
- 35 27. Un kit para su uso en la preparación de una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y en sus extremos 3', comprendiendo el kit polinucleótidos adaptadores con apareamiento erróneo como se define en uno cualquiera de los párrafos 1 a 18 y cebadores de oligonucleótido capaces de hibridar con los polinucleótidos adaptadores con apareamiento erróneo.
- 40 28. Un kit según el párrafo 27, en el que los cebadores de oligonucleótido son capaces de hibridar con una región no apareada de los polinucleótidos adaptadores con apareamiento erróneo.
29. Un kit según el párrafo 28, en el que los polinucleótidos adaptadores con apareamiento erróneo son
45 adaptadores bifurcados como se define en uno cualquiera de los párrafos 13 a 15.

LISTADO DE SECUENCIAS

- <110> Solexa Limited
- 50 <120> Método de preparación de bibliotecas de polinucleótidos plantilla
- <130> P81137WO00
- <160> 23
- 55 <170> FastSEQ para Windows Versión 4.0
- <210> 1
- <211> 33
- 60 <212> ADN
- <213> Secuencia artificial
- <220>
- <221> misc_feature
- 65 <222> oligonucleótido sintético

ES 2 732 253 T3

<400> 1
acacttttc cctacacgac gctctccga tct 33

5 <210> 2
<211> 33
<212> ADN
<213> Secuencia artificial
<220>
10 <221> misc_feature
<222> oligonucleótido sintético

<400> 2
gatcggaaga gctcgtatgc cgtcttctgc ttg 33

15 <210> 3
<211> 44
<212> ADN
<213> Secuencia artificial

20 <220>
<221> misc_feature
<222> oligonucleótido sintético

25 <400> 3
aatgatacgg cgaccaccga gatctacact ctttcctac acga 44

30 <210> 4
<211> 21
<212> ADN
<213> Secuencia artificial

35 <220>
<221> misc_feature
<222> oligonucleótido sintético

<400> 4
caagcagaag acggcatacg a 21

40 <210> 5
<211> 296
<212> ADN
<213> *Homo sapiens*

45 <400> 5

tgatacggcg accaccgaga tctacactct ttcctacac gacgctcttc cgatctgtgg 60
ggaccgtcct gtgcattgta ggggtgtcaa cagcatccct gaacctccacc tacaagatgc 120
cagtagcgaa tcccctcagc cctcatctcc ttgccatagt tgtgtcaacc aaaatcatct 180
ccacacattg ttagatgttt actgggaggc agactcactc ccacttgaga accactgtac 240
tagaaatata accaagagaa tgagatcgga agagctcgta tgccgtcttc tgcttg 296

50 <210> 6
<211> 431
<212> ADN
<213> *Homo sapiens*

55 <400> 6

ES 2 732 253 T3

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctgt 60
gggctttggt ctttgagagg ttgcagtcaa catgattctt taagaccaga accctgcaca 120
cttcttgggc tgtatttctt acattccctt tctattttaa ccataatcca tcttacctac 180
ttccagcata gtggtcatat ttaattttta caaaaccatt ttgccacttg ctgccaacta 240
tgttctttat aaagcagact ttgagatgga ggctagtgtt cagaggggat gcttaggaga 300
actttggaga ttaatactta tggcaggtaa gggaaggaag caggattaga cagaaatatt 360
gaactgtgat acaaagtcag caaagacttt agtcaataga tcggaagagc tcgtatgccg 420
tcttctgctt g 431

```

5 <210> 7
 <211> 518
 <212> ADN
 <213> *Homo sapiens*
 <400> 7

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatcttt 60
cgattccctt caatgattat tccattcgag tacattcgat gattccattc gattctatat 120
gatgatgatt gcattcgagt ccgtgtatta ttccattcca ttccattaga tgattccatt 180
cgagtccatt cgatgattct cttcgattcc gttcgataat tacgcttgat tccgcttgat 240
gttgattcca ttcgagtcca ttcaatgtta attccattcg attctaagcg atgattccat 300
tcctttccat tagaagatga ttccattcga gaccattcga tgattgcatt caactcattc 360
gatgacgatt ccattcaatt ctgttcaatg attccattag attccatttg atgatgagtc 420
cattcgattc catttgatga tgattccatg cgattccatt agatgatgac ccctttcatt 480
tccaagatcg gaagagctcg tatgccgtct tctgcttg 518

```

10
 15 <210> 8
 <211> 387
 <212> ADN
 <213> *Homo sapiens*
 <400> 8

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatcttt 60
aaatgctagg catattgtgt accccacatt ggttttagc cagctctatg tcatagggcc 120
cttacccttt acctatttat tgttagtata atgtccataa acaagccaat ggctcagcat 180
gaactgatgc taaagaaagc tcatgcctga gtgataaatt aagtgacctc agctatttct 240
cttcagtgtt gtgaaagtta tttttaacag taggtttcct ggtagattct ctaaccactc 300
ggtatttcac atggccaac ttggttaact cgactggtta cggcaaatgc tgaagatcgg 360
aagagctcgt atgccgtctt ctgcttg 387

```

20
 25 <210> 9
 <211> 536
 <212> ADN
 <213> *Homo sapiens*
 <400> 9

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgaagctct tccgatctta 60
aggaagttag gtagataatt tttgtttagg ccatataget ttgattttct gataacaatt 120
ttataaactt agaaattttc atgtaagata caggaatact ggaagcaaaa aaaagaaggt 180
gctttaaact tagggattga aaaaatagta atttaggttg aaaatgctgc ttgaaagtta 240
atgctgatag cactactaca catgatgatt ttttctgaa ggaaagcttt atctgggcct 300
tcaatttagg aatttttctc tttggttttt aaaagctgcc atattcactt gagcttcatg 360
ggaaagatgc aaataactaa aacaaatgaa caaaaaccat gttgagggtca ggaacttatt 420
tcaagaaagc aagttctagg ttttctttta aagtgacagt agagccttag gcctcaaacc 480
atctacaacc atgttaacag taagatcgga agagctcgta tgccgtcttc tgcttg 536

```

30 <210> 10
 <211> 364
 <212> ADN

ES 2 732 253 T3

<213> *Homo sapiens*

<400> 10

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tcagatcttc 60
ctgcctcagc ctcccagagta gttgggattg caggcatgtg ccaccatgcc ctgctaattt 120
ttgtatTTTT aactaaagga ggggttttgc catgttggcc aggctggctc tgaatgcctg 180
acctcagggtg atccgcccac ctacgcctcc cagtgtggg attacagggtg tgagccactg 240
cgcccagctg agggtaacta tttttaatgt ggctgatgaa tgtaactatc ctgtcccctg 300
tctctgtccc cagctgcaga gccctcgtcg agatcgggag agctcgtatg ccgtcttctg 360
cttg 364

```

5

<210> 11

<211> 374

<212> ADN

10 <213> *Homo sapiens*

<400> 11

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctaa 60
ggtttaaatt ttctatatag cctgaaaagt ttgaatgttt aattcaaact aatttattga 120
gcaatggcat ttaagaaaat gaaagatac aaagggactt tcacagatg ataagtggat 180
aagagagaaa aatgcagaca gatgagccag agttgtgtaa aagctgggag gctagcaggg 240
ccttgtagat agccaagctg actggggaac agagataaat gggagcagag atcagaaagt 300
tcacacctac cctgctgcc gtggtgaaag gagacttgca agatcgggag agctcgtatg 360
ccgtcttctg cttg 374

```

15

<210> 12

<211> 379

<212> ADN

20 <213> *Homo sapiens*

<400> 12

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctaa 60
ggaatcttta ttttctacat ttgagtttgg aaaactgagc tagcacatct aaatccatct 120
aattttggtc attggtttta acaagttcat cttatTTTT taaacatctg atctttattt 180
tatagaatag actacacaaa gtcttttggg aaattaaaat attttaactt ccaacaattt 240
tcagatttta cttataaaaa aattttaaact cctctacttt actcgcctct ttattatttc 300
tgactttcta gctacttaa gttaaggagg aaattaacct ctctaagatc ggaagagctc 360
gtatgccctc ttctgcttg 379

```

25

<210> 13

<211> 455

<212> ADN

<213> *Homo sapiens*

30

<400> 13

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctcg 60
ccataaccac agcccaggcc tccgtagcca cagcaccat agccatggcc accattgtag 120
tttctgtagt agctgccaca catgggtgtg gttgttgagg tcaccttgg gtaggaagga 180
gtgtagggtga cttcagtatg gacacttctc cgcagagggc cttttatatg cctcagtgaa 240
tcagaacata gcgtgccctt gcaaaaatat ctctaaaggc ctttcattgt gctgagaagt 300
tctggccctt acgtatctct ctgatttcat atcctgctac tctcctccat ttatctataa 360
tgctcaaact ctgctggctt tttgtctttt aaaatgcage aggtttcttc tcacaataag 420
gagatcggaa gagctcgtat gccgtcttct gcttg 455

```

35

<210> 14

<211> 495

<212> ADN

<213> *Homo sapiens*

ES 2 732 253 T3

<400> 14

```

aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctag 60
agagaagtta ttaagacaa gtaaggatc aggttgtag tcaaatacca cagaatctat 120
ctattgttag caatattaag tatattttct taaatgaagg attctccatt taccatatgc 180
cccttgata atttcoagag aatttaattt tttaaaagga attttacca attaaattat 240
tgtttgatc aaaagaggac ccactgaaca cttattcat tattaaaatg tatcataaaa 300
cttaattatg gagctgggta caggggctca tgectgtaat cccagcactt tgggaggctg 360
aggcaggagg actgcttgag tccaggagtt tgagaccagc ctgggtaaca tggtgaaacc 420
ctgtctctac aaaaaataca aaaaattagc caggtgtggt gagatcgga gagctcgtat 480
gccgtcttct gcttg 495

```

5 <210> 15
 <211> 491
 <212> ADN
 <213> *Homo sapiens*

10 <400> 15

```

aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatctcc 60
cagggaaaggc aaaagattgg acaccctct tcaactat aaattcctcc catagttagt 120
ttggcctatg ctgggaaatg aacaaggggtg gctttgaggt tagaagcaa atggagtcag 180
ttaggtcaga ctttttttca ctatcatact tttctatgt cagatttctc tcacttgtaa 240
tttttgcaag ggtggtttca gagccactaa gcttggtgta attttttact gcaatagaaa 300
actaatgcat taggtaacc ctttttttcc cctctgattg cttgctctgg ggtaagccag 360
ctgccatgtt gaatttttca ttttgttact gagcaatcaa tgggctcact gcccgacgtg 420
catagaggcc aatactgtgg cactagtttt tgagaaaaga tcggaagagc tcgtatgccg 480
tcttctgctt g 491

```

15 <210> 16
 <211> 603
 <212> ADN
 <213> *Homo sapiens*

20 <400> 16

```

aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttt 60
tcagatctta tatgtattaa ctcagaacac acaccttta tcacacatat tttttcatgt 120
aatttatcta aatcttatag aaaaggtcc atttgcattt tctcttatta gactcctgat 180
ttcaaataat atattactta tgagtatttt tctgtgctgt agttattcat tcttatagat 240
atgtaacata attccttttg caaaggtaaa aattgagcta tctcttggtg aggatttgtt 300
gatctctgtc taaagtttca aaaataaaga actttaaaag caaaatgtaa attccttca 360
agtttttagta aaattacttc aaacttagta gcttaaacaa tacagattta ttatgttaca 420
gttctgtaag acagaaatct gacttgatca caccatggta aaaccaagat actgccaggg 480
ttggtttttt cttggggggg gtctgtggga agagtttgtt tcctttgggt ttccacagcc 540
cagaggctgc ttgcattcct ttgatcacta gatcggaaga gctcgtatgc cgtcttctgc 600
ttg 603

```

25 <210> 17
 <211> 442
 <212> ADN
 <213> *Homo sapiens*

<400> 17

ES 2 732 253 T3

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctta 60
ttgttttaggg aataatgaca agggaaaaaa gtctgtagat attcagtaca gaggcaccga 120
tctttttaaaa tttctgaaga ttttttactc atgcttggtt gaatccacag atgcagaacc 180
cataggttca gagggccagc tgtgctttga aatatattagc ttgtgttttt attagaaaga 240
aaactctgag gccaggcagc gtggctcagc cctgtaatcc cagcactttg ggaggctgag 300
gtgggcggat cacaaggtga ggagatcgag accattctgg ctaacatggt gaaacctgt 360
ctctactaaa aatacaaaaa aattagccgg gcgtggtagt gagcacctag atcggaagag 420
ctcgtatgcc gtcttctgct tg 442

```

<210> 18
 <211> 537
 <212> ADN
 <213> *Homo sapiens*

<400> 18

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctgc 60
aattggtaaa acagtaagca atgaaacaga cacttctcaa atattccaag atggtaacag 120
cttttcagtg tgtatgatcc aataaagcca ttggaagtag gctttaatag tcaaaaaaga 180
ctattcagtt agataggaac tatttgcta taactattgg ccaaaaatag gttaaaaaat 240
tgttttaaat ttgtgcttta caaaacatgt ggactttttt agaaaatgtg tcaaatttca 300
aaagaaatat agacattatg gaaaggtcag ttaagcacag ccctaactct gaaaacataa 360
ctatgaaaga tactagctgt tacttgtaac caaaaggaaa aaaaagatat tagtaaccaa 420
taattagcaa acaatgccca tatatttctt tttttttttt ttttttttga gacaggggct 480
tactcaggct ggatgtgato atgagatcgg aagagctcgt atgccgtctt ctgcttg 537

```

<210> 19
 <211> 381
 <212> ADN
 <213> *Homo sapiens*

<400> 19

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatcttt 60
agaagctcta ttatactgga aaagagatat gagacccttc ctactttaag aatcaatgaa 120
gccgggtgtg gtggctcagc cctgtactcc cagcattttg agaggccaag ctgggcagat 180
cacctgaggt cgggagtttg agaccagcct agccaacatg atgaaatcct gtttctacta 240
ataacacaaa aattagccgg gtgtgggtggc gcacatctgg aatcccagct actccagagg 300
ctgaggcagg aaaattgctt gaagctggga agcagaagtt gcagtgaaga tccgaagagc 360
tcgtatgccg tcttctgctt g 381

```

<210> 20
 <211> 650
 <212> ADN
 <213> *Homo sapiens*

<400> 20

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctgg 60
ctggactgaa taggatagcc ttagctgtaa aattgggctg atctttcaaa tggactcatg 120
cttgccgaat gactcacgct cctgtttaca aatcagctct gtgaagaaat gcagagtggg 180
aggetctgct tgccagacgg agaccctaga cctccagggg cggagaacgg agtacttctt 240
ctgggtgctg gcttcccttc ctggggcagc atctctcagc ttctggttgg tggctctcaa 300
aatccagaca caaggtcagc tgcagccagc gtggggccctg gagtagctcc agttatgggg 360
cagcaatggc cccctctcat ttgagagct cactttgctt gtggatggtt ttaatccatc 420
tgataaaact tgaggcccat gggaatacca tatactatgg taacctatga cactgctcta 480
aagatgtggc tgctgttgta taactttttt ctttattttt gtcaatttcc tattttccag 540
agtcttgcat acccactatg tctactgtga tagtgaacgt aaaaacatac aagatgttgg 600
tgttatcctc aatctcagat cggaaagagct cgtatgccgt cttctgcttg 650

```

<210> 21
 <211> 605

ES 2 732 253 T3

<212> ADN
<213> *Homo sapiens*

<400> 21

5

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctct 60
gtgagaaatc aatgtctgct gtttataagc cgccgggctg tgatatcctg tgagagtggc 120
cccagtggat gaagacagat gctctcaagg agcgcagatg acgccgggtc cgaaggactc 180
ggcaccaccg ccggaggccg gcaacatggg caaggggctt ctcacggctg acctgtttcc 240
tcatcagcac atcaggacaa taagagctcc cacttcacag gtggggaaga gccaacgtgg 300
tgaagaatga ataaagcagc tcgtggaaag tgctgtgcat gaggcctggc aaccgggtccc 360
tgctctgagg tcacctgcca cggagctgct gacaggacca ttaaaaacac aattgtgcaa 420
gtgctcacc caccctcacag cagcagaatc tccaccagcc aagcattgga gacgatcctt 480
gcatccatag acatgaacag atgagcaaaa cgtgggtctat acggacgatg aaatagcact 540
cagccctaag aagaaataaa atcccgcacag agagatcggg agagctcgtg tgccgtcttc 600
tgctt

```

<210> 22
<211> 439
<212> ADN
<213> *Homo sapiens*

10

<400> 22

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctgc 60
tgagaggaga ggagaggctg gcagggggct gatgcaggag gtgatagggc tcccgggtgat 120
aagaggtgag aagaacagtc tctgtgtgcc tagagaagag attaccagaa gtctgctatc 180
tgtttgctcg cggatgtcgg acaggcagga tcgggtgatgg caggtcttgg ggaagatt 240
atcaggagct aaaagctgtc ttcaccttgg ctgctaagaa ctcctctcgg atcttcttag 300
aattccaaat cggacttttc tcttagcagt ggctacatcc ttaacctcaa aaataccctg 360
attagcagat ctacctccat gaaatagaca attcttgaca aactaagatc ggaagagctc 420
gtatgccgtc ttctgcttg

```

15

<210> 23
<211> 698
<212> ADN
<213> *Homo sapiens*

20

<400> 23

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctgt 60
tatttttcat actgatgatt ctaggggcgt tccctggact tggctgggac cagaactgtc 120
tcacagctgc cagcgtctcc caccoccatg agaaccatcg tgctgaagat gaagagccaa 180
agccaaactg accagcctct gcaaccoccaa gcatcaaccg gtgaggacag cgtgatgtgg 240
agccgcggag accccacgca gtgctgcgag gggcaaccac atgcaggggg gcagaggtgt 300
gggggagcaa gcagatgcca ccggcatgcc taggagccac agggagcatg cgggctgggc 360
agtgatggga atgaacgtga actccagtcc tgcccaga agtccctcgg cccctccttc 420
tcaattcagg gcacaaagtg gtaactgcag catgcagagg agtcgaggag tcttccctcg 480
tcccaggagc agcacctcgg gcacagtctc ggtcccacag aacagccaag tgtgggttgg 540
tggctagag acctcogaag atccagtggg ggaaggatgg gcagcagagg gtctactctc 600
tgaaaataag gggaaaggat tttccctccc cactgccaag gtcccagcta ctggacgtgg 660
gtggagatcg gaagagctcg tatgccgtct tctgcctt

```

REIVINDICACIONES

1. Un método de generación de una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3', comprendiendo el método:
- 5 proporcionar uno o más dúplex de polinucleótido diana con extremos romos, añadir un solo desoxinucleótido "A" a ambos extremos 3' de los dúplex de polinucleótido diana, produciendo así un saliente de una base en 3',
- 10 ligar polinucleótidos adaptadores con apareamiento erróneo idéntico en los dos extremos de cada uno de uno o más dúplex de polinucleótido diana para formar una o más construcciones adaptador-diana, en el que cada adaptador con apareamiento erróneo se forma a partir de dos cadenas polinucleotídicas hibridadas que forman un complejo bimolecular que comprende al menos una región bicatenaria y una región no apareada, y en el que cada adaptador con apareamiento erróneo comprende un saliente de una sola base en 3' que comprende un nucleótido "T" en el extremo bicatenario, que es complementario al saliente de una sola base en 3' que comprende un nucleótido "A" en los dúplex de polinucleótido diana,
- 15 realizar una reacción inicial de extensión de cebador en la que un oligonucleótido cebador hibrida con una parte de adaptador de cada una de las construcciones adaptador-diana y se extiende por adición secuencial de nucleótidos para formar productos de extensión complementarios a al menos una cadena de cada una de las construcciones adaptador-diana,
- 20 en el que los productos de extensión, y opcionalmente los productos de amplificación derivados de los mismos, proporcionan en su conjunto una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3';
- en el que los dúplex de polinucleótido diana a ligar son una mezcla compleja de fragmentos de ADN genómico que representan un genoma completo o esencialmente completo, y
- 25 en el que la biblioteca de moléculas polinucleotídicas plantilla generadas es representativa del genoma completo o esencialmente completo.
2. El método según la reivindicación 1, en el que la reacción inicial de extensión del cebador comprende:
- 30 a) hibridar un primer oligonucleótido cebador con una parte de adaptador de cada una de las construcciones de adaptador-diana,
- b) extender los cebadores por adición secuencial de nucleótidos para formar productos de extensión complementarios a al menos una cadena de cada una de las construcciones adaptador-diana, y
- 35 c) someter los productos obtenidos en la etapa b) a condiciones desnaturalizantes, separando de este modo los productos de extensión de cadenas de las construcciones adaptador-diana.
3. El método según cualquiera de las reivindicaciones 1 a 2, en el que los cebadores usados en la reacción inicial de extensión del cebador hibridan con una región no apareada de los adaptadores.
- 40 4. El método según cualquiera de las reivindicaciones 1 a 3, en el que la reacción de extensión del cebador se lleva a cabo como parte de una reacción en cadena de la polimerasa y los productos de amplificación de la reacción de PCR se recogen para proporcionar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y 3'.
- 45 5. El método según la reivindicación 4, en el que la reacción en cadena de la polimerasa se realiza usando primeros cebadores de oligonucleótido que son capaces de hibridar con partes de adaptador de cada una de las construcciones adaptador-diana y segundos cebadores de oligonucleótido que son capaces de hibridar con una región de las cadenas extendidas producidas por extensión de los primeros cebadores de oligonucleótido, siendo esta región complementaria a una parte de adaptador de las construcciones adaptador-diana.
- 50 6. El método según la reivindicación 5, en el que los primeros cebadores de oligonucleótido son capaces de hibridar con la región no apareada en la primera cadena de las construcciones adaptador-diana y los segundos cebadores de oligonucleótido son capaces de hibridar con un producto de extensión complementario a una región no apareada en una segunda cadena de las construcciones adaptador-diana.
- 55 7. El método según la reivindicación 5, en el que los primer y segundos cebadores de oligonucleótido tienen secuencias de nucleótidos esencialmente idénticas, en el que opcionalmente los primer y segundos cebadores de oligonucleótido son capaces de hibridar con una región bicatenaria en las partes de adaptador de las construcciones de adaptador-diana.
- 60 8. El método según cualquiera de las reivindicaciones 1 a 7, en el que los polinucleótidos adaptadores con apareamiento erróneo son adaptadores bifurcados formados por hibridación de primeras y segundas cadenas polinucleotídicas parcialmente complementarias, en donde una secuencia de 5 o más nucleótidos consecutivos en el extremo 3' de la primera cadena es complementaria a una secuencia de 5 o más nucleótidos consecutivos en el
- 65 extremo 5' de la segunda cadena de tal manera que se forma una región bicatenaria de 5 o más pares de bases

consecutivos hibridando las dos cadenas y en donde una secuencia de al menos 10 nucleótidos consecutivos en el extremo 5' de la primera cadena y una secuencia de al menos 10 nucleótidos consecutivos en el extremo 3' de la segunda cadena no son complementarias de tal manera que la región no apareada de al menos 10 nucleótidos consecutivos en cada cadena permanece como una forma monocatenaria cuando hibrida la región bicatenaria.

- 5
9. El método según cualquiera de las reivindicaciones 1 a 8, en el que los dúplex de polinucleótido diana son fragmentos de ADN genómico, en el que opcionalmente los dúplex de polinucleótido diana son fragmentos de un genoma completo.
- 10 10. Un método de amplificación de ácidos nucleicos en fase sólida de moléculas polinucleotídicas plantilla que comprende:
preparar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3' usando el método según cualquiera de las reivindicaciones 1 a 9, y realizar una reacción de amplificación de ácido nucleico en fase sólida en el que se amplifican dichas moléculas
- 15 polinucleotídicas plantilla.
11. El método según la reivindicación 10, en el que dichas moléculas polinucleotídicas plantilla se amplifican por PCR en fase sólida usando cebadores de amplificación directos e inversos, en el que opcionalmente dicha biblioteca de moléculas polinucleotídicas plantilla se prepara usando el método de cualquiera de las reivindicaciones 4 a 8 y la
- 20 PCR en fase sólida se lleva a cabo usando cebadores idénticos a los usados en la preparación de la biblioteca de moléculas polinucleotídicas plantilla, o en el que dicha biblioteca de moléculas polinucleotídicas plantilla se prepara usando el método de cualquiera de las reivindicaciones 4 a 8 y una PCR en fase sólida se lleva a cabo usando cebadores que no son idénticos a los usados en la preparación de la biblioteca de moléculas polinucleotídicas plantilla o en el que dicha biblioteca de moléculas polinucleotídicas plantilla se prepara usando el método de
- 25 cualquiera de las reivindicaciones 1 a 3 y una PCR en fase sólida se lleva a cabo usando cebadores de amplificación directos capaces de hibridar con una parte de adaptador de los productos de extensión formados en la reacción inicial de extensión del cebador y cebadores inversos que son capaces de hibridar con productos de extensión formados por la extensión de cebadores de amplificación directos hibridados con los productos de extensión formados en la reacción inicial de extensión con cebador.
- 30
12. Uso de una biblioteca de moléculas polinucleotídicas plantilla preparada según el método de cualquiera de las reivindicaciones 1 a 9 como plantilla para amplificación por PCR en fase sólida.
13. Uso de una biblioteca de moléculas polinucleotídicas plantilla preparada según el método de la reivindicación 9
- 35 como plantilla para la amplificación del genoma completo.
14. Un método de amplificación del genoma completo que comprende:
preparar una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y en sus extremos 3' usando el método según la reivindicación 9, y llevar a cabo una reacción de amplificación de
- 40 ácido nucleico en la que se amplifican dichas moléculas polinucleotídicas plantilla.
15. Un kit para su uso en la preparación de una biblioteca de moléculas polinucleotídicas plantilla que tienen secuencias comunes en sus extremos 5' y secuencias comunes en sus extremos 3', comprendiendo el kit polinucleótidos adaptadores con apareamiento erróneo como se define en cualquiera de las reivindicaciones 1 a 9 y
- 45 cebadores de oligonucleótido capaces de hibridar con los polinucleótidos adaptadores con apareamiento erróneo, en donde los cebadores oligonucleotídicos son cebadores con cola que comprenden una parte de cola en 5' que no hibrida con los polinucleótidos adaptadores con apareamiento erróneo, en el que opcionalmente los cebadores oligonucleotídico son capaces de hibridar con una región no apareada de los polinucleótidos adaptadores con apareamiento erróneo, en el que preferentemente los polinucleótidos adaptadores con apareamiento erróneo son
- 50 adaptadores bifurcados como se define en la reivindicación 8.

Fig. 1

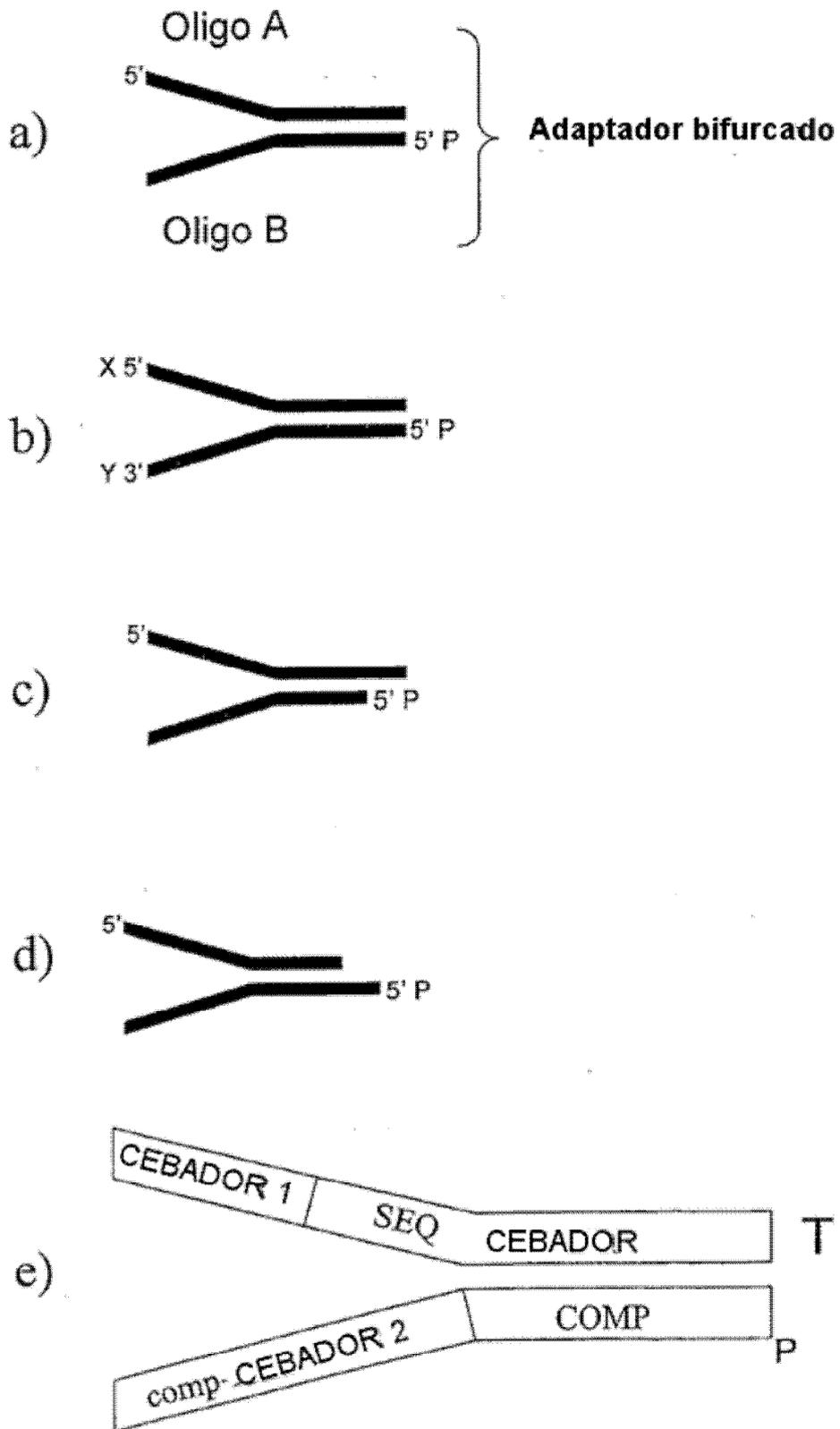


Fig. 2a

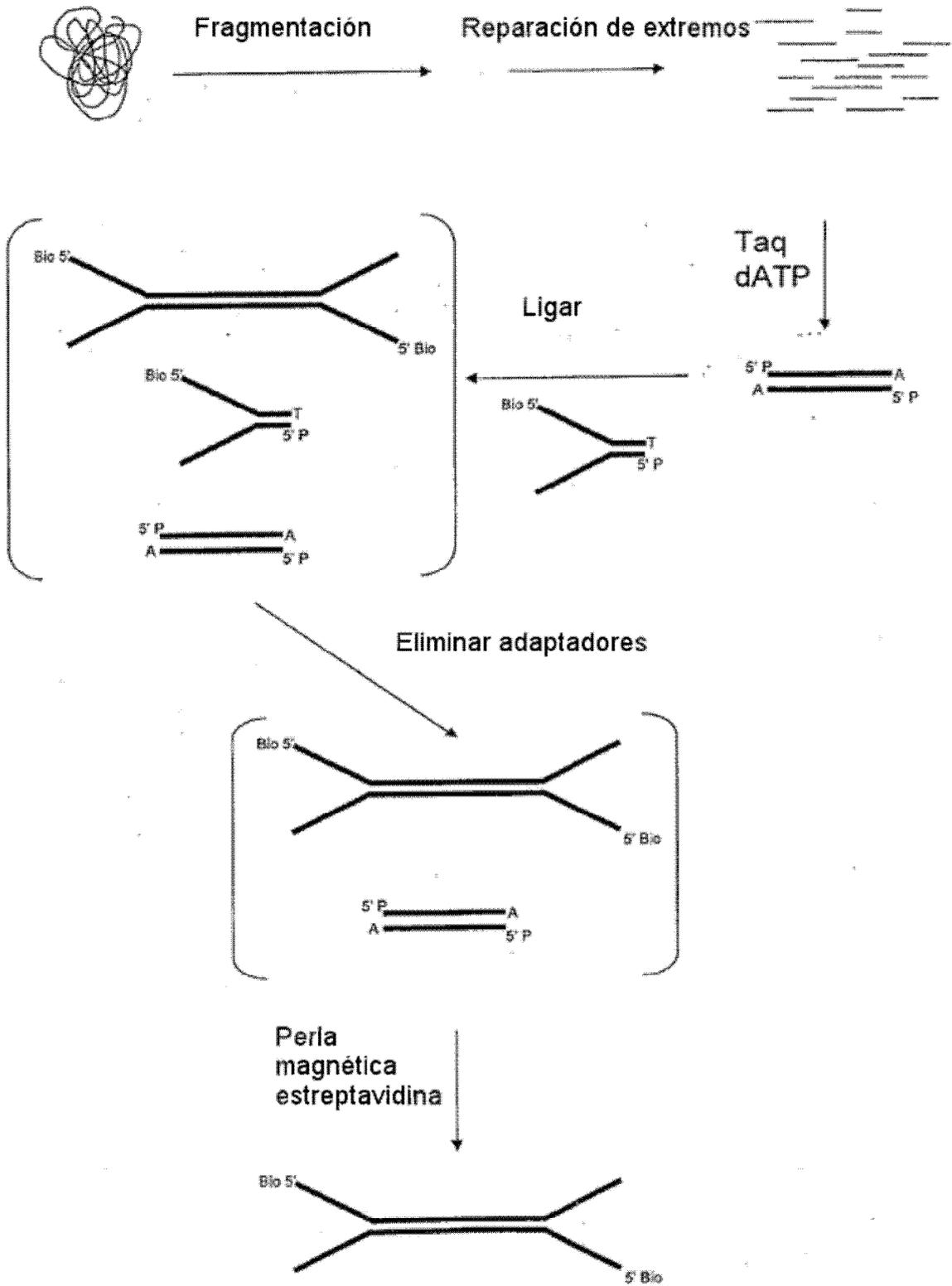


Fig. 2b

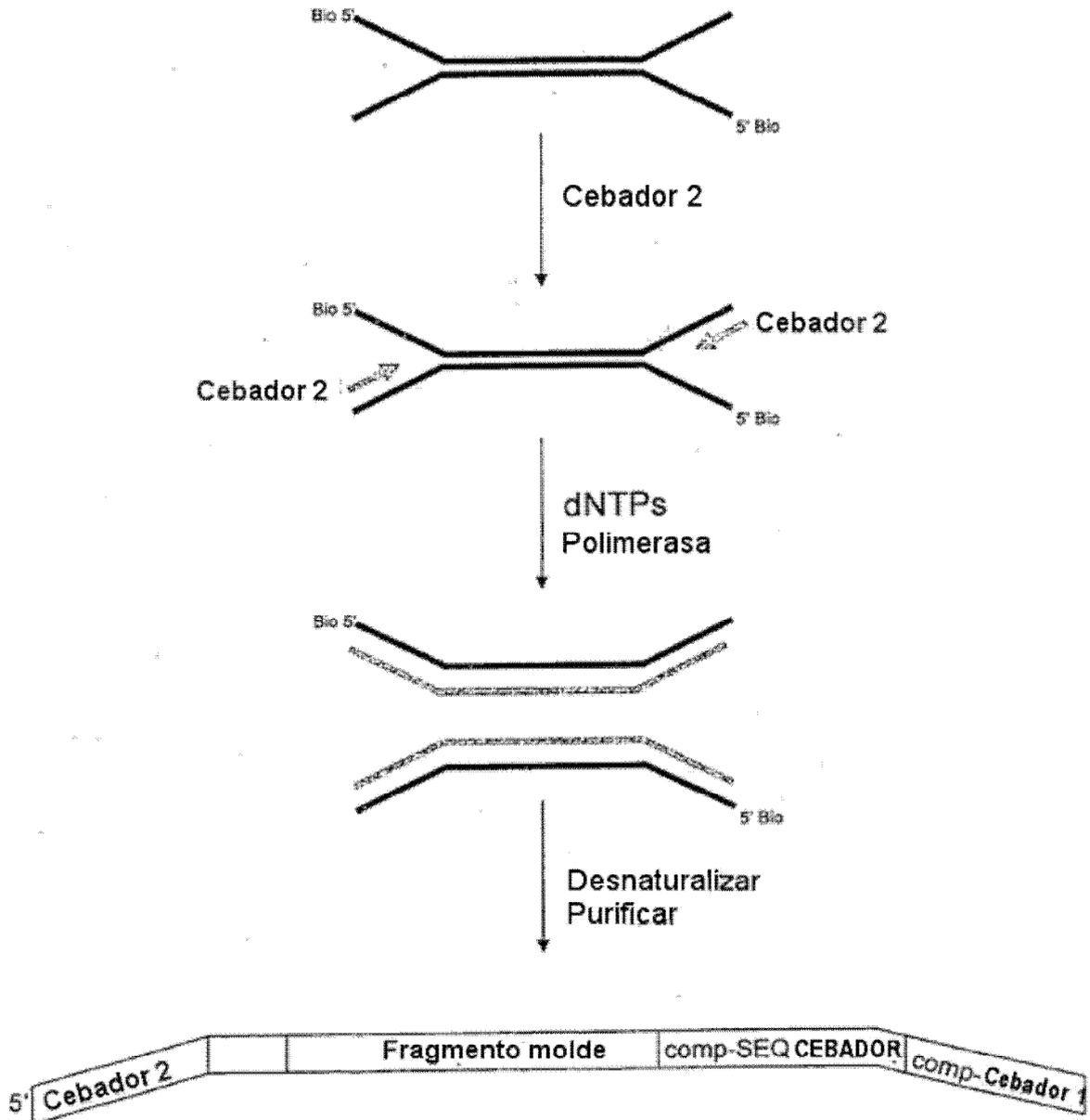


Fig. 3

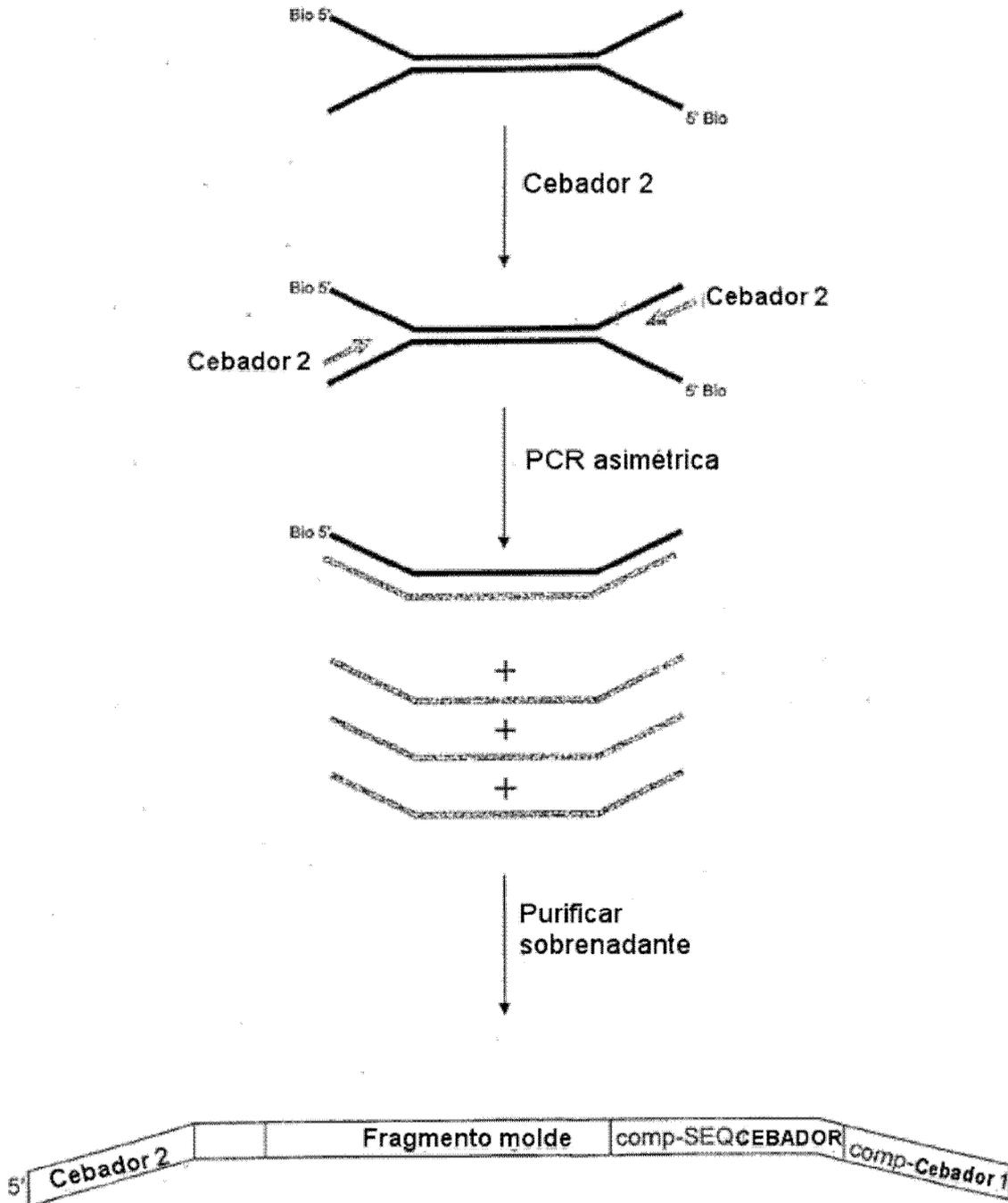


Fig. 4

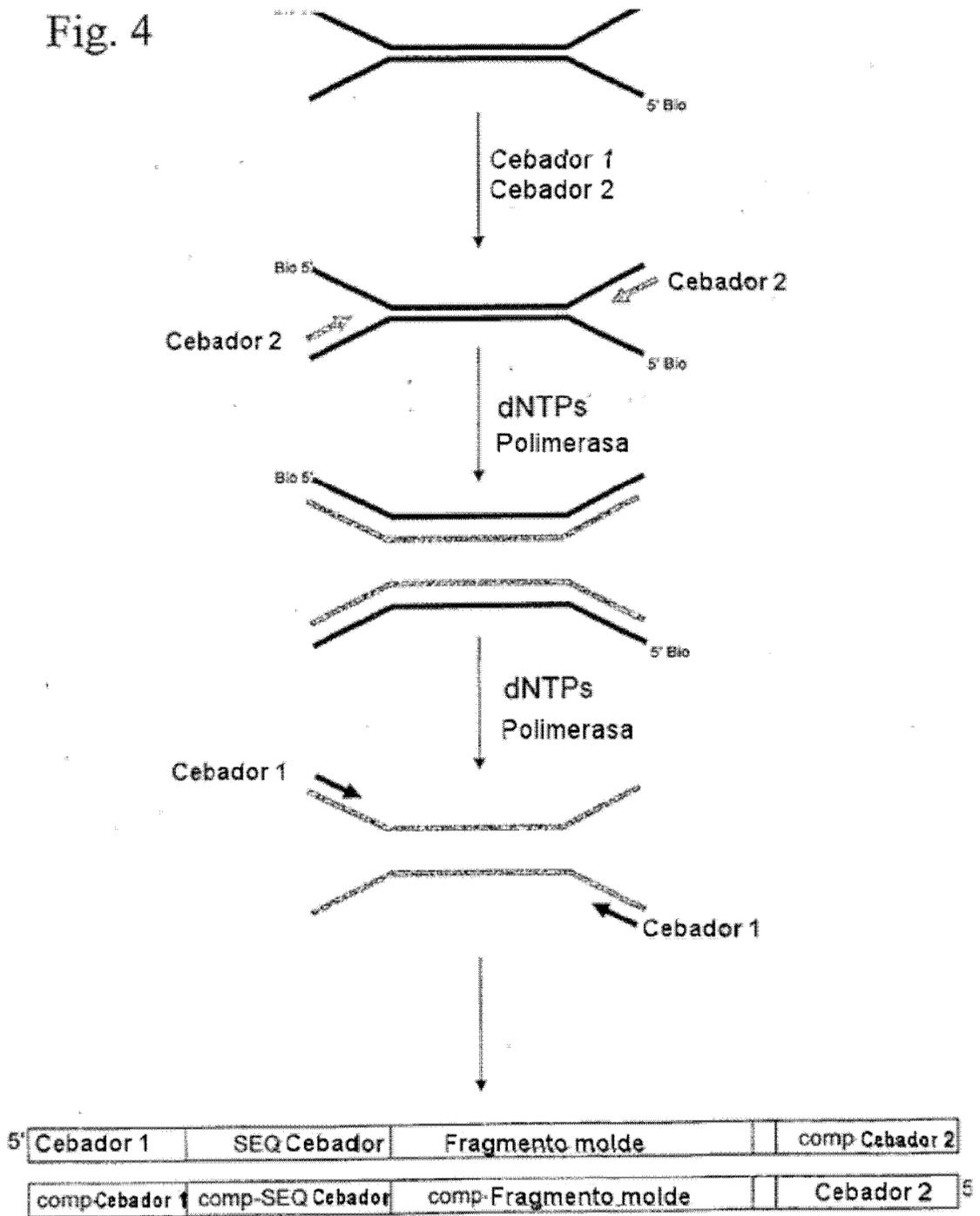
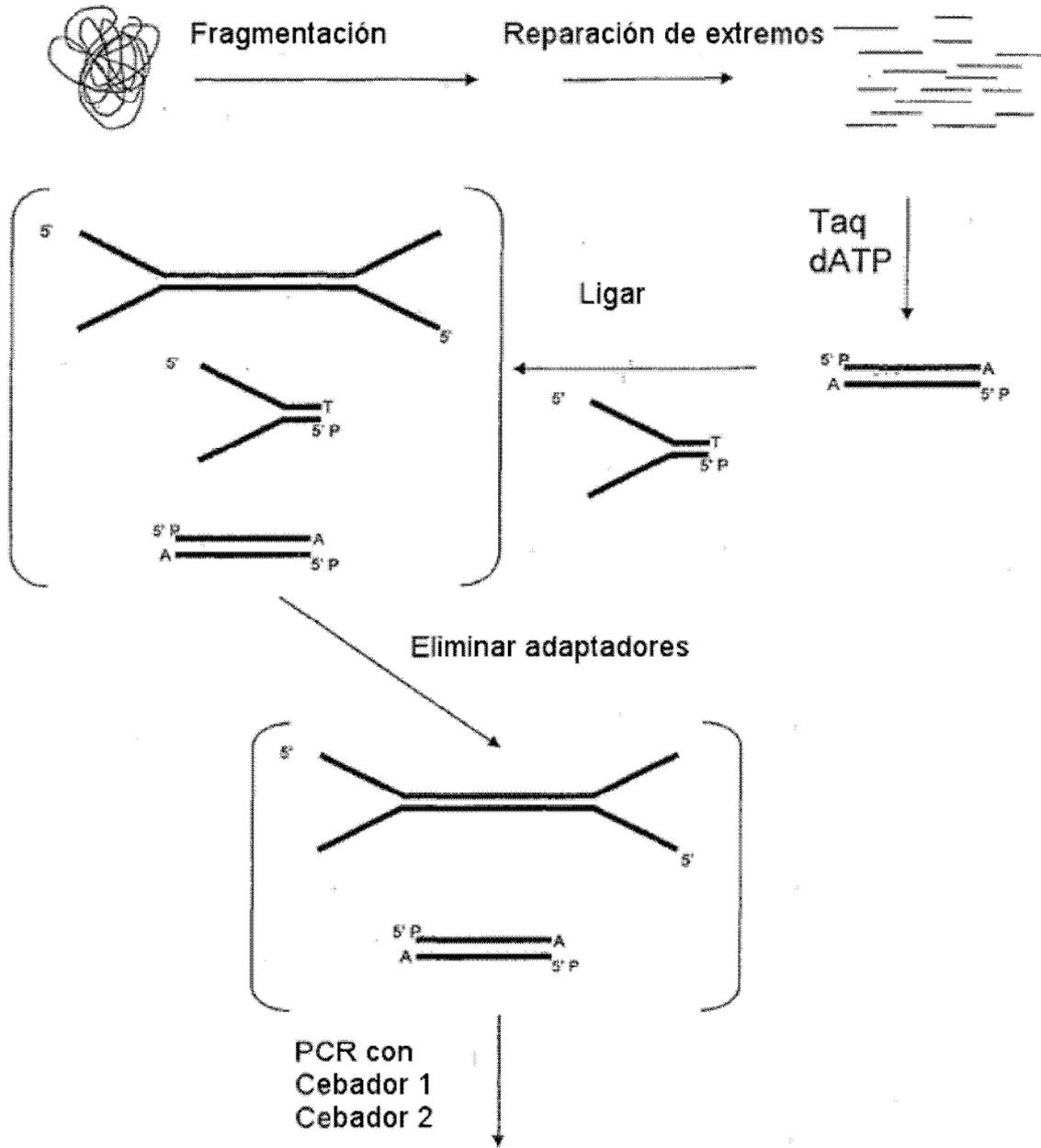


Fig. 5



5'	Cebador 1	SEQ Cebador	Fragmento molde	comp-Cebador 2
	comp-Cebador 1	comp-SEQ Cebador	comp-Fragmento molde	Cebador 2

Fig. 6

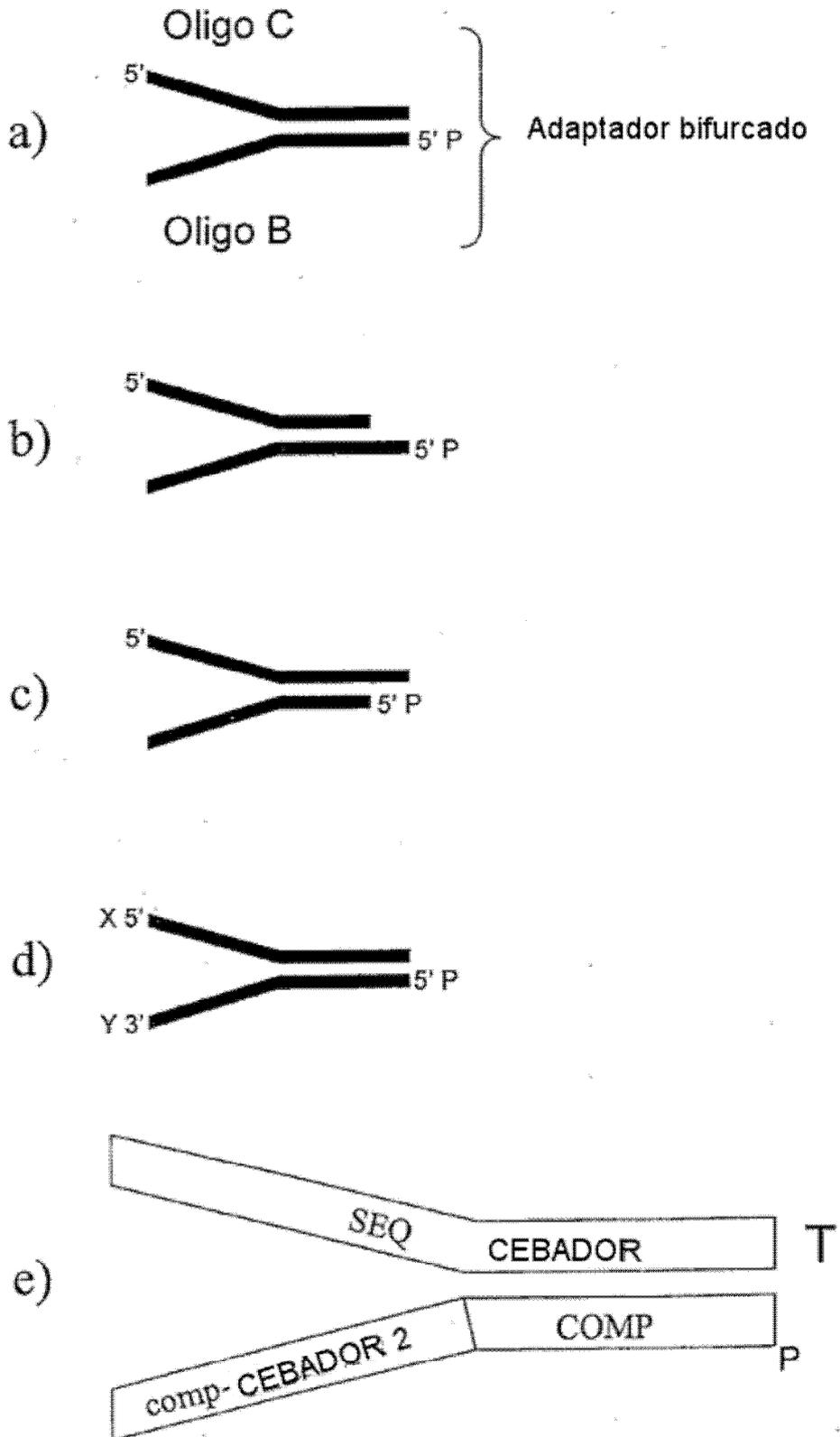


Fig. 7a

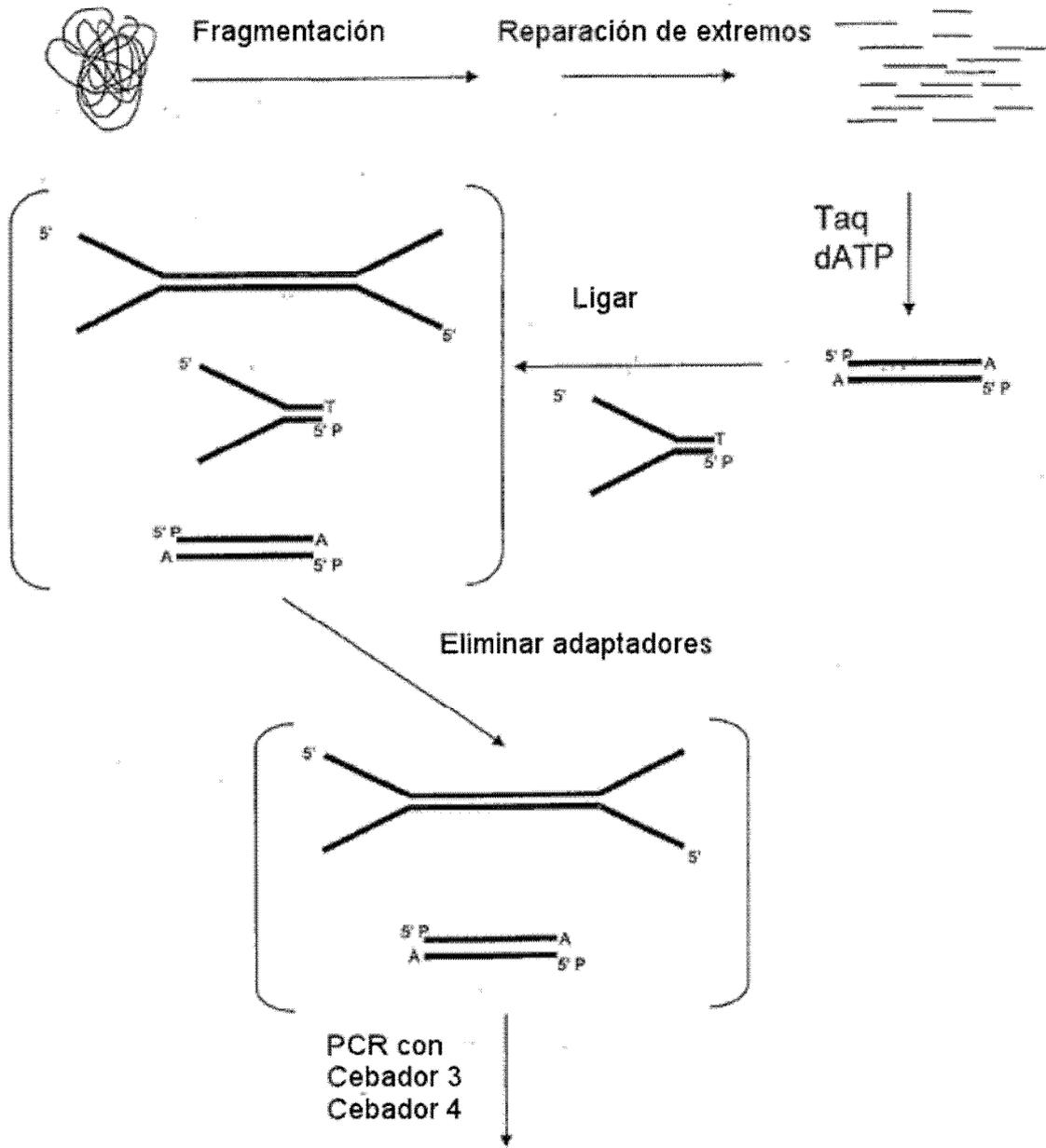


Fig. 7b

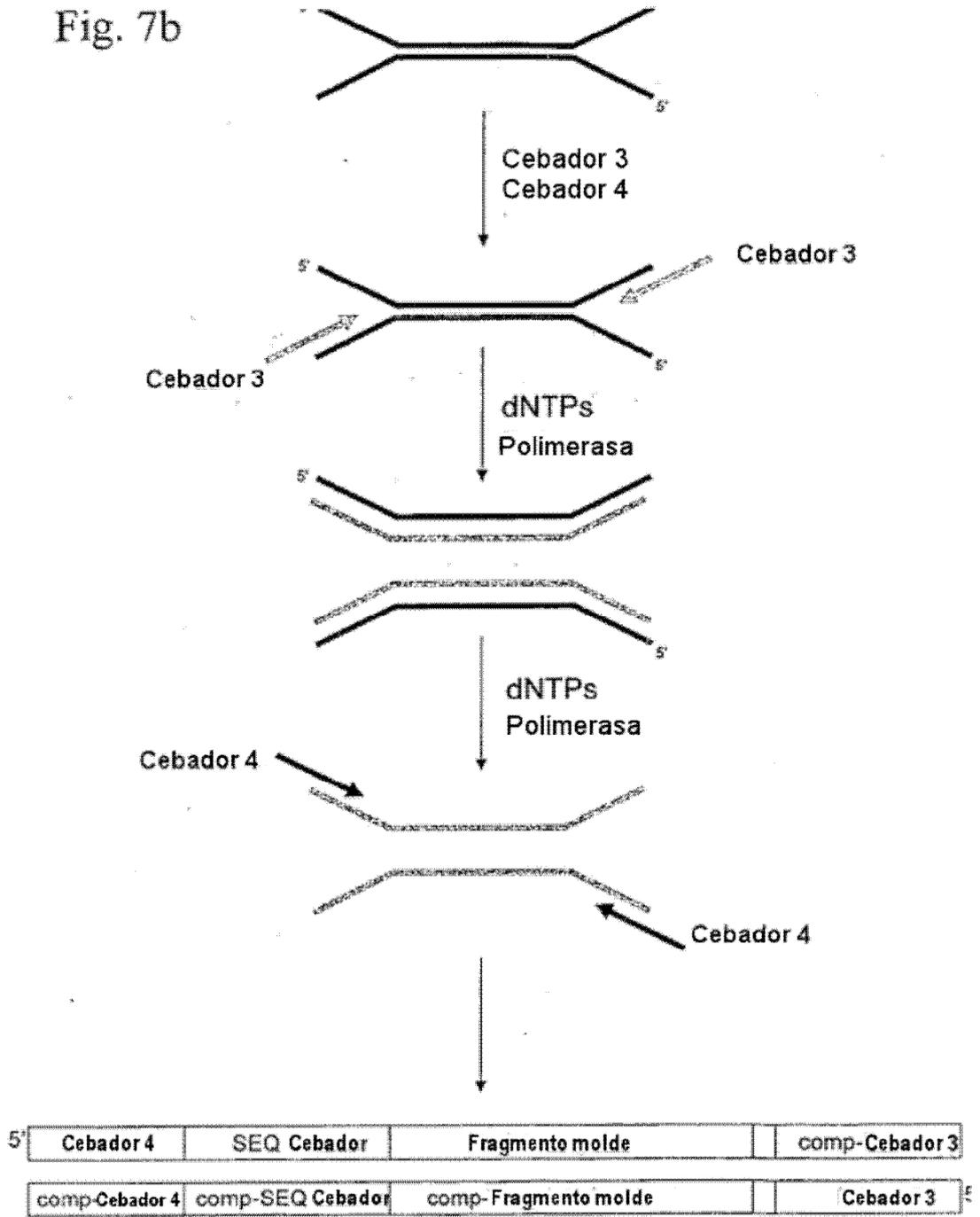


Fig. 8

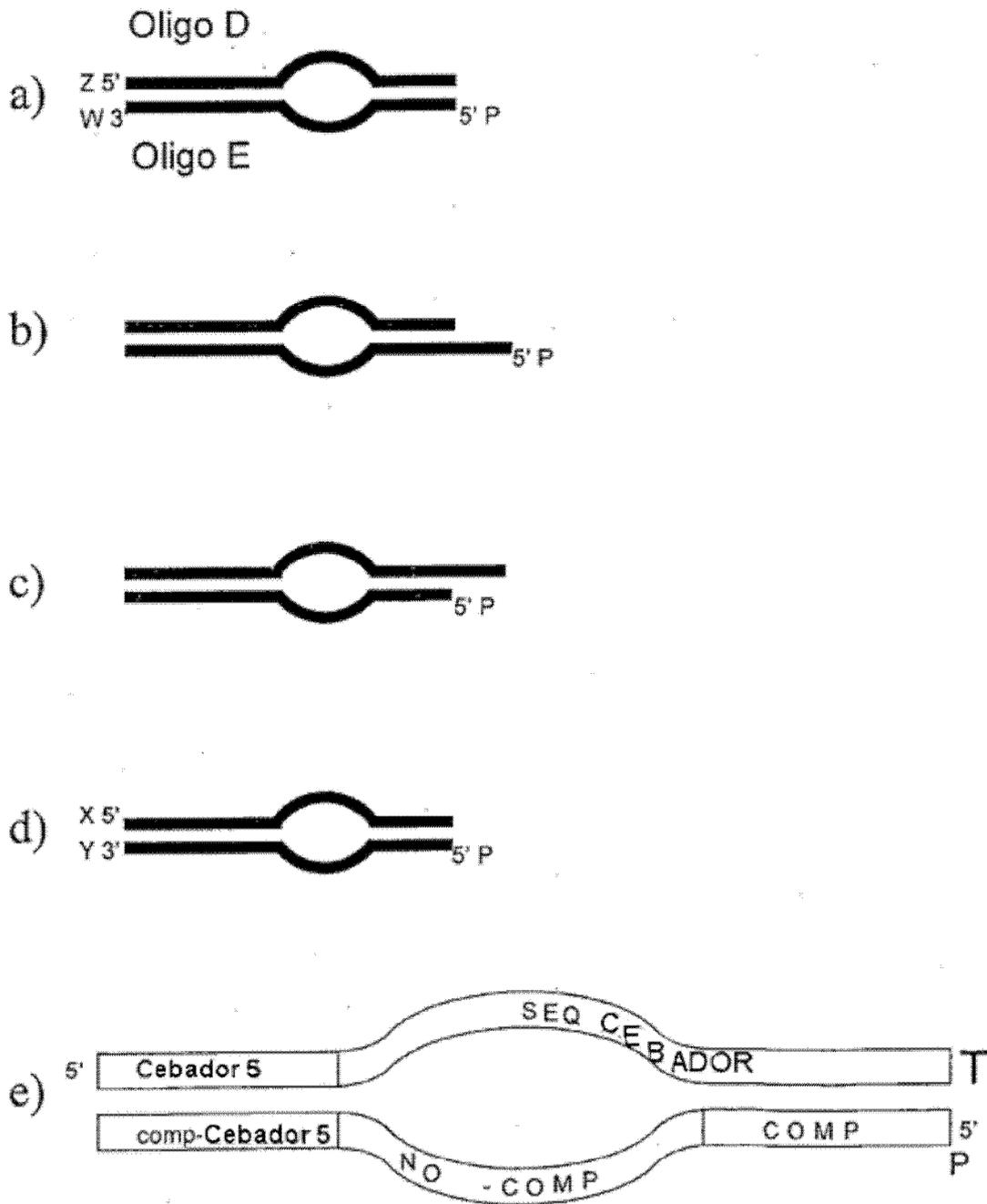


Fig. 9a

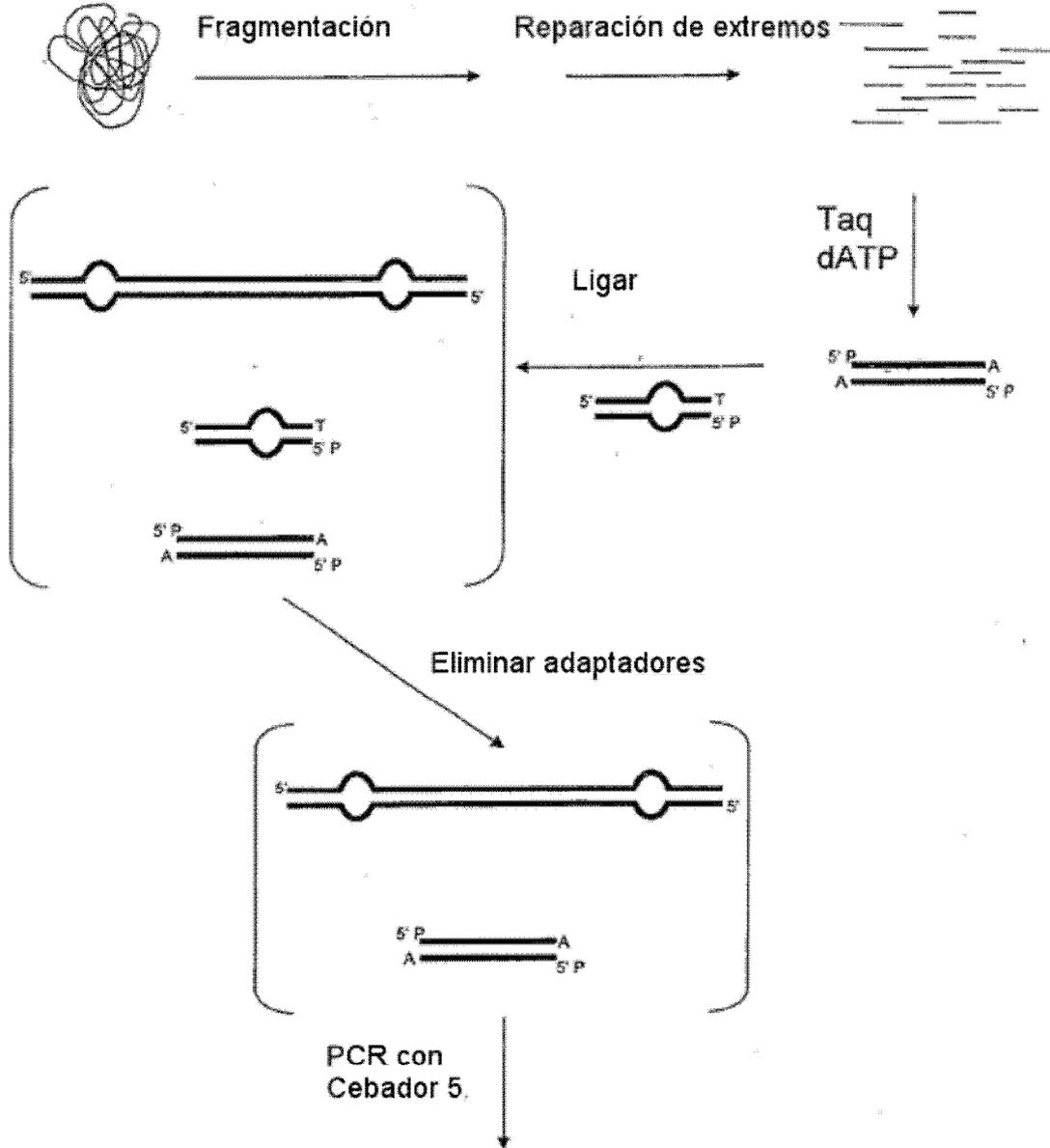
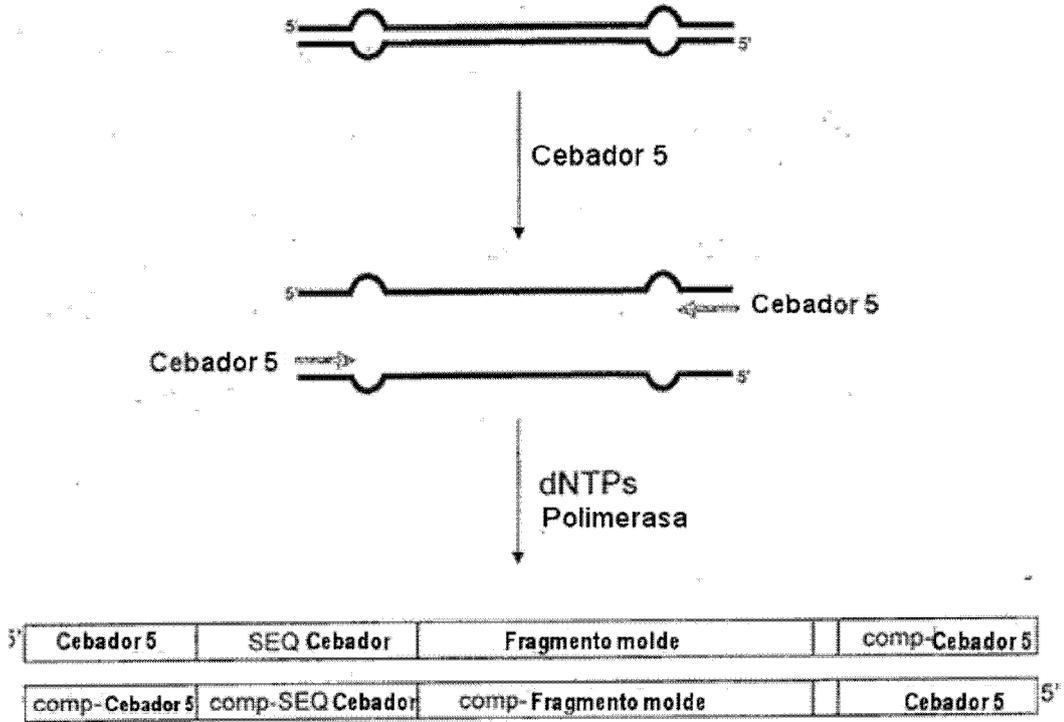


Fig. 9b



REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante es únicamente para la comodidad del lector. No forma parte del documento de la patente europea. A pesar del cuidado tenido en la recopilación de las referencias, no se pueden excluir errores u omisiones y la EPO niega toda responsabilidad en este sentido.

Documentos de patentes citados en la descripción

- 10 • WO 9844151 A [0005] [0006] [0008] [0102]
- WO 0018957 A [0005] [0006] [0008] [0102]
- WO 05068656 A [0047]

15 Literatura diferente de patentes citada en la descripción

- **MITRA, R.D ; CHURCH, G.M.** *Nucleic Acids Research*, 1999, vol. 27 (24 [0006]
- 20 • **SAMBROOK et al.** *Molecular Cloning, A Laboratory Manual*. Cold Spring Harbor Laboratory Press, 2001 [0040] [0041] [0111]
- *Current Protocols* [0040] [0041] [0111]