

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 424 155**

51 Int. Cl.:

C12Q 1/68

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **08.10.2007 E 07824052 (0)**

97 Fecha y número de publicación de la concesión europea: **22.05.2013 EP 2084295**

54 Título: **Procedimiento para la secuenciación de un molde polinucleotídico**

30 Prioridad:

06.10.2006 US 850210 P
01.02.2007 US 898910 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
27.09.2013

73 Titular/es:

ILLUMINA CAMBRIDGE LIMITED (100.0%)
Chesterford Research Park Little Chesterford
Saffron Walden
Essex CB10 1XL, GB

72 Inventor/es:

RIGATTI, ROBERTO;
OST, TOBIAS y
FASHENA, SARAH

74 Agente/Representante:

ZEA CHECA, Bernabé

ES 2 424 155 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

Procedimiento para la secuenciación de un molde polinucleotídico

5 **Campo de la invención**

[0001] La invención se refiere a procedimientos para la secuenciación por pares de un molde de polinucleótido de doble hebra, a procedimientos que tienen como resultado la determinación secuencial de las secuencias de nucleótidos en dos regiones distintas y separadas del molde de polinucleótido.

10

Antecedentes de la invención

[0002] En esta solicitud se hace referencia a varias publicaciones y documentos de patente con el fin de describir más completamente el estado de la técnica a la que pertenece esta invención.

15

[0003] Los avances en el estudio de moléculas biológicas han ido precedidos, en parte, por la mejora en las tecnologías usadas para caracterizar las moléculas o sus reacciones biológicas. En particular, el estudio de los ácidos nucleicos ADN y ARN se ha beneficiado del desarrollo de tecnologías que se utilizan para el análisis de secuencia.

20

[0004] Un procedimiento para la secuenciación de una molde de polinucleótido implica la realización de múltiples reacciones de extensión utilizando una ADN polimerasa para incorporar sucesivamente nucleótidos marcados a una hebra molde. En una reacción de "secuenciación por síntesis" de este tipo, se crea una nueva hebra de nucleótidos con bases emparejadas a la hebra molde en la dirección 5' a 3' por la incorporación sucesiva de nucleótidos individuales complementarios a la hebra molde. Si se utiliza de forma simultánea, los nucleósidos trifosfato sustrato utilizados en la reacción de secuenciación pueden ser bloqueados para evitar el exceso de incorporación y marcados diferentes, lo que permite la determinación de la identidad del nucleótido incorporado a medida que se añaden nucleótidos sucesivos.

25

[0005] Con el fin de llevar a cabo la secuenciación precisa se puede añadir una modificación estructural terminadora de hebra reversible o "grupo de bloqueo" a los nucleótidos sustrato para garantizar que los nucleótidos se incorporan de uno en uno de una manera controlada. A medida que se incorpora cada nucleótido único, el grupo de bloqueo impide la incorporación de más nucleótidos en la hebra de polinucleótido. Una vez que la identidad del nucleótido marcado incorporado en último lugar se ha determinado, se eliminan el resto marcador y el grupo de bloqueo, permitiendo que el siguiente nucleótido marcado bloqueado se incorpore en una ronda posterior de secuenciación.

30

[0006] En determinadas circunstancias, la cantidad de datos de la secuencia que se puede obtener con fiabilidad con el uso de técnicas de secuenciación por síntesis, particularmente cuando se usan nucleótidos marcados bloqueados puede ser limitada. En algunas circunstancias, se prefiere limitar el "ciclo" de secuenciación a un número de bases que permite la realineación de la secuencia con el genoma humano, habitualmente alrededor de 25-30 ciclos de incorporación. Aunque los ciclos de secuenciación de esta longitud son extremadamente útiles, en particular en aplicaciones, tales como, por ejemplo, análisis de SNP y genotipado, sería ventajoso en muchas circunstancias poder obtener de manera fiable más datos de la secuencia para la misma molécula molde.

35

[0007] La técnica de secuenciación de "extremos emparejados" o secuenciación "por pares" se conoce habitualmente en la técnica de la biología molecular, en particular en el contexto de la secuenciación aleatoria del todo el genoma. La secuenciación de extremos emparejados permite la determinación de dos "lecturas" de la secuencia de dos lugares en un solo dúplex polinucleotídico. La ventaja del procedimiento de extremos emparejados es que existe significativamente más información que se puede obtener a partir de la secuenciación de dos tramos cada uno de "n" bases de un solo molde que de la secuenciación de "n" bases de cada uno de dos moldes independientes de forma aleatoria. Con el uso de herramientas de software adecuadas para el montaje de la información de la secuencia, es posible hacer uso del conocimiento de que las secuencias con "extremos emparejados" no son completamente aleatorias, sino que se sabe que se producen en una sola dúplex, y por lo tanto están vinculadas o emparejadas en el genoma. Se ha demostrado que esta información ayuda en gran medida al montaje de secuencias del genoma completo en una secuencia de consenso.

40

45

50

[0008] La secuenciación de extremos emparejados normalmente se ha realizado mediante el uso de vectores de clonación aleatorios circulares especializados. Después de cortar el vector en un sitio simple específico, el molde de ADN a secuenciar (típicamente ADN genómico) se inserta en el vector y los extremos se vuelven a cerrar para formar una nueva construcción. Las secuencias del vector que flanquean el inserto de ADN incluyen sitios de unión para cebadores de secuenciación que permiten la secuenciación del inserto de ADN en hebras opuestas. Sin embargo, la necesidad de cebadores de secuenciación en ambos extremos del fragmento de molde hace que el uso de técnicas de secuenciación basadas en matrices sea extremadamente difícil. Con las técnicas basadas en matrices, que por lo general se basan en un molde de hebra sencilla, por lo general sólo es posible secuenciar desde un extremo de un molde de nucleótidos, ya que la hebra complementaria no está unida a la superficie.

55

60

65

[0009] Se han descrito varios procedimientos para la secuenciación de extremos doble de una molde de polinucleótido que se pueden llevar a cabo sobre un soporte sólido, por ejemplo, en los documentos US20060024681, US20060292611, WO06110855, WO06135342, WO03074734, WO07010252, WO07091077y WO00179553.

5

[0010] El documento DE10051564A1 es una solicitud de patente alemana de Axaron Bioscience, que divulga procedimientos de secuenciación utilizando moldes de ácidos nucleicos inmovilizados y cebadores de colonias, que pueden hibridarse al extremo 3' del molde, en el que los moldes se copian por extensión del cebador y la secuenciación se lleva a cabo a partir de un extremo.

10

[0011] Los documentos WO 98/44151y WO 00/18957 describen procedimientos de amplificación de ácido nucleico que permiten inmovilizar los productos de amplificación en un soporte sólido con el fin de formar matrices compuestas de agregados o "colonias" formadas a partir de una pluralidad de hebras idénticas de polinucleótidos inmovilizadas y una pluralidad de hebras complementarias inmovilizadas idénticas. Las moléculas de ácidos nucleicos presentes en las colonias de ADN en las matrices agrupadas preparadas de acuerdo con estos procedimientos pueden proporcionar moldes para las reacciones de secuenciación, por ejemplo, como se describe en el documento WO 98/44152. Es ventajoso permitir la secuenciación eficiente de ambas hebras de tales agregados, como se describe en detalle en los procedimientos de la presente memoria.

15

20 Descripción resumida de la invención

[0012] Los presentes inventores han desarrollado ahora procedimientos tal como se definen en las reivindicaciones para la secuenciación con extremos emparejados de moldes de polinucleótidos de doble hebra, incluyendo moldes de doble hebra presentes en las matrices agregadas, tales como las descritas en este documento. Los procedimientos permiten la secuenciación de dos regiones distintas, una en cada extremo de las hebras complementarias de un dúplex de polinucleótido diana. Utilizando los procedimientos de la invención, es posible obtener dos lecturas unidas o emparejadas de la información de la secuencia de cada molde de doble hebra en una matriz agregada, en lugar de una sola secuencia de lectura de una hebra del molde.

25

[0013] De acuerdo con un procedimiento de la invención, se proporciona un procedimiento para la secuenciación por pares de una primera y segunda regiones de un polinucleótido de doble hebra diana, en el que dichas primera y segunda regiones están en el mismo polinucleótido de doble hebra diana, comprendiendo el procedimiento:

30

(a) proporcionar un soporte sólido que tiene inmovilizado sobre él una pluralidad de polinucleótidos molde de doble hebra formados cada uno a partir de una primera y segunda hebras molde complementarias unidas al soporte sólido en sus extremos 5' y múltiples copias de uno o más cebadores inmovilizados en el extremo 5' capaces de hibridar con el extremo 3' de la primera hebra del molde;

35

(b) tratar la pluralidad de polinucleótidos molde de doble hebra de tal manera que las primeras hebras molde hibridan con cebadores inmovilizados en el extremo 5';

40

(c) llevar a cabo una primera lectura de secuenciación para determinar la secuencia de una primera región del polinucleótido molde;

(d) llevar a cabo una reacción de extensión para extender uno o más de los cebadores inmovilizados para copiar la primera hebra del molde para generar una segunda hebra del molde inmovilizada;

45

(e) tratar la pluralidad de una primera y segunda hebras del molde inmovilizadas para eliminar la primera hebra del molde del soporte sólido;

(f) llevar a cabo una segunda lectura de secuenciación para determinar la secuencia de una segunda región del polinucleótido molde, en el que determinar las secuencias de la primera y segunda regiones del polinucleótido diana logra la secuenciación por pares de dichas primera y segunda regiones de dicho polinucleótido de doble hebra diana.

50

[0014] Un segundo procedimiento de la invención proporciona un procedimiento para la secuenciación de regiones de los extremos A y B de un polinucleótido de doble hebra diana, en el que dichas regiones terminales A y B están en el mismo polinucleótido de doble hebra diana, comprendiendo el procedimiento:

55

(a) proporcionar un soporte sólido que tiene inmovilizado sobre él una pluralidad de polinucleótidos molde de doble hebra formados cada uno a partir de una primera y segunda hebras molde complementarias unidas al soporte sólido en sus extremos 5';

60

(b) tratar los polinucleótidos molde de doble hebra, de manera que cada polinucleótido molde de doble hebra se corta en al menos dos lugares para generar dos fragmentos A y B molde de doble hebra más cortos inmovilizados en un extremo, en el que A y B ya no están conectados directamente;

(c) tratar los dos fragmentos A y B molde de doble hebra más cortos inmovilizados en un extremo para hacer de los dos extremos no inmovilizados un dúplex de extremos romos;

65

(d) tratar los dos dúplex de extremos romos de tal manera que los dos extremos romos A y B estén conectados para formar una secuencia de nucleótidos de doble hebra que contiene ambos extremos A y B del fragmento diana original en una secuencia contigua acortada, inmovilizado en ambos extremos;

(e) escindir una hebra de la secuencia de nucleótidos de doble hebra que contiene ambos extremos distales A y B del fragmento diana original en una secuencia contigua acortada inmovilizado en ambos extremos para generar una secuencia de nucleótidos diana de hebra sencilla que contiene ambos extremos distales A y B del fragmento diana original en una secuencia contigua acortada, en el que dicha secuencia de nucleótidos diana de hebra sencilla se inmoviliza en un extremo 5' o 3';

(f) hibridar un cebador de secuenciación con la secuencia de nucleótidos diana de hebra sencilla que contiene ambos extremos distales A y B del fragmento diana original en una secuencia contigua acortada y

(g) llevar a cabo una reacción de secuenciación única para determinar una secuencia contigua de ambos extremos A y B del fragmento diana original.

[0015] Un tercer procedimiento de la invención proporciona la secuenciación por pares de una primera y segunda regiones de un polinucleótido de doble hebra diana, en el que dicha primera y segunda regiones están en el mismo polinucleótido de doble hebra diana, comprendiendo el procedimiento:

(a) proporcionar un soporte sólido que tiene inmovilizado sobre él una pluralidad de polinucleótidos molde de doble hebra formados cada uno a partir de una primera y segunda hebras molde complementarias unidas al soporte sólido en sus extremos 5' y múltiples copias de uno o más cebadores inmovilizados en el extremo 5' capaces de hibridar con el extremo 3' de la primera hebra del molde;

(b) tratar la pluralidad de polinucleótidos molde de doble hebra de tal manera que las primeras hebras del molde hibriden con un cebador que está inmovilizado en el soporte sólido en su extremo 5';

(c) llevar a cabo una primera secuenciación para determinar la secuencia de una primera región del polinucleótido molde;

(d) llevar a cabo una reacción de extensión para extender uno o más de los cebadores inmovilizados con el extremo de la primera hebra del molde para generar una segunda hebra del molde inmovilizado;

(e) tratar la pluralidad de polinucleótidos molde de tal manera que la primera hebra del molde se separa del soporte sólido dejando la segunda hebra del molde de hebra sencilla inmovilizada;

(f) llevar a cabo una segunda lectura de secuenciación para determinar la secuencia de una segunda región del polinucleótido molde, en el que determinar las secuencias de la primera y segunda regiones del polinucleótido diana logra la secuenciación por pares de dichas primera y segunda regiones de dicho polinucleótido de doble hebra diana.

[0016] En un ejemplo más específico, dicho tercer procedimiento es un procedimiento para la secuenciación por pares de una primera y segunda regiones de un polinucleótido de doble hebra diana, en el que dichas primera y segunda regiones están en el mismo polinucleótido de doble hebra diana, comprendiendo el procedimiento:

(a) proporcionar un soporte sólido que tiene inmovilizado sobre él una pluralidad de polinucleótidos molde de doble hebra formados cada uno a partir de una primera y segunda hebras molde complementarias unidas al soporte sólido en sus extremos 5';

(b) tratar la pluralidad de polinucleótidos molde de doble hebra de tal manera que una de las hebras se liberan de la superficie, dejando una única primera hebra del molde de hebra sencilla inmovilizada sobre el soporte sólido en su extremo 5';

(c) hibridar un cebador con dicha primera hebra del molde y llevar a cabo una primera reacción de secuenciación para controlar la incorporación de nucleótidos marcados en el cebador hibridado usando ciclos de extensión del cebador con una polimerasa y nucleótidos marcados para generar un primer cebador de secuenciación extendida y determinar la secuencia de una primera región del polinucleótido molde;

(d) eliminar dicho primer cebador de secuenciación extendida;

(e) hibridar la primera hebra del molde inmovilizado con cebadores inmovilizados y extendiéndose dichos cebadores inmovilizados para regenerar dicha pluralidad de polinucleótidos molde de hebra doble formados cada uno a partir de las primera y segunda hebras del molde complementarias unidas al soporte sólido en sus extremos 5';

(f) el tratamiento de la pluralidad de polinucleótidos molde de tal manera que la primera hebra del molde se elimina de la superficie, dejando la segunda hebra del molde inmovilizada en forma de hebra sencilla;

(g) hibridar un segundo cebador de secuenciación con las segundas hebras del molde inmovilizadas y

(h) llevar a cabo una segunda secuenciación para controlar la incorporación de nucleótidos marcados en el segundo cebador de secuenciación mediante ciclos de extensión del cebador con una polimerasa y nucleótidos marcados para generar un segundo cebador de secuenciación extendida y determinar la secuencia de una segunda región del polinucleótido del molde, en el que determinar las secuencias de la primera y segunda regiones del polinucleótido diana logra la secuenciación por pares de dichas primera y segunda regiones de dicho polinucleótido de doble hebra diana.

[0017] Además abarcadas dentro de las realizaciones de la invención están matrices agrupadas preparadas de acuerdo con el procedimiento tal como se define en las reivindicaciones.

[0018] Se describe adicionalmente en la presente memoria un procedimiento para mejorar la calidad de los datos de una reacción de secuenciación en un molde inmovilizado, comprendiendo el procedimiento hibridar el molde con un

cebador inmovilizado, de tal manera que el molde se inmoviliza a través de ambos extremos.

Breve descripción de las figuras

5 [0019]

La **Figura 1** muestra un esquema de lectura de extremos emparejados usando un primer procedimiento de la invención mediante una enzima de corte.

10 La **Figura 2** muestra un esquema de lectura emparejada usando un primer procedimiento de la invención mediante un cebador uracilo.

La **Figura 3** muestra un esquema de lectura emparejada usando un segundo procedimiento de la invención.

15 La **Figura 4a** muestra un esquema de lectura emparejada usando un tercer procedimiento de la invención. La figura muestra la primera lectura de secuenciación que se produce a partir de un cebador hibridado, aunque la hebra también puede estar hibridada con el cebador bloqueado con el fosfato bloqueado en la superficie, y por lo tanto se puede inmovilizar a través de ambos extremos (como se muestra en la Figura 4c).

20 La **Figura 4b** muestra un esquema de lectura emparejada usando el tercer procedimiento de la invención mediante la utilización de tres cebadores de injerto.

La **Figura 4c** muestra el procedimiento de la Figura 4a, en donde el molde se inmoviliza a través de ambos extremos durante la lectura SBS 1.

25

La **Figura 5** muestra una representación de los agregados como se cultivan en la superficie.

La **Figura 6** muestra la secuencia completa de una hebra del molde usada para la amplificación.

30 La **Figura 7** muestra datos de dos ciclos de incorporación de nucleótidos en el cebador de secuenciación inmovilizado.

35 La **Figura 8** muestra los datos de las lecturas de secuenciación obtenidas a partir del procedimiento mostrado en la Figura 4b. La muestra utilizada fue un BAC fragmentado de 140 KB fragmentado a un tamaño de inserción promedio de 80 pares de bases. Las dos lecturas se obtuvieron a partir de cualquiera de los extremos del fragmento. Se muestran los datos numéricos de un solo bloque de un análisis de secuenciación. Ambas lecturas se alinean claramente contra de la secuencia de BAC, con sólo el 4% de la lectura derivada de *E. coli* que contaminaba la muestra original.

40 La **Figura 9** muestra un esquema de un procedimiento que utiliza dos enzimas de corte y la resíntesis de la hebra.

La **Figura 10** muestra la etapa opcional de extensión del cebador antes de la resíntesis de la hebra, lo que mejora la etapa de la resíntesis de la hebra.

45

La **Figura 11** muestra un procedimiento alternativo para bloquear reversiblemente el cebador inmovilizado.

La **Figura 12** muestra un esquema de lecturas por pares de indexación, en el que los moldes de muestra se prepararon utilizando una secuencia o marcador de indexación.

50

Descripción detallada de la invención

[0020] La invención proporciona un procedimiento para la secuenciación de dos regiones de un molde de polinucleótido de doble hebra diáfila, al que se hace referencia en la presente memoria como las primera y segunda regiones para determinación de la secuencia. Las primera y segunda regiones para determinación de la secuencia se encuentran en ambos extremos de las hebras complementarias del molde de polinucleótido de doble hebra, a las que se hace referencia en la presente memoria, respectivamente, como primera y segunda hebras del molde. Una vez que se conoce la secuencia de una hebra, también se conoce la secuencia de su hebra complementaria, por lo tanto, el término dos regiones pueden aplicarse por igual a ambos extremos de un molde de hebra sencilla, o ambos extremos de un molde de hebra doble, en el que una primera región y su complemento son conocidos, y se conocen una segunda región y su complemento.

[0021] El punto de partida para el procedimiento de la invención es la provisión de una pluralidad de dúplex de polinucleótidos molde inmovilizados sobre un soporte sólido. Los polinucleótidos del molde se pueden inmovilizar en forma de una matriz de moléculas de molde individuales amplificadas o "agregados". Cada uno de los dúplex dentro

de un agregado particular, comprende la misma región diana de doble hebra a secuenciar. Los dúplex están formados cada uno de una primera y segunda hebras del molde complementaria que están unidos al soporte sólido en o cerca de sus extremos 5'. Habitualmente, se proporcionarán los dúplex de polinucleótidos molde en forma de una matriz agrupada.

5

[0022] Un punto de partida alternativo es una pluralidad de moldes de hebra sencilla que se unen a la misma superficie que una pluralidad de cebadores que son complementarios al extremo 3' del molde inmovilizado. Los cebadores pueden ser bloqueados reversiblemente para evitar la extensión. Los moldes de hebra sencilla pueden ser secuenciados usando un cebador hibridado con el extremo 3'. El cebador de secuenciación puede ser retirado después de la secuenciación y los cebadores inmovilizados desbloqueados para liberar un hidroxilo en 3' extensible. Estos cebadores se pueden usar para copiar el molde usando resíntesis de hebra con puente para producir un segundo molde inmovilizado que es complementario con el primero. La eliminación del primer molde de la superficie permite secuenciar el segundo molde de hebra sencilla reciente, de nuevo desde el extremo 3'. Por lo tanto, se pueden secuenciar ambos extremos del molde inmovilizado original. Dicha técnica permite lecturas de extremos emparejados donde los moldes se amplifican usando un único cebador inmovilizado extensible, por ejemplo como se describe en Polony technology (Nucleic Acids Research 27, 24, e34 (1999)) o PCR en emulsión (Science 309, 5741, 1728-1732 (2005); Nature 437, 376-380 (2005)). Los detalles descritos en la presente memoria a continuación en relación con las etapas individuales aplican haciendo los ajustes necesarios a las etapas descritas anteriormente, por ejemplo, los cebadores pueden ser bloqueados usando las mismas técnicas como se describe en las secciones pertinentes a continuación.

[0023] Si la amplificación se realiza en perlas, ya sea con un único o múltiples cebadores extensibles, las perlas pueden ser analizadas en solución, en pocillos individuales de una placa de microtitulación o picotitulación, inmovilizadas en pocillos individuales, por ejemplo, en un dispositivo de tipo de fibra óptica, o como una matriz inmovilizada sobre un soporte sólido. El soporte sólido puede ser una superficie plana, por ejemplo un portaobjetos de microscopio, en el que las perlas se depositan al azar y se mantienen en su lugar con una película de polímero, por ejemplo agarosa o acrilamida.

[0024] Cuando se hace referencia a la inmovilización o unión de moléculas (por ejemplo, ácidos nucleicos) a un soporte sólido, los términos "inmovilizados" y "unidos" se utilizan forma intercambiable en la presente memoria y ambos términos pretenden incluir la unión directa o indirecta, covalente o no covalente, a menos indique lo contrario, de forma explícita o por el contexto. En determinadas realizaciones de la invención, se puede preferir la unión covalente, pero en general todo lo que se requiere es que las moléculas (por ejemplo, ácidos nucleicos) permanezcan inmovilizadas o unidos al soporte en las condiciones en las que se haya previsto utilizar el soporte, por ejemplo, en aplicaciones que requieren la amplificación y/o secuenciación de ácidos nucleicos.

[0025] Determinadas realizaciones de la invención pueden utilizar soportes sólidos que comprenden un sustrato inerte o matriz (por ejemplo, portaobjetos de vidrio, perlas de polímero, etc.) que se han "funcionalizado", por ejemplo, mediante la aplicación de una capa o recubrimiento de un material intermedio que comprende grupos reactivos que permiten la unión covalente de biomoléculas, tales como polinucleótidos. Ejemplos de tales soportes incluyen, pero no se limitan a, hidrogeles de poliácridamida soportados sobre un sustrato inerte tal como vidrio. En dichas realizaciones, las biomoléculas (por ejemplo, polinucleótidos) pueden estar directamente unidas covalentemente al material intermedio (por ejemplo, el hidrogel), pero el propio material intermedio puede estar unido no covalentemente al sustrato o matriz (por ejemplo, el sustrato de vidrio). La expresión "unión covalente a un soporte sólido" ha de interpretarse en consecuencia como que abarca este tipo de disposición.

[0026] Con el fin de secuenciar dos regiones de un polinucleótido de doble hebra diana dado usando los procedimientos de la invención, los agregados pueden manipularse, ya sea para escindir la región del ADN diana entre los dos extremos, o para obtener datos de la secuencia de ambas hebras del dúplex. Si la porción central del agregado se escinde, y se unen los extremos, es posible leer la secuencia contigua acortada del agregado en una sola reacción de secuenciación que lee a lo largo de ambos extremos del molde original. Esto es similar a la idea de la clonación basada en vector (ditag) en solución, pero utilizando el agregado en la superficie, que comprende una única secuencia para asegurar que los dos extremos del fragmento original se unen. En el caso de vectores en solución, los fragmentos de ADN se ligan en constructos circulares y, por lo tanto, los dos extremos (o marcadores) de los fragmentos permanecen en la misma molécula, incluso cuando el círculo se abre. Esto garantiza que los dos extremos de un mismo fragmento original se unen de nuevo (para formar ditags), sin reordenamiento inter-molecular. En el caso de agregados, el hecho de que los extremos de los fragmentos de dúplex permanezcan unidos a la superficie garantiza que se unan los extremos del fragmento original, lo que también impide que se mezclen secuencias de diferentes moléculas, o agregados. Ver en la Figura 3 un esquema de secuenciación por pares que requiere la ligación de extremos acortados. Al inicio del procedimiento, los dos extremos A y B están unidos covalentemente a través de una serie de bases intermedias, por ejemplo 100-1000 bases. Una primera etapa de tratamiento corta las hebras del agregado para dejar sólo los extremos A y B como dúplex cortos independientes que ya no están conectados de forma covalente. Una sola hebra de cada dúplex está unida a la superficie y la otra hebra hibridada. Si los extremos de los dúplex se hacen romos y se ligan de nuevo, los extremos AB (y AA y BB) se vuelven a unir covalentemente sin las bases intermedias.

[0027] Procedimientos de tratamiento de los dúplex para cortar las hebras en posiciones específicas, pero independientes de la secuencia, incluyen el uso de una enzima de restricción. Las enzimas de restricción adecuadas son bien conocidas en la técnica anterior, por ejemplo, Mme1, que corta 18 y 20 bases de forma remota a su sitio de unión. Es deseable que la enzima de restricción llegue tan lejos como sea posible en la región desconocida del ADN, 5 prefiriéndose una enzima que corta entre 10-50 bases, o más. En una realización particular, se considera una enzima que corta entre 15-30 bases remotas de su sitio de unión, tal como EcoP15.

[0028] Los sitios de reconocimiento para la enzima de restricción particular utilizada pueden ser diseñados en el cebador utilizado para la amplificación en fase sólida y unirse a los extremos de los fragmentos diana en la etapa de 10 preparación de la muestra. La unión de extremos conocidos universales a una biblioteca de fragmentos de DNA por ligación permite la amplificación de una gran variedad de diferentes secuencias en una sola reacción de amplificación. Las secuencias de la porción de la secuencia conocida del molde de ácido nucleico pueden estar diseñadas de forma que las enzimas de restricción de tipo 2s se unan a la región conocida y corten en la región desconocida del molde amplificado.

15 **[0029]** El tratamiento de los dúplex donde ambas hebras se inmovilizan con la o las enzimas de restricción dará lugar a dos fragmentos de dúplex acortados en cada agregado, estando cada dúplex anclado a la superficie por uno de sus extremos 5'. El tratamiento con enzimas de restricción habitualmente deja un saliente en una de las hebras. El saliente de hebra sencilla se puede llenar o eliminar con una polimerasa o exonucleasa, respectivamente, para 20 asegurar los extremos del dúplex son romos.

[0030] Los fragmentos del dúplex con extremos romos se pueden unir a continuación mediante ligación para producir dúplex donde ambos extremos 5' están inmovilizados. Tales dúplex acortados contienen las secuencias derivadas de los dos extremos del fragmento original, pero sin ningún tipo de bases intermedias. Un solo análisis de 25 secuenciación de, por ejemplo, 50 bases, podría obtener la secuencia de 25 bases de cada extremo del fragmento original.

[0031] Para facilitar la secuenciación, es preferible que una de las hebras se retire de la superficie para permitir la hibridación eficaz de un cebador de secuenciación con la hebra restante inmovilizada. A continuación se describen 30 procedimientos adecuados para la linealización y se describen en más detalle en el documento Número de solicitud WO07010251.

[0032] La linealización también sirve para limitar el problema de la ligación de dos fragmentos del mismo extremo en lugar de extremos opuestos. Para cada fragmento con extremos A y B, los "muñones" A y B se quedan en la 35 superficie después de la etapa de escisión. No hay nada que impida una ligadura de A con A o la ligadura de B con B para dar dos copias repetidas de uno de los extremos, en lugar de la copia deseada de dos extremos. La linealización de una de las hebras ayuda a evitar este problema, ya que para las tres posibilidades A-A, A-B o B-B, a un dúplex le faltarán ambas y a uno no le faltarán ninguna y, por lo tanto, permanece como un dúplex que no está disponible para la hibridación. Procedimientos para la linealización y la secuenciación de los agregados se describen 40 en su totalidad a continuación.

[0033] En el procedimiento de la invención, los datos de la secuencia se pueden obtener de ambos extremos del dúplex inmovilizado por un procedimiento en el que el dúplex se trata para liberar un resto hidroxilo en 3' que se puede utilizar como cebador de extensión. El cebador de extensión se puede utilizar a continuación para leer la 45 primera secuencia de una hebra del molde. Después de la primera lectura, la hebra se puede extender hasta copiar completamente todas las bases hasta el final de la primera hebra. Esta segunda copia se queda unida a la superficie en el extremo 5'. Si la primera hebra se retira de la superficie, la secuencia de la segunda hebra se puede leer. Esto da una lectura de la secuencia de los dos extremos del fragmento original.

50 **[0034]** En el tercer procedimiento, los datos de la secuencia se pueden obtener de ambos extremos de un dúplex molde obteniendo una lectura de secuencia de una hebra del molde a partir de un cebador en solución, copiando la hebra utilizando cebadores inmovilizados, liberando la primera hebra y secuenciando la segunda hebra copiada.

[0035] Los procedimientos para generar un hidroxilo libre en 3' en una sola hebra de un dúplex inmovilizado en 55 ambos extremos incluyen cualquier tratamiento con una enzima de corte, o tratamiento químico para eliminar un nucleótido específico. Las enzimas de corte adecuadas son bien conocidas en la técnica, y cortarían preferiblemente en un sitio que es 3'-remoto respecto a su sitio de unión para evitar tener que secuenciar las bases que se derivan desde el sitio cortado conocido. La enzima de corte debe cortar sólo una de las hebras inmovilizadas, en el extremo más cercano a la superficie. Véase en la figura 1 un esquema de una metodología de secuenciación de extremos 60 emparejados utilizando una enzima de corte. Ejemplos de enzimas de restricción adecuadas incluirían Nt.BstNBI y Nt.AlwI, que no tienen bases de la secuencia definida más allá del hidroxilo en 3' liberado.

[0036] Procedimientos de eliminación de nucleótidos del dúplex implican el diseño del molde de tal manera que la base inmediatamente adyacente a la región diana desconocida puede ser eliminada para conseguir un sitio abásico. 65 Un "sitio abásico" se define como una posición de nucleótido en una hebra de polinucleótidos de la que el

componente de base se ha eliminado. Sitios abásicos pueden ocurrir de forma natural en el ADN en condiciones fisiológicas mediante la hidrólisis de residuos de nucleótidos, pero también pueden formarse químicamente en condiciones artificiales o por la acción de las enzimas. Una vez formados, los sitios abásicos pueden ser escindidos (por ejemplo, mediante tratamiento con una endonucleasa u otra enzima de escisión de hebra sencilla, exposición al calor o álcali), proporcionando un medio para la escisión específica del sitio de una hebra de polinucleótidos.

[0037] En una realización no limitativa un sitio abásico se puede crear en una posición predeterminada en una hebra de un dúplex de polinucleótido molde y luego escindirse incorporando primero desoxiuridina (U) en un sitio de escisión pre-determinado en una hebra del dúplex de polinucleótido molde. Esto se puede lograr, por ejemplo, incluyendo U en uno de los cebadores utilizados para la preparación del dúplex polinucleótido molde por amplificación en fase sólida. La enzima uracilo ADN glicosilasa (UDG) se puede usar a continuación para eliminar la base uracilo, generando un sitio abásico en una hebra. La hebra de polinucleótido que incluye el sitio abásico puede escindirse a continuación en el sitio abásico por tratamiento con endonucleasa (por ejemplo, EndoIV endonucleasa, AP liasa, FPG glicosilasa/AP liasa, EndoVIIIglicosilasa/AP liasa), calor o álcalis. En una realización particular, se usa el reactivo USER™ disponible en New England Biolabs (M5505S) para la creación de una única brecha de nucleótido en una base uracilo en hebra del dúplex. El tratamiento con enzimas endonucleasa da lugar a un resto 3'-fosfato en el sitio de escisión, que se puede eliminar con una fosfatasa adecuada, tal como la fosfatasa alcalina. Ver en la Figura 2 una ilustración esquemática de una reacción de secuenciación de extremos emparejados que utiliza linealización resíntesis de linealización con USER.

[0038] Los sitios abásicos también pueden ser generados en desoxirribonucleótidos no naturales/modificados distintos de desoxiuridina y escindidos de manera análoga por tratamiento con endonucleasa, calor o álcali. Por ejemplo, la 8-oxo-guanina se puede convertir en un sitio abásico por la exposición a la FPG glicosilasa. La desoxiinosina se puede convertir en un sitio abásico por la exposición a AlkA glicosilasa. Los sitios abásicos generados de esta manera pueden escindirse después, habitualmente mediante tratamiento con una endonucleasa adecuada (por ejemplo, EndoIV, AP liasa). Un ejemplo adicional incluye el uso de una citosina no metilada específica. Si el resto de las secuencias del cebador y los dNTP utilizados en la formación de un agregado son residuos de citosina metilados, entonces, las citosinas no metiladas se pueden convertir específicamente en residuos de uracilo mediante tratamiento con bisulfito. Esto permite el uso del mismo tratamiento con 'USER' para linealizar ambas hebras del agregado, ya que uno de los cebadores puede contener un uracilo y uno puede contener la citosina que se puede convertir en un uracilo (efectivamente una especie 'uracilo protegido'). Si el nucleótido no natural/modificado a ser incorporado en un cebador de amplificación para su uso en la amplificación en fase sólida, entonces, el nucleótido no natural/modificado debe ser capaz de ser copiado por la polimerasa utilizada para la reacción de amplificación.

[0039] En una realización, las moléculas a ser escindidas pueden exponerse a una mezcla que contiene el glicosilado apropiado y una o más endonucleasas adecuadas. En tales mezclas, el glicosilado y la endonucleasa estarán presentes habitualmente en una relación de actividad de al menos aproximadamente 2:1.

[0040] Este procedimiento de escisión tiene ventajas particulares en relación con la creación de moldes para la secuenciación de ácidos nucleicos. En particular, la escisión de un sitio abásico generado por tratamiento con un reactivo, tal como USER libera automáticamente un grupo 3' fosfato libre en la hebra escindida que, después de tratamiento con fosfatasa puede proporcionar un punto de iniciación para la secuenciación de una región de la hebra complementaria. Por otra parte, si el ácido nucleico de doble hebra inicial contiene sólo una base escindible (por ejemplo, uracilo) en una hebra, entonces se puede generar una única "mella" en una única posición en esta hebra del dúplex. Dado que la reacción de escisión requiere un residuo, por ejemplo, desoxiuridina, que no existe de forma natural en el ADN, pero que es de otra manera independiente del contexto de la secuencia, si se incluye sólo una base no natural, no hay posibilidad de que la escisión mediada por la glicosilasa se produzca en otro lugar en posiciones no deseadas en el dúplex. Por el contrario, si el ácido nucleico de doble hebra tuviera que escindirse con una endonucleasa de "corte" que reconoce una secuencia específica, hay una posibilidad de que la enzima pueda crear mellas en "otros" sitios en el dúplex (además del sitio de escisión deseado) si éstos poseen la secuencia de reconocimiento correcta. Esto podría presentar un problema si se crean mellas en la hebra que se desea copiar, en lugar de la hebra que se elimina total o parcialmente y es un riesgo particular si la porción diana de la molécula de ácido nucleico de doble hebra tiene una secuencia desconocida. Estas limitaciones se pueden superar preparando una muestra con el sitio de reconocimiento para dos enzimas de corte diferentes. Si ambas preparaciones se procesan por separado usando el procedimiento, las regiones del genoma escindidas por una enzima de corte deben estar representadas en la otra preparación.

[0041] El hecho de que no haya ningún requisito para que el residuo no natural (por ejemplo, uracilo) se sitúe en un contexto de secuencia detallado con el fin de proporcionar un sitio para la escisión usando este enfoque es ventajoso en sí. En particular, si el sitio de escisión se va a incorporar en un cebador de amplificación para ser usado en la producción de una matriz agrupada por amplificación en fase sólida, sólo es necesario sustituir un nucleótido natural (por ejemplo, T) en el cebador con un nucleótido no natural (por ejemplo, U) con el fin de permitir la escisión. No hay necesidad de diseñar el cebador para incluir una secuencia de reconocimiento de una enzima de restricción de varios nucleótidos de longitud. Los cebadores de oligonucleótidos incluyendo nucleótidos U y otros nucleótidos no

naturales, tales como los enumerados anteriormente, se pueden preparar fácilmente usando técnicas y aparatos convencionales para la síntesis química de oligonucleótidos.

5 **[0042]** Otra ventaja obtenida mediante la escisión de sitios abásicos en una molécula de doble hebra generada por la acción de UDG en uracilo es que la primera base incorporada en una reacción de "secuenciación por síntesis" que se inicia en el grupo hidroxilo en 3' libre formado por la escisión en tal sitio siempre será T. Por lo tanto, si el dúplex polinucleótido molde forma parte de una matriz agrupada compuesto de muchas de tales moléculas, todas las cuales se escinden de esta manera para producir moldes de secuenciación, entonces la primera base universalmente incorporado en toda la matriz será T. Esto puede proporcionar un ensayo de secuenciación independiente para la intensidad del agregado individual al inicio de un "ciclo" de secuenciación.

15 **[0043]** En casos particulares, puede ser ventajoso llevar a cabo un tratamiento de bloqueo con un didesoxinucleótido trifosfato y una polimerasa y/o transferasa terminal. Después de una amplificación en fase sólida, quedan en la superficie un gran número de cebadores de amplificación no usados, además de los extremos 3' libres de cada una de las hebras del molde. El tratamiento con un nucleótido bloqueado de este tipo asegura que estos grupos funcionales 3'-OH libres no estén disponibles para la extensión durante las etapas subsiguientes de la polimerasa. En los procedimientos en los que se crean grupos hidroxilo en 3' con el tratamiento químico o endonucleasas de restricción, es ventajoso bloquear cualquier grupo hidroxilo en 3' residual antes de que se creen los grupos hidroxilo en 3' deseados. En el caso de tratamiento con el reactivo USER, dado que se libera un grupo fosfato, la etapa de bloqueo se puede realizar ya sea antes o después del tratamiento con USER, ya que el grupo fosfato actuará como un grupo protector para evitar el bloqueo de los restos hidroxilo en 3' deseados. El grupo fosfato se puede eliminar después de que se ha realizado la etapa de bloqueo.

25 **[0044]** El acto de tratamiento con enzimas de restricción o de generación de un sitio abásico y escisión da lugar a un extremo 5' libre en la hebra que ya no está inmovilizada sobre la superficie. Esta hebra puede eliminarse completamente de la superficie por el tratamiento con una 5'-3' exonucleasa, tales como la exonucleasa lambda o T7. Este tratamiento hace que la hebra molde sea de hebra sencilla y esté disponible para la posterior copia. Si la polimerasa que se usa para extender el grupo hidroxilo en 3' tiene una actividad de desplazamiento de la hebra, entonces, puede no ser necesario el tratamiento de la 5'-3' exonucleasa.

30 **[0045]** En la realización de la invención, la utilización de un cebador inmovilizado para la secuenciación, es ventajoso para extender el cebador hidroxilo libre en 3' con una pluralidad de bases complementarios con el molde antes de iniciar la secuenciación. Esto aumenta la temperatura de fusión del dúplex inmovilizado y ayuda a prevenir la hebra del molde rehibride con otros cebadores inmovilizados durante la secuenciación, lo que daría lugar a problemas de fase dentro del grupo. La etapa de ligación se lleva a cabo después de que la etapa de la fosfatasa haya eliminado el grupo fosfato del cebador inmovilizado. La adición de 20-30 bases de la secuencia puede ser realizada por una reacción de ligación con un cebador modificado en 5'-fosfato hibridado adyacente al hidroxilo libre en 3'. Una ligasa, tal como la ADN ligasa de T4 se puede utilizar para sellar la brecha. En el caso del tratamiento con USER que elimina el nucleótido U, la base en 5' del cebador será la T que sustituirá al UR escindido. Para que la etapa de hibridación se lleve a cabo de manera eficiente, la hebra no inmovilizada en 5' debe de eliminarse por tratamiento de exonucleólisis 5'-3' como se describió anteriormente. Dichos cebadores extendidos inmovilizados con un hidroxilo libre en 3' se describen como cebadores anclados al 5' extendidos o como cebadores inmovilizados extendidos y la generación de dichos cebadores extendidos es sólo una de las etapas involucradas en el tratamiento de la pluralidad de polinucleótidos del molde de doble hebra, de modo que la primera hebra del molde se hibrida a un cebador que está inmovilizado en el soporte sólido en su extremo 5'.

45 **[0046]** El grupo hidroxilo libre en 3' en el cebador anclado en 5' o cebador extendido se puede utilizar para iniciar rondas de secuenciación para determinar la secuencia de las bases en el molde hibridado. La secuenciación se puede llevar a cabo usando cualquier técnica adecuada "secuenciación por síntesis", en la que se añaden nucleótidos sucesivamente al grupo hidroxilo libre en 3', lo que tiene como resultado la síntesis de una hebra de polinucleótidos en la dirección 5' a 3'. La naturaleza del nucleótido añadido se determina preferiblemente después de cada adición. Los procedimientos alternativos de secuenciación incluyen la secuenciación mediante ligación, por ejemplo como se describe en los documentos US6306597o WO06084132.

55 **[0047]** El uso de USER de acuerdo con la presente invención produce una primera hebra del molde linealizado y un césped de cebadores que quedan con restos de bloqueo de fosfato. Sin tratamiento posterior con fosfatasa, el césped de cebador inmovilizado no es adecuado para la extensión de la polimerasa. Se puede utilizar un cebador de secuenciación en solución para iniciar una secuencia de lectura para la primera hebra del molde. Si los grupos fosfato se eliminan del césped de cebadores, entonces las primeras hebras del molde, una vez eliminado el primer cebador de secuenciación extendido, se pueden hibridar con el primer césped imprimación y someter a uno o más ciclos de extensión, desnaturalización y rehibridación. La desnaturalización puede llevarse a cabo térmicamente o isotérmicamente, por ejemplo, utilizando desnaturalización química. El desnaturalizante químico puede ser urea, hidróxido o formamida u otro reactivo similar. Esto regenera los dúplex de molde donde ambas hebras están inmovilizadas. La primera hebra puede eliminarse mediante una etapa de tratamiento de linealización ortogonal adecuada, tal como la escisión del diol o la eliminación de un residuo de 8-oxo-G y después la desnaturalización de

la primera hebra del molde, un segundo cebador de secuenciación puede hibridar con la segunda hebra del molde y secuenciar la segunda hebra del molde. Esta estrategia de linealización ortogonal permite leer desde ambos extremos del molde.

5 **[0048]** El primer procedimiento de la invención proporciona una reacción de secuenciación llevada a cabo utilizando un cebador inmovilizado. El cebador de secuenciación inmovilizada puede ser desnaturalizado y la primera hebra del molde utilizada para volver a iniciar nuevas rondas de ampliación o amplificación como antes. De nuevo, esto genera los dúplex de molde donde se inmovilizan ambas hebras. La primera hebra puede ser eliminada por una etapa de tratamiento de linealización ortogonal adecuada, tal como la escisión de un diol o la eliminación de un residuo de 8-oxo-G y después la desnaturalización de la primera hebra del molde, un segundo cebador de secuenciación puede hibridar con la segunda hebra del molde y secuenciar la segunda hebra del molde. Esta estrategia de linealización ortogonal permite leer desde ambos extremos del molde.

15 **[0049]** Ambos procedimientos detallados anteriormente permiten la repoblación del agregado entre la primera y segunda lecturas. Esto es beneficioso para aumentar el nivel de la señal obtenida para la segunda lectura del agregado. Cualquier procedimiento que tenga como resultado que ambos cebadores de amplificación queden retenidos en la superficie durante la primera lectura secuenciación está dentro del alcance del tercer procedimiento de la invención, aunque es ventajoso si los cebadores se pueden tratar, como en el caso de USER, para hacer que el hidroxilo en 3' no esté disponible para la extensión del cebador durante la reacción de secuenciación. Un procedimiento alternativo de bloqueo de los cebadores de injerto no extendidos sería el uso de un nucleótido portador de un grupo 3' no extendible. El tratamiento de la superficie con un nucleótido de este tipo y transferasa terminal bloquearía la superficie para una extensión adicional. Después del primer ciclo de secuenciación, los nucleótidos pueden ser desprotegidos para permitir ciclos adicionales de la copia de la hebra molde. Si los nucleótidos llevan una modificación didesoxi, esta se puede eliminar mediante una exonucleasa, o una polimerasa con actividad exonucleasa. Por lo tanto, los agregados podrían construirse con dos cebadores injertados como se describe, los cebadores no utilizados bloqueados durante la primera reacción de secuenciación, entonces desbloqueados para permitir la copia posterior de las hebras del molde.

30 **[0050]** Una parte de los cebadores de amplificación se puede unir a la superficie con una modificación bloqueando el hidroxilo en 3' de la extensión en los ciclos de amplificación. Esto en efecto significa que la superficie se trata con tres o más cebadores de amplificación en lugar de dos. Al menos dos de los cebadores de amplificación deben comprender regiones de secuencia idéntica, pero al menos un cebador no será susceptible a las condiciones utilizadas para eliminar el segundo cebador durante el proceso de linealización y contendrá un resto bloqueante del 3'. El resto bloqueante puede adoptar la forma de un bloque químico, tal como un grupo azidometilo que se puede eliminar con un reactivo de fosfina, un grupo enzimáticamente eliminable, tal como un grupo fosfato que se puede eliminar con una fosfatasa o puede estar en la forma de un grupo nucleósido que se puede eliminar usando exonucleólisis 3'-5'. Tales modificaciones de nucleósidos incluyen sitios abásicos, que se pueden eliminar, como se ha descrito o nucleótidos didesoxi 2', 3' que se pueden eliminar mediante una polimerasa con actividad exonucleasa. Otras modificaciones incluyen el uso de una secuencia de oligonucleótidos que pueden formar una región autocomplementaria con una secuencia de reconocimiento para una enzima de restricción, como se muestra en la figura 11. El tratamiento con la enzima de restricción debe cortar la hebra horquilla y liberar una secuencia más corta con un grupo hidroxilo en 3'. El tratamiento de la superficie después de finalizado el primer ciclo de secuenciación para desbloquear los cebadores permitirá que la primera hebra restante hibride con los cebadores desprotegidos y vuelva a copiar las hebras ya secuenciadas. Una segunda lectura de secuenciación puede obtenerse eliminando la primera hebra, hibridando un cebador de secuenciación y leyendo la segunda hebra recién sintetizada.

50 **[0051]** Un procedimiento de secuenciación particular, que se puede utilizar en los procedimientos de la invención se basa en el uso de nucleótidos modificados que pueden actuar como terminadores de hebra reversibles. Dichos terminadores de hebra reversibles comprenden extraíbles grupos de bloqueo 3' extraíbles. Una vez que un nucleótido modificado así se ha incorporado en la hebra en crecimiento del polinucleótido complementario a la región del molde que se está secuenciado, no existe ningún grupo 3'-OH libre disponible para dirigir más la extensión de la secuencia y, por lo tanto, la polimerasa no puede añadir más nucleótidos. Una vez que se ha determinado la naturaleza de la base incorporada en la hebra en crecimiento, el bloque de 3' puede ser eliminado para permitir la adición del siguiente nucleótido sucesivo. Al ordenar los productos derivados que utilizan estos nucleótidos modificados, es posible deducir la secuencia de ADN del molde de ADN. Tales reacciones pueden realizarse en un solo experimento si cada uno de los nucleótidos modificados ha unido a la misma un marcador diferente, conocida por corresponder a la base en particular, para facilitar la discriminación entre las bases añadidas en cada etapa de incorporación. Los marcadores adecuados se describen en la solicitud PCT copendiente PCT/GB/2007/001770.

60 **[0052]** Alternativamente, se puede llevar a cabo una reacción separada que contiene cada uno de los nucleótidos modificados añadidos individualmente.

65 **[0053]** Los nucleótidos modificados pueden llevar un marcador para facilitar su detección. En una realización particular, el marcador es un marcador fluorescente. Cada tipo de nucleótidos puede llevar un marcador fluorescente

diferente. Sin embargo, el marcador detectable no necesita ser un marcador fluorescente. Cualquier marcador se puede utilizar que permite la detección de la incorporación de los nucleótidos en la secuencia de ADN.

- 5 **[0054]** Un procedimiento para la detección de nucleótidos marcados con fluorescencia comprende el uso de luz láser de una longitud de onda específica para los nucleótidos marcados o el uso de otras fuentes adecuadas de iluminación. La fluorescencia del marcador en un nucleótido incorporado puede ser detectada por una cámara CCD o otros medios de detección adecuados. Medios de detección adecuados se describen en la solicitud copendiente PCT/US2007/007991.
- 10 **[0055]** Una vez que la primera lectura de la secuencia ha finalizado y se ha determinado una longitud de lectura suficiente de leer, se puede copiar el resto de la hebra. Si el grupo hidroxilo en 3' se creó originalmente con una enzima de corte, entonces será posible volver a crear un grupo hidroxilo en 3' nuevo en la misma posición y extenderse desde esta posición, sin embargo, es igualmente posible seguir copiando la primera hebra del molde desde el grupo hidroxilo en 3' de los nucleótidos incorporados como parte de la reacción de secuenciación.
- 15 **[0055]** Una reacción de extensión con los cuatro nucleótidos no marcados y una polimerasa copiará todas las bases del primer molde. Los cebadores inmovilizados pueden contener un sitio de restricción para una enzima de corte y el tratamiento con la enzima de restricción pueden acortar los cebadores inmovilizados, o los dúplex inmovilizados del molde para liberar un grupo hidroxilo en 3', como se muestra en la Figura 9.
- 20 **[0056]** En el curso del desarrollo de los protocolos incluidos en la presente memoria, se observó sorprendentemente que los resultados de la secuenciación mejoraban en las situaciones en las que el extremo 5' del molde era capaz de hibridación con un cebador inmovilizado. En resumen, la secuenciación como se muestra en la figura 4c era mejor que la secuenciación como se muestra en la figura 4a. En los resultados experimentales mostrados en esta memoria, mejor significa que el nivel de señal retenida a través de múltiples ciclos era mayor, lo que significa que la
- 25 **[0056]** curva de extinción con ambos extremos inmovilizados era menos pronunciada y la señal más alta que la observada cuando sólo se inmoviliza un solo extremo. Esto puede ser debido al hecho de que cualquier rotura de hebra del molde que se produzca durante el proceso de secuenciación no tiene como resultado automáticamente la pérdida del molde de la superficie, ya que el extremo 3' del molde también se mantiene en su lugar por hibridación, aunque pueden existir también otros factores que causan este efecto.
- 30 **[0057]** La primera hebra del molde se puede unir a la superficie de una manera que permite la eliminación selectiva. Si la primera hebra del molde se elimina de la superficie y las hebras del dúplex se desnaturalizan, por ejemplo por tratamiento con hidróxido o formamida, entonces, la segunda hebra copiada permanece inmovilizada como una sola hebra linealizada. Como se conocen las secuencias de los extremos de esta hebra, es posible hibridar un cebador
- 35 **[0057]** de secuenciación con la segunda hebra del molde y repitiendo los ciclos de secuenciación como se ha descrito anteriormente, determinar la secuencia de una segunda lectura para el otro extremo del molde de la primera lectura.
- [0058]** La eliminación selectiva, o linealización de la primera hebra del molde se puede lograr en un número de distintas maneras. La linealización para permitir la hibridación de un cebador de secuenciación en solución no tiene
- 40 **[0058]** que dejar un hidroxilo en 3' funcional en la hebra del molde y puede escindir o bien una hebra o ambas hebras. Por lo tanto, tal como se utiliza aquí, el término "linealización" se refiere a la eliminación selectiva de una hebra complementaria. Si uno de los cebadores de amplificación se inmoviliza de tal manera que se puede escindir de la superficie, el ADN de doble hebra resultante se puede convertir en ADN de hebra sencilla utilizando calor o condiciones químicas de desnaturalización para dar una molécula de hebra sencilla que contiene un sitio de
- 45 **[0058]** hibridación del cebador. La molécula de hebra sencilla puede hibridar con un cebador de secuenciación en solución para permitir una lectura de secuenciación de la hebra del molde inmovilizada. Dicho sitio de escisión es un sitio que permite la escisión controlada de la primera hebra del molde por medios químicos, enzimáticos o fotoquímicos.
- [0059]** Cualquiera, reacción de escisión enzimática, química o fotoquímica adecuada puede usarse para escindir. En
- 50 **[0059]** el documento WO07010251 se describen varios procedimientos adecuados.
- [0060]** La reacción de escisión puede tener como resultado la eliminación de una parte o la totalidad de la hebra a ser escindida. Medios de escisión adecuados incluyen, por ejemplo, digestión con enzimas de restricción, en cuyo caso el sitio de escisión es un sitio de restricción apropiado para la enzima que dirige la escisión de una o ambas
- 55 **[0060]** hebras de un molde dúplex; la digestión con ARNasa o la escisión química de un enlace entre un desoxirribonucleótido y un ribonucleótido, en cuyo caso el sitio de escisión puede incluir uno o más ribonucleótidos, la reducción química de un enlace disulfuro con un agente reductor (por ejemplo, TCEP), en cuyo caso el sitio de escisión debería incluir un enlace disulfuro apropiado; la escisión química de un conector de diol con peryodato, en cuyo caso el sitio de escisión debería incluir un conector de diol; la generación de un sitio abásico y posterior
- 60 **[0060]** hidrólisis, etc.
- [0061]** En una realización, la escisión puede ocurrir en un sitio de escisión en una o ambas hebras de un dúplex de polinucleótido molde que comprende uno o más o de cualquier combinación de nucleótidos no naturales, ribonucleótidos o modificaciones químicas de un no nucleótido.
- 65 **[0061]**

[0062] Técnicas de escisión adecuadas para su uso en el procedimiento de la invención incluyen, pero no se limitan a, las siguientes:

i) escisión química

5

[0063] El término "escisión química" abarca cualquier procedimiento que utiliza un ácido no nucleico y un reactivo químico no enzimático con el fin de promover/lograr la escisión de una o ambas hebras de un dúplex de polinucleótido molde. Si es necesario, una o ambas hebras del dúplex de polinucleótido del molde puede incluir uno o más restos químicos y/o nucleótidos no naturales y/o uniones de esqueleto no naturales con el fin de permitir una reacción de escisión química. En una realización particular, la modificación o modificaciones requeridas para permitir la escisión química se puede incorporar en un cebador de amplificación utilizado para formar el dúplex de polinucleótido molde por amplificación de ácidos nucleicos en fase sólida.

10

[0064] En una realización preferida, pero no limitante, una hebra del dúplex de polinucleótido molde (o el cebador de amplificación del que se deriva esta hebras si se ha formado por amplificación en fase sólida) puede incluir un conector de diol que permite la escisión por tratamiento con peryodato (por ejemplo, peryodato de sodio). Se apreciará que en el sitio de escisión se puede incluir más de un diol.

15

[0065] Unidades de conector de diol basados en la química de fosforamidita adecuados para su incorporación en las hebras de polinucleótidos se pueden adquirir en Fidelity Systems Inc. (Gaithersburg, MD, EE.UU.). En un polinucleótido se pueden incorporar una o más unidades de diol usando procedimientos convencionales para la síntesis química automatizada de ADN. Por lo tanto, los cebadores de oligonucleótidos que incluyen uno o más conectores diol se pueden preparar convenientemente por síntesis química.

20

[0066] Con el fin de posicionar el conector de diol a una distancia óptima del soporte sólido, se pueden incluir una o más moléculas separadoras entre el conector de diol y el sitio de unión al soporte sólido. La molécula separadora puede ser un resto químico no nucleótido. Unidades separadoras adecuadas basadas en la química de la fosforamidita para su uso junto con conectores diol también se suministran por Fidelity Systems, Inc. Un separador adecuado para su uso con conectores diol, es el separador designado arm 26, identificado en los ejemplos que se acompañan. Para permitir la fijación a un soporte sólido en el extremo 5' de la hebra de polinucleótido, el arm 26 puede ser modificado para incluir un grupo fosforotioato. El grupo fosforotioato puede unirse con facilidad durante la síntesis química de una hebra de "polinucleótido" incluyendo el separador y las unidades de diol.

25

30

[0067] Se pueden usar otras moléculas separadoras como una alternativa a arm 26. Por ejemplo, se puede incluir un tramo de nucleótidos "separadores" no objetivo. Por lo general, se pueden incluir de 1 a 20, más particularmente de 1 a 15 o de 1 a 10, e incluso más en particular 2, 3, 4, 5, 6, 7, 8, 9 o 10 nucleótidos separadores. En una realización particular, 10 nucleótidos separadores se colocan entre el punto de unión al soporte sólido y el conector de diol. En otra realización particular, se utilizan separadores poliT, aunque se pueden utilizar otros nucleótidos y combinaciones de los mismos. En otra realización particular, el cebador puede incluir nucleótidos separadores 10T.

35

40

[0068] El conector de diol se escinde por tratamiento con un "agente de escisión", que puede ser cualquier sustancia que promueve la escisión del diol. En una realización particular, el agente de escisión es peryodato, preferiblemente peryodato de sodio acuoso (NaIO_4). Después del tratamiento con el agente de escisión (por ejemplo, peryodato) para escindir el diol, el producto escindido se puede tratar con un "agente de protección terminal" con el fin de neutralizar las especies reactivas generadas en la reacción de escisión. Agentes de protección terminal adecuados para este propósito incluyen aminas, tales como etanolamina o propanolamina. Ventajosamente, el agente de protección terminal (por ejemplo, propanolamina) puede incluirse en una mezcla con el agente de escisión (por ejemplo, peryodato), de modo que las especies reactivas se protejan tan pronto como se formen.

45

[0069] La combinación de un enlace de diol y un agente de escisión (por ejemplo, peryodato) para lograr la escisión de al menos una hebra de un dúplex de polinucleótido molde se puede usar con ventaja particular para la linealización del dúplex de molde en hidrogeles de poli(acrilamida) de soporte sólido, porque el tratamiento con peryodato es compatible con la integridad de ácido nucleico y con las sustancias químicas de la superficie del hidrogel. La utilidad de los conectores diol/periyodato como un procedimiento de linealización no se limita, sin embargo, limitada a las superficies de hidrogel de poli(acrilamida), sino que también se extiende a la linealización, de dúplex inmovilizados sobre otros soportes y superficies sólidas, incluyendo soportes recubiertos con silanos funcionalizados (etc).

50

55

[0070] En una realización adicional, la hebra a ser escindida (o el cebador de amplificación del que se deriva esta hebra si se ha preparado por amplificación en fase sólida) puede incluir un grupo disulfuro que permite la escisión con un agente reductor químico, por ejemplo, clorhidrato de tris-(2-carboxietil)-fosfato (TCEP).

60

ii) escisión de sitios abásicos

[0071] El uso de los sitios abásicos se ha descrito anteriormente con el fin de generar un resto hidroxilo libre en 3'

65

que actúe como un cebador de secuenciación. Si ambos cebadores de amplificación se modifican de tal manera que se pueden escindir de forma secuencial, la segunda escisión se puede usar para escindir la primera hebra de la superficie. El primer (o segundo) cebador podría contener una base uracilo, que puede ser escindida por una enzima (UDG) y el segundo (o primer) cebador podría contener una base de 8-oxo-guanina que se puede escindir por una segunda enzima ortogonal enzima, la FPG glicosilasa. La segunda escisión del sitio abásico se podría utilizar para dejar un cebador de secuenciación unido a una superficie, de tal manera que se incorpore una base G como el primer ciclo de secuenciación, o las hebras dúplex escindidas se pueden desnaturalizar para permitir la hibridación de un cebador de secuenciación en solución.

10 iii) escisión de ribonucleótidos

[0072] La incorporación de uno o más ribonucleótidos en una hebra polinucleotídica que comprende de otro modo desoxirribonucleótidos (con o sin restos adicionales químicos no nucleótidos adicionales, bases no naturales o uniones de esqueleto no naturales) puede proporcionar un sitio para la escisión usando un agente químico capaz de escindir selectivamente el enlace fosfodiéster entre un desoxirribonucleótido y un ribonucleótido o usando una ribonucleasa (ARNasa). Por lo tanto, se pueden producir moldes de secuenciación por escisión de una hebra de un dúplex de polinucleótido molde en un sitio que contiene uno o más ribonucleótidos consecutivos utilizando un agente de escisión químico de este tipo o una ARNasa. En una realización particular, la hebra que se escinde contiene un único ribonucleótido, que proporciona un sitio de escisión química.

[0073] Agentes de escisión química adecuados capaces de escindir selectivamente el enlace fosfodiéster entre un desoxirribonucleótido y un ribonucleótido incluyen iones metálicos, por ejemplo iones de metales de tierras raras (especialmente La^{3+} , particularmente Tm^{3+} , Yb^{3+} o Lu^{3+} (Chen et al. Biotechniques. 2002, 32: 518-520; Komiyama et al. Chem. Commun. 1999, 1443-1451)), $\text{Fe}(3)$ o $\text{Cu}(3)$ o la exposición a un pH elevado, por ejemplo, el tratamiento con una base tal como hidróxido de sodio. Por "escisión selectiva del enlace fosfodiéster entre un desoxirribonucleótido y un ribonucleótido" se entiende que el agente de escisión química no es capaz de escindir el enlace fosfodiéster entre dos desoxirribonucleótidos en las mismas condiciones.

[0074] La composición de la base del o los ribonucleótidos habitualmente no es material, pero se puede seleccionar con el fin de optimizar la escisión química (o enzimática). A modo de ejemplo, se prefieren habitualmente rUMP o rCMP si la escisión se va a llevar a cabo por exposición a iones metálicos, especialmente iones de metales de tierras raras.

[0075] El o los ribonucleótidos habitualmente se incorporarán en una hebra de un dúplex de polinucleótido molde (o el cebador de amplificación del que se deriva esta hebra si se ha preparado por amplificación en fase sólida) y se puede situar en una región del dúplex que es de hebra única cuando las dos hebras complementarias del dúplex están hibridadas (es decir, en una porción saliente en 5'). Si el dúplex de polinucleótido molde dúplex se prepara mediante amplificación en fase sólida utilizando cebadores de amplificación directo e inverso, uno de los cuales contiene al menos un ribonucleótido, las enzimas ADN polimerasas convencionales utilizadas para la amplificación no son capaces de copiar moldes de ribonucleótidos. Por lo tanto, los productos de amplificación contendrán una región 5' saliente que comprende el o los ribonucleótidos y cualquier resto del cebador de amplificación en dirección 5' del o los ribonucleótidos.

[0076] El enlace fosfodiéster entre un ribonucleótido y un desoxirribonucleótido, o entre dos ribonucleótidos también pueden ser escindido por una ARNasa. Para este fin se puede utilizar cualquier ribonucleasa endolítica con una especificidad por el sustrato apropiada. Si el o los ribonucleótidos están presentes en una región que es de hebra sencilla cuando las dos hebras complementarias de la molécula de doble hebra están hibridadas (es decir, en una porción saliente en 5'), entonces, la ARNasa será una endonucleasa que tiene especificidad por hebras sencillas que contienen ribonucleótidos. Para la escisión con ribonucleasa se prefiere incluir dos o más ribonucleótidos consecutivos y preferiblemente de 2 a 10 o de 5 a 10 ribonucleótidos consecutivos. La secuencia precisa de los ribonucleótidos habitualmente no es importante, a menos que ciertas ARNasas tengan una especificidad por la escisión después de ciertos residuos. Las ARNasas adecuadas incluyen, por ejemplo, la ARNasa A, que escinde después de los residuos C y U. Por lo tanto, cuando se escinde con ARNasa el sitio de escisión debe incluir al menos un ribonucleótido que sea C o U.

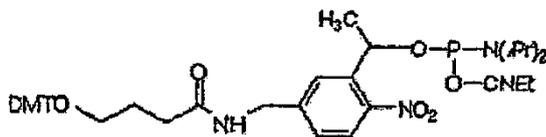
[0077] Los polinucleótidos que incorporan uno o más ribonucleótidos pueden sintetizarse fácilmente usando técnicas estándar para la síntesis química de oligonucleótidos con precursores de ribonucleótidos apropiados. Si el dúplex polinucleótido molde se prepara por amplificación de ácido nucleico en fase sólida, entonces es conveniente incorporar uno o más ribonucleótidos en uno de los cebadores para ser utilizados en la reacción de amplificación.

60 iv) escisión fotoquímica

[0078] El término "escisión fotoquímica" engloba cualquier procedimiento que utiliza la energía lumínica con objeto de conseguir la escisión de una o de ambas hebras de la molécula de ácido nucleico de doble hebra.

65

[0079] Puede proporcionarse un sitio predeterminado para la escisión fotoquímica mediante una unidad separadora química no nucleotídica en una de las hebras de la molécula de doble hebra. Algunos separadores fotoquímicos escindibles adecuados incluyen el separador PC de fosforamidita (4-(4,4'-dimetoxitritiloxi)butiramidometil)-1-(2-nitrofenil)-etil]-2-cianoetil-(N, N-diisopropil)-fosforamidita suministrado por Glen Research, Sterling, Virginia, EE.UU. 5 (número cat. 10-4913-XX), que tiene la estructura:



[0080] La unidad separadora puede ser escindida mediante exposición a una fuente de luz UV.

10 **[0081]** Esta unidad separadora puede estar unida al extremo 5' de un polinucleótido, junto con un grupo de tiofosfato que permite la unión a una superficie sólida, usando técnicas estándar para la síntesis química de oligonucleótidos. Convenientemente, esta unidad separadora puede incorporarse en un cebador de amplificación directo o inverso para usarse en la síntesis de un dúplex polinucleotídico molde fotoescindible mediante amplificación en fase sólida.

15 v) finalizadores de PCR

[0082] En otra forma de realización de la invención, el dúplex de polinucleótido molde puede prepararse mediante amplificación en fase sólida usando cebadores directos e inversos, uno de los cuales contiene un "finalizador de PCR". Un "finalizador de PCR" es cualquier fracción (nucleótido o no nucleótido) que evita la ultralectura por parte de la polimerasa usada para la amplificación, de forma que no puede copiar más allá de ese punto. El resultado es que las hebras amplificadas derivadas por la extensión del cebador que contiene el finalizador de PCR contendrán una porción saliente en 5'. Este saliente en 5' (distinto al propio finalizador de PCR) puede estar formado por desoxirribonucleótidos naturales, predominantemente con uniones de esqueleto naturales, es decir, simplemente puede ser una ampliación de un ADN monohebra. A continuación, la molécula puede ser escindida en la región saliente en 5' mediante el uso de un reactivo de escisión (por ejemplo, una enzima) que sea selectiva para la escisión del ADN monohebra pero no del ADN de doble hebra, por ejemplo nucleasa de judía mungo.

[0083] El finalizador de PCR puede ser esencialmente cualquier fracción que evite la ultralectura de la polimerasa que se va a usar para la reacción de amplificación. Algunos finalizadores de PCR adecuados incluyen, pero no se limitan a, hexaetilenglicol (HEG), sitios abásicos y cualquier nucleótido no natural o modificado que evite la ultralectura de la polimerasa, incluyendo análogos de ADN tales como ácido nucleico peptídico (PNA).

[0084] Pueden introducirse sitios abásicos estables durante la síntesis química de oligonucleótidos usando unidades separadoras apropiadas que contienen el sitio abásico estable. A modo de ejemplo, pueden incorporarse separadores abásicos de furano (5'-O-dimetoxitritil-1',2'-didesoxirribosa-3'-[(2-cianoetil)-(N,N-diisopropil)]fosforamidita) disponibles comercialmente en Glen Research, Sterling, Virginia, Estados Unidos, durante la síntesis química del oligonucleótido con objeto de introducir un sitio abásico. Dicho sitio puede por lo tanto introducirse fácilmente en un cebador oligonucleotídico que se va a usar en una amplificación en fase sólida. Si se incorpora un sitio abásico en un cebador de amplificación directo o inverso, el producto resultante de la amplificación tendrá un saliente en 5' en una hebra que incluirá el sitio abásico (en la forma monohebra). El sitio abásico monohebra puede escindirse entonces por la acción de un agente químico adecuado (por ejemplo, exposición a un álcali) o una enzima (por ejemplo, endonucleasa AP-VI, Shida y col. Nucleic Acids Research, 1996, Vol. 24, 4572 - 4576).

45 vi) escisión de un péptido conector

[0085] También puede introducirse un sitio de escisión en una hebra de dúplex de polinucleótido molde preparando una estructura conjugada en la que se une una molécula peptídica a una hebra del dúplex (o el cebador de amplificación del que se deriva esta hebra si se ha preparado por amplificación en fase sólida). La molécula peptídica puede ser escindida subsiguientemente por una enzima peptidasa con la especificidad apropiada, o mediante cualquier otro medio adecuado de escisión química no enzimática o fotoquímica. Habitualmente, el conjugado entre el péptido y el ácido nucleico se formará uniendo covalentemente el péptido sólo a una hebra del dúplex de polinucleótido molde, estando conjugada la porción peptídica al extremo 5' de esta hebra, adyacente al punto de unión a la superficie sólida. Si el dúplex de polinucleótido molde se prepara mediante amplificación en fase sólida, el conjugado peptídico puede incorporarse en el extremo 5' de uno de los cebadores de amplificación. Obviamente, el componente peptídico de este cebador no será copiado durante la amplificación, por lo que el producto de la amplificación "en puente" incluirá un péptido escindible en 5' "sobresaliendo" de una hebra.

[0086] Los conjugados entre péptidos y ácidos nucleicos en los que el péptido está conjugado en el extremo 5' del ácido nucleico pueden prepararse usando técnicas conocidas habitualmente en la materia. En una de dichas 20

técnicas, los componentes de péptido y de ácido nucleico de la secuencia deseada de aminoácidos y nucleótidos pueden sintetizarse por separado, por ejemplo, mediante técnicas automatizadas de síntesis química estándar, y conjugarse después en una disolución acuosa/orgánica. A modo de ejemplo, el sistema OPeC™ disponible comercialmente en Glen Research, se basa en la "ligación nativa" de un péptido N-terminal funcionalizado con tioéster a un oligonucleótido de 5'-cisteinilo. Se usa S-benciltiosuccinato de pentafluorofenilo en la etapa de acoplamiento final del ensamblaje de péptidos en fase sólida basado en Fmoc estándar. La desprotección con ácido trifluoroacético genera, en disolución, péptidos sustituidos con un grupo S-benciltiosuccinilo N-terminal. Se usa O-trans-4-(N-a-Fmoc-S-terc-butilsulfenil-1-cisteinil)aminociclohexil-O-2-cianoetil-N,N-diisopropilfosforamidita en la etapa final de acoplamiento en el ensamblaje estándar de oligonucleótidos de fosforamidita en fase sólida. La desprotección con una disolución acuosa amoníaco genera en disolución oligonucleótidos funcionalizados con 5'-S-terc-butilsulfenil-L-cisteinilo. El tiobencilo terminal del Péptido Modificado se convierte en el análogo de tiofenilo mediante el uso de tiofenol, mientras que el Oligonucleótido Modificado se reduce usando tris (carboxietil) fosfina. El acoplamiento de estos dos intermedios, seguido por la etapa de "ligación nativa", conduce a la formación del Conjugado Oligonucleótido-Péptido.

15 **[0087]** La hebra conjugada que contiene el péptido y el ácido nucleico puede unirse covalentemente a un soporte sólido usando cualquier técnica de unión covalente adecuada conocida en la materia que sea compatible con la superficie elegida. Si la estructura del conjugado de péptido/ácido nucleico es un cebador de amplificación que se va a usar para una amplificación en fase sólida, la unión al soporte sólido debe dejar libre el extremo 3' del componente ácido nucleico.

25 **[0088]** El componente peptídico puede diseñarse para que sea escindible por cualquier enzima peptidasa elegida, de las que muchas son conocidas en la materia. La naturaleza de la peptidasa no está particularmente limitada, sólo es necesaria para que la peptidasa escinda en algún sitio del componente peptídico. De forma similar, la longitud de la secuencia de aminoácidos del componente peptídico no está particularmente limitada, excepto porque se requiere que sea "escindible" por la peptidasa elegida.

30 **[0089]** La longitud y la secuencia precisa del componente de ácido nucleico tampoco están particularmente limitadas, pueden ser de cualquier secuencia deseada. Si el componente de ácidos nucleico va a funcionar como un cebador en una amplificación en fase sólida, entonces su longitud y secuencia de nucleótidos se elegirán para permitir la hibridación con el molde que se va a amplificar.

vii) digestión enzimática con la endonucleasa de restricción/endonucleasa de corte

35 **[0090]** La escisión de polinucleótidos de doble hebra con endonucleasa de restricción es una técnica de uso rutinario en la técnica de la biología molecular. Las endonucleasas de corte son enzimas que escinden selectivamente o "cortan" una hebra de un dúplex de polinucleótido y también son bien conocidas en la técnica de la biología molecular. La invención no está limitada con respecto a la naturaleza de la enzima. Esencialmente, se puede utilizar cualquier endonucleasa de restricción o corte, siempre que una secuencia de reconocimiento adecuado pueda incluirse en el sitio de corte. En el caso de la invención que usa dos lecturas de secuenciación, la elección de la enzima de corte tendrá que ser diferente a la utilizada en el primer ciclo de extensión y esto permite usar cebadores de amplificación con sitios de reconocimiento para dos enzimas diferentes.

45 **[0091]** El procedimiento de la invención se describe con más detalle como sigue.

[0092] Se puede usar cualquier soporte sólido adecuado y cualquier unión adecuada conocidos en la técnica, de los cuales se describen a continuación varios a modo de ejemplo. En una realización particular, la unión al soporte sólido se logra a través de unión covalente.

50 **[0093]** Los dúplex de polinucleótidos habitualmente se forman a partir de dos hebras de polinucleótidos complementarias comprendidas por desoxirribonucleótidos unidos por enlaces fosfodiéster, pero pueden incluir adicionalmente uno o más ribonucleótidos y/o restos químicos no nucleótidos y/o nucleótidos de origen no naturales y/o uniones de esqueleto de origen no natural. En particular, el ácido nucleico de doble hebra puede incluir restos químicos no nucleótidos, por ejemplo, conectores o espaciadores, en el extremo 5' de una o ambas hebras. A modo de ejemplo no limitativo, el ácido nucleico de doble hebra puede incluir nucleótidos metilados, bases de uracilo, grupos fosforotioato, ribonucleótidos, conectores diol, enlaces disulfuro, péptidos, etc. Pueden incluirse tales modificaciones no de ADN o no naturales con el fin de permitir la escisión o para conferir alguna otra propiedad deseable, por ejemplo, para permitir la unión covalente a un soporte sólido, o para actuar como separadores para situar un sitio de escisión a una distancia óptima desde el soporte sólido.

60 **[0094]** Los dúplex molde pueden incluir también secuencias no diana, por ejemplo, en los extremos 5' y 3' de ambas hebras, la región flanqueante del polinucleótido diana. Si los dúplex molde se forman mediante amplificación en fase sólida, estas secuencias no diana pueden derivar habitualmente de los cebadores usados para la amplificación en fase sólida.

65

[0095] Los dúplex de polinucleótido forman parte de un único agregado o de una colonia que comprendida por muchos de dichos primero y segundo dúplex y el propio agregado o colonia habitualmente formará parte de una matriz de dichos agregados o colonias. Los términos "agregado" y "colonia" se utilizan indistintamente a lo largo de toda la memoria y se refieren a un sitio discreto sobre un soporte sólido que comprende una pluralidad de hebras de ácido nucleico inmovilizadas idénticas y una pluralidad de hebras de ácido nucleico complementarias inmovilizadas idénticas. El término "matriz agrupada" se refiere a un conjunto formado por tales grupos o colonias.

[0096] Una característica clave de la invención es que ambos ciclos de secuenciación pueden ocurrir en el mismo agregado o colonia en una matriz agregada. En una matriz de este tipo, cada dúplex dentro de cada colonia comprenderá el mismo polinucleótido de doble hebra diana, aunque se pueden formar diferentes colonias de dúplex que comprenden diferentes polinucleótidos diana de doble hebra. En una realización particular se formarán al menos un 90%, más particularmente al menos un 95% de las colonias en una matriz agrupada dada a partir de dúplex de molde que comprenden diferentes polinucleótidos diana de doble hebra, aunque dentro de cada colonia individual en la matriz, todos los dúplex molde comprenderán el mismo polinucleótido de doble hebra diana.

[0097] Los polinucleótidos amplificados pueden ser tratados de tal manera que se permita la extensión de un cebador hibridado. Cada dúplex de polinucleótido en la matriz contiene las mismas regiones de reconocimiento de cebador universales para permitir que se utilicen los mismos cebadores para secuenciar cada agregado. A continuación, se hibrida un primer cebador de secuenciación con la primera hebra del molde y procede una reacción de secuenciación a través de la incorporación sucesiva de nucleótidos u oligonucleótidos al primer cebador de secuenciación, lo que resulta en la determinación de la secuencia de una primera región del polinucleótido diana. En el caso del procedimiento en el que ambos extremos del fragmento original se hacen contiguos (véase la Figura 3), el primer ciclo de secuenciación da información de la secuencia de ambos extremos del fragmento original. En este punto, el procedimiento se ha completado, no hay necesidad de procesamiento adicional de la muestra.

[0098] Por el contrario, los otros procedimientos de la invención requieren la realización de dos reacciones de secuenciación. La primera reacción de secuenciación se inicia ya sea por el grupo hidroxilo en 3' del cebador inmovilizado que se libera desde el interior del dúplex inmovilizado, o se inicia por un cebador de secuenciación añadido de la solución. La segunda reacción de secuenciación se inicia por un cebador de secuenciación que, o bien puede estar inmovilizado o se aplica en solución. La hibridación del cebador de secuenciación en solución con la hebra molde se consigue poniendo en contacto el cebador y la hebra molde en condiciones que promueven la hibridación del cebador con el molde. Tales condiciones habitualmente serán bien conocidas por los expertos en la técnica de la biología molecular.

[0099] Cuando la primera reacción de secuenciación se haya completado, si el primer cebador de secuenciación se inmoviliza, esto se puede utilizar para ampliar aún más la hebra para copiar cada base en la plantilla. Esto se lleva a cabo usando los cuatro trifosfatos de nucleótidos nativos, dATP, dGTP, dCTP y dTTP y una polimerasa adecuada. Si la hebra complementaria se eliminó por tratamiento con una etapa con 5'-3' exonucleasa antes de la reacción de secuenciación, entonces, la reacción de extensión puede llevarse a cabo a cualquier temperatura a la que la hebra ya es de hebra sencilla y, por lo tanto, cualquier polimerasa como Klenow, Taq o ven es adecuada. Si la hebra no se ha eliminado, entonces se requiere una polimerasa con actividad de desplazamiento de la hebra, tales como la Bst polimerasa.

[0100] Antes de llevar a cabo la reacción de extensión, puede ser ventajoso extender el cebador inmovilizado, tal como se muestra en la figura 10. La extensión se puede realizar usando un oligonucleótido hibridado con una secuencia que se extiende más allá del extremo 3' del cebador inmovilizado, cuya secuencia es también la misma secuencia que la región correspondiente en el extremo del molde. Esta parte extendida puede servir como una base para la extensión del cebador inmovilizado y, por lo tanto, el cebador extendido es complementario con la hebra de molde inmovilizada. Los cebadores extendidos pueden mejorar la eficiencia de la etapa de resíntesis de la hebra debido a su mayor longitud.

[0101] Una vez que se ha generado una secuencia complementaria de la primera hebra, la primera hebra puede eliminarse de la superficie como se describe anteriormente. Un segundo cebador de secuenciación se hibrida a continuación con la hebra copiada del molde y una reacción de secuenciación procede mediante la adición sucesiva de nucleótidos al segundo cebador de secuenciación, lo que tiene como resultado la determinación de la secuencia de una segunda región del polinucleótido diana. Véase la Figura 1.

[0102] Como se describió anteriormente, la secuenciación puede llevarse a cabo usando cualquier técnica adecuada de "secuenciación por síntesis", en la que se añaden nucleótidos u oligonucleótidos sucesivamente a un grupo hidroxilo libre en 3', habitualmente proporcionada por hibridación de un cebador de secuenciación, lo que tiene como resultado la síntesis de un hebra de polinucleótido en la dirección 5' a 3'. En una realización particular, la naturaleza del nucleótido u oligonucleótido añadido se determina después de cada adición.

[0103] Un procedimiento de secuenciación particular, que se puede utilizar en los procedimientos de la invención se basa en el uso de nucleótidos modificados que pueden actuar como terminadores de hebra reversibles. Los

nucleótidos para su uso en la invención se describen completamente en el documento W004018497. Una vez que el nucleótido modificado ha sido incorporado en la hebra de polinucleótidos en crecimiento complementaria con la región del molde que se está secuenciado no existe un grupo 3'-OH libre disponible para dirigir más la extensión de la secuencia y, por lo tanto, la polimerasa no puede añadir más nucleótidos. Una vez que se ha determinado la naturaleza de la base incorporada en la hebra de crecimiento, el bloque 3' se puede eliminar para permitir la adición del nucleótido sucesivo. Al ordenar los productos derivados que utilizan estos nucleótidos modificados, es posible deducir la secuencia de ADN del molde de ADN. Tales reacciones pueden realizarse en un solo experimento si cada uno de los nucleótidos modificados ha unido a la misma un marcador diferente, que se sabe que corresponde a la base en particular, lo que facilita la discriminación entre las bases añadidas en cada etapa de incorporación.

10 Alternativamente, se puede llevar a cabo una reacción separada que contiene cada uno de los nucleótidos modificados, que se añaden por separado.

[0104] Los nucleótidos modificados pueden llevar un marcador para facilitar su detección. En una realización particular, el marcador es un marcador fluorescente. Cada tipo de nucleótido puede llevar un marcador fluorescente diferente. Los marcadores fluorescentes adecuados para su uso en la presente invención se describen en la solicitud PCT PCT/GB2007/001770. Sin embargo, el marcador detectable no necesita ser un marcador fluorescente. Se puede usar cualquier marcador que permita la detección de la incorporación de los nucleótidos en la secuencia de ADN.

20 **[0105]** Un procedimiento para la detección de los nucleótidos marcados con fluorescencia comprende el uso de luz láser de una longitud de onda específica para los nucleótidos marcados, o el uso de otras fuentes adecuadas de iluminación. La fluorescencia del marcador sobre el nucleótido puede ser detectada por una cámara CCD u otros medios de detección adecuados. Un sistema de formación de imágenes adecuado para determinar la señal de fluorescencia de nucleótidos incorporados se describe en el Número de solicitud PCT PCT/US07/007991.

25 **[0106]** Los procedimientos de la invención no se limitan al uso del procedimiento de secuenciación descrito anteriormente, pero se pueden utilizar conjuntamente con esencialmente cualquier metodología de secuenciación que se base en la incorporación sucesiva de nucleótidos en una hebra de polinucleótidos. Las técnicas adecuadas incluyen, por ejemplo, Pyrosequencing™, FISSEQ (resecuenciación fluorescente in situ), MPSS (secuenciación masiva de firmas en paralelo) y la secuenciación mediante procedimientos basados en la ligación.

[0107] El polinucleótido de doble hebra diana que se va a secuenciar usando el procedimiento de la invención puede ser cualquier polinucleótido que se desee secuenciar. El polinucleótido diana puede ser de secuencia conocida, desconocida o parcialmente conocida, tal como, por ejemplo, en aplicaciones de resecuenciación. Usando el procedimiento de preparación de moldes descrito con detalle a continuación es posible preparar matrices de moldes partiendo esencialmente de cualquier poli nucleótido de doble hebra diana de secuencia conocida, desconocida o parcialmente conocida. Con el uso de matrices es posible secuenciar en paralelo múltiples dianas con la misma o con diferente secuencia. Una aplicación particularmente preferida del procedimiento por pares es la secuenciación de fragmentos de ADN genómico. El procedimiento proporciona ventajas particulares en la identificación de los reordenamientos del genoma, ya que las dos regiones de secuencia obtenidas para cada molécula diana utilizando el procedimiento se sabe que están asociadas dentro de una cierta distancia la una de la otra en el genoma, dependiendo del tamaño de la molécula diana de partida.

Preparación de los moldes a secuenciar

45 **[0108]** Moldes adecuados para la secuenciación utilizando el procedimiento de la invención se pueden preparar por amplificación de ácidos nucleicos en fase sólida para producir colonias de ácido nucleico. Los moldes para ser amplificados comprenderán habitualmente regiones desconocidas flanqueadas por extremos conocidos, por ejemplo, preparadas de acuerdo con los procedimientos descritos en la solicitud WO07052006.

50 **[0109]** Por ejemplo, los moldes pueden derivar de una muestra de ADN genómico, o de una biblioteca de ADNc. La amplificación se puede realizar usando procedimientos análogos a los descritos en los documentos WO 98/44151, WO 00/18957, WO0206456 o WO07107710.

55 **[0110]** Los moldes para la amplificación se pueden preparar para contener una secuencia de marcador, por ejemplo como se muestra en la Figura 12. El uso de una secuencia marcador, por ejemplo, como se describe en la solicitud WO05068656, permite analizar múltiples muestras diferentes en el mismo ciclo de secuenciación preservando al mismo tiempo la identidad de cada muestra. La Figura 12 muestra el marcador que se lee al final de la primera lectura, aunque el marcador se puede leer igualmente al final de la segunda lectura, por ejemplo, usando un cebador de secuenciación complementario con la hebra marcada P7. La invención no se limita a dos lecturas por agregado, se pueden obtener tres o más lecturas por agregados simplemente deshibridando un primer cebador de secuenciación extendida y rehibridando un segundo cebador antes o después de la repoblación del agregado/etapa de resíntesis del agregado. Los procedimientos de preparación de muestras adecuadas para la indexación se describen en la solicitud copendiente US60/899221, presentada el 1 de febrero de 2007.

65

[0111] Para que tenga lugar la amplificación, una mezcla de al menos dos cebadores de amplificación se inmoviliza o se "injerta" sobre la superficie de un soporte sólido adecuado.

[0112] Los cebadores de amplificación son moléculas de oligonucleótidos que tienen las siguientes estructuras:

5

Cebador directo: A-L-X-S1

Cebador inverso: A-L-Y-S2

[0113] En las que A representa un resto que permite la unión a un soporte sólido, L es un resto conector o separador opcional, X e Y son sitios de escisión opcionales como se describió anteriormente para permitir la eliminación posterior de una u otra de las hebras de la superficie y S1 y S2 son secuencias de polinucleótidos que permiten la amplificación de una molécula de ácido nucleico molde que comprende el polinucleótido de doble hebra diana.

[0114] La mezcla de cebadores comprenderá habitualmente sustancialmente cantidades iguales de los cebadores directo e inverso.

[0115] L representa un conector que puede incluirse, pero no es estrictamente necesario. El conector puede ser una hebra carbonada, tal como aquellas de fórmula $(CH_2)_n$ en la que "n" es desde 1 hasta aproximadamente 1500, por ejemplo menos de aproximadamente 1000, preferiblemente menos de 100, por ejemplo, de 2-50, particularmente de 5-25. Sin embargo, puede emplearse una variedad de otros conectores con la única restricción en sus estructuras de que los conectores sean estables en las condiciones en las que pretenden usarse subsiguientemente los polinucleótidos, por ejemplo, las condiciones usadas en la amplificación y secuenciación del ADN.

[0116] También pueden usarse conectores que no consistan únicamente en átomos de carbono. Dichos conectores incluyen polietilenglicol (PEG) con una fórmula general $(CH_2-CH_2-O)_m$, en la que m es desde aproximadamente 1 hasta 600, preferiblemente menos de aproximadamente 500.

[0117] Los conectores formados principalmente a partir de hebras de átomos de carbono y a partir de PEG pueden modificarse para que contengan grupos funcionales que interrumpan las hebras. Algunos ejemplos de dichos grupos incluyen cetonas, ésteres, aminas, amidas, éteres, tioéteres, sulfóxidos, sulfonas. Por separado o en combinación, con la presencia de dichos grupos funcionales pueden emplearse restos de alqueno, alquino, aromáticas o heteroaromáticas, o restos alifáticos cíclicos (por ejemplo, ciclohexilo). Los anillos de ciclohexilo o de fenilo pueden estar conectados, por ejemplo, a PEG o a una hebra de $(CH_2)_n$ a través de sus posiciones 1 y 4.

[0118] Como alternativa a los conectores descritos anteriormente, que se basan principalmente en hebras lineales de átomos de carbono saturados, opcionalmente interrumpidas con átomos de carbono insaturados o con heteroátomos, pueden contemplarse otros conectores que se basan en unidades de ácidos nucleicos o de monosacáridos (por ejemplo, glucosa). También está en el ámbito de la invención utilizar péptidos como conectores.

[0119] En una realización adicional, el conector puede comprender uno o más nucleótidos que forman parte del cebador de amplificación, pero que no participan en ninguna reacción llevada a cabo en o con el cebador (por ejemplo, una reacción de hibridación o amplificación). Dichos nucleótidos también pueden denominarse en la presente memoria polinucleótidos "separadores". Habitualmente, pueden incluirse entre 1 y 20, más preferiblemente entre 1 y 15 o entre 1 y 10, y más particularmente 2, 3, 4, 5, 6, 7, 8, 9 ó 10 nucleótidos separadores. En una realización particular, el cebador incluirá 10 nucleótidos separadores. En algunas realizaciones se puede usar separadores de poliT, aunque pueden usarse otros nucleótidos y combinaciones de los mismos. En una realización particular, el cebador puede incluir nucleótidos separadores de 10T.

[0120] El uno o más nucleótidos separadores sirven para separar la porción del cebador requerido para hibridarse con una diana y la amplificación directa lejos del sitio de unión al soporte sólido (es decir, S1 o S2). La inclusión de nucleótidos separadores en el extremo 5' puede mejorar notablemente el rendimiento de la hibridación de polinucleótidos complementarios con la región S1 o S2. En una realización más particular, el polinucleótido incluye nucleótidos separadores 10T y un grupo 5' fosforotioato para la fijación al soporte sólido (resto A), aunque se pueden usar otros restos de como se discute a continuación.

55

[0121] Las secuencias S1 y S2 en los cebadores directo e inverso son secuencias de polinucleótidos que, en combinación, la amplificación dirigen la amplificación de un molde por reacción de amplificación en puente en fase sólida. El molde a amplificar debe comprender en sí mismo (cuando se considera como una sola hebra) en el extremo 3' una secuencia capaz de hibridar con la secuencia S1 en los cebadores directos y en el extremo 5' una secuencia cuyo complemento es capaz de hibridarse con la secuencia S2 del cebador inverso.

60

[0122] La identidad precisa de las secuencias S1 y S2 en los oligonucleótidos cebadores directo e inverso será dependiente de la naturaleza del molde que se pretende amplificar. S1 y S2 deben ser capaces de hibridarse a secuencias cognadas en hebras complementarias del molde a amplificar. El término "hibridación" abarca la unión específica de secuencia entre el cebador y el molde. La unión de un cebador con su secuencia cognada en el molde

65

debe ocurrir en condiciones típicas utilizadas para la hibridación cebador-molde en la PCR estándar. Habitualmente, las condiciones de hibridación son 5 x SSC a 40 °C, después de una etapa inicial de desnaturalización. No es esencial para la hibridación que las secuencias S1 y S2 sean exactamente complementarias con sus secuencias cognadas en el molde a amplificar, aunque esto se prefiere.

5 **[0123]** S1 y S2 pueden tener una secuencia diferente o idéntica y tendrán típicamente alrededor de 20 a 30 nucleótidos de longitud. Los cebadores pueden incluir bases naturales y no naturales de ADN, también ribonucleótidos o cualquier combinación de los mismos y también pueden incluir uniones de esqueleto no naturales, tales como disulfuros o fosforotioatos.

10 **[0124]** El sitio de escisión de X y/o Y puede estar dentro de la secuencia S1 o S2, o si el conector L es por sí mismo una escisión de polinucleótido que puede formar parte de la región conector L. En otras realizaciones, el sitio de corte puede estar formado en la unión de las secuencias de L y S1 o L y S2, o en la unión entre el resto A y el conector L (si está presente) o entre el resto A y la secuencia S1 o S2 (si L no está presente).

15 **[0125]** El resto A puede ser cualquier resto químico que permita la inmovilización de un cebador de oligonucleótido sobre un soporte sólido. La superficie del soporte sólido puede estar ella misma funcionalizada para permitir la unión de los cebadores. Cualquier medio de unión covalente o no covalente adecuado puede ser utilizado, de los cuales muchos son conocidos en la técnica.

20 **[0126]** A modo de ejemplo, las albúminas biotiniladas (BSA) pueden formar una unión estable de grupos de biotina por fisiorción de la proteína sobre las superficies. También se puede llevar a cabo la modificación covalente utilizando silanos, que se han utilizado para unir moléculas a un soporte sólido, por lo general un portaobjetos de vidrio. A modo de ejemplo, se puede utilizar una mezcla de tetraetoxisilano y trietoxi-bromoacetamidopropil-silano
25 (por ejemplo, en una proporción de 1:100) para preparar portaobjetos de vidrio funcionalizadas que permiten la unión de moléculas, tales como ácidos nucleicos que incluyen una funcionalidad tiofosfato o fosforotioato. Las moléculas de biotina se pueden unir a las superficies utilizando especies adecuadamente reactivas, tales como éster de biotina-PEG-succinimidilo que reacciona con una superficie de amino. Una mezcla de cebadores de amplificación puede ponerse a continuación en contacto con el soporte sólido funcionalizado.

30 **[0127]** En realizaciones alternativas se pueden usar hidrogeles de poliacrilamida funcionalizados para unir cebadores en los que el resto A es un grupo nucleófilo que contiene azufre. Los ejemplos de polinucleótidos que contienen nucleófilos de azufre adecuados se describen en Zhao et al (Nucleic Acids Research, 2001, 29(4), 955-959) y Pirrung et al (Langmuir, 2000, 16, 2185-2191) e incluyen, por ejemplo, tioles simples, tiofosfatos y tiofosforamidatos. En realizaciones particulares, los hidrogeles se forman a partir de una mezcla de (i) un primer comonomero que es acrilamida, metacrilamida, metacrilato de hidroxietilo o N-vinil pirrolidina; y

(ii) un segundo comonomero que es una acrilamida funcionalizada o un acrilato de fórmula (I) :

40
$$H_2C=C(H)-C(=O)-A-B-C \text{ (I) ;}$$

o un metacrilato o una metacrilamida de fórmula (II): o

45
$$H_2C=C(CH_3)-C(=O)-A-B-C \text{ (II)}$$

en la que:

A es NR u O, en la que R es hidrógeno o un grupo hidrocarbilo saturado opcionalmente sustituido que comprende de 1 a 5 átomos de carbono;

50 -B- es un birradical alquileno opcionalmente sustituido de fórmula $-(CH_2)_n-$ en la que n es un número entero de 1 a 50; y en la que n = 2 o más, uno o más birradicales etileno opcionalmente sustituidos $-CH_2CH_2-$ de dicho birradical alquileno pueden estar independientemente sustituidos por restos de etenileno y etinileno; y en la que n = 1 o más, uno o más birradicales metileno $-CH_2-$ pueden estar independientemente sustituidos por un birradical hidrocarbonado monocíclico o policíclico opcionalmente sustituido que comprende de 4 a 50 átomos de carbono, o un correspondiente birradical heteromonocíclico o heteropolicíclico, en el que al menos 1 CH_2 o CH_2 está sustituido por un átomo de oxígeno, de azufre o de nitrógeno, o un grupo NH; y

55 C es un grupo para la reacción con un compuesto, para unir covalentemente el compuesto al hidrogel) para formar un producto polimerizado. En una realización particular, el hidrogel se forma por la copolimerización de acrilamida y N-(5-bromoacetamidilpentil)acrilamida (Embrapa), como se describe en la solicitud WO05065814.
60

[0128] El término "soporte sólido", como se usa en la presente memoria, se refiere al material al que se unen las moléculas de polinucleótidos. Algunos soportes sólidos adecuados están disponibles comercialmente y serán evidentes para el experto. Los soportes pueden elaborarse a partir de materiales tales como vidrio, cerámica, sílice y
65 silicio. También pueden usarse soportes con una superficie de oro. Los soportes comprenden habitualmente una

superficie plana, o al menos, una estructura en la que los polinucleótidos que se van a analizar están aproximadamente en el mismo plano. Alternativamente, el soporte sólido puede no ser plano, por ejemplo, una microperla. Puede usarse cualquier tamaño adecuado. Por ejemplo, los soportes podrían ser del orden de 1 - 10 cm en cada dirección.

5

[0129] Para proceder a la reacción de injerto del cebador, se aplica una mezcla de los cebadores de amplificación sobre un soporte sólido (funcionalizado adecuadamente) en unas condiciones que permitan la reacción entre el resto A y el soporte. El resultado de la reacción de injerto es una distribución sustancialmente uniforme de los cebadores sobre el soporte sólido.

10

[0130] En ciertas realizaciones, el molde que se va a amplificar puede injertarse sobre el soporte sólido junto con los cebadores de amplificación en una única reacción de injerto. Esto puede conseguirse añadiendo a la mezcla de cebadores moléculas de molde que incluyan el resto A en el extremo 5' para formar una mezcla de cebador-molde. Esta mezcla se injerta a continuación sobre el soporte sólido en una única etapa. A continuación, puede proceder la amplificación usando el molde y los cebadores inmovilizados en una reacción análoga a la descrita en el documento WO 00/18957. La primera etapa de dicha reacción será la hibridación entre los moldes unidos a la superficie y los cebadores de amplificación unidos a la superficie.

15

[0131] Si únicamente se injerta una mezcla de cebadores en el soporte sólido y el molde que se va a amplificar puede añadirse libre en disolución, entonces puede proceder la reacción de amplificación sustancialmente según se describe en el documento WO 98/44151. Resumiendo, tras la unión de los cebadores, el soporte sólido se pone en contacto con el molde que se va a amplificar en unas condiciones que permitan la hibridación entre el molde y los cebadores inmovilizados. Habitualmente el molde se añade libre en disolución en unas condiciones de hibridación adecuadas, que serán apreciables por el lector experto. Las condiciones típicas de hibridación son, por ejemplo, 5 x 25 SSC a 40 °C, tras una etapa de desnaturalización inicial. A continuación, puede proceder la amplificación en fase sólida, siendo la primera etapa una etapa de extensión del cebador en el que se añaden nucleótidos en el extremo 3' del cebador inmovilizado hibridado con el molde para producir una hebra complementaria completamente extendida. Esta hebra complementaria incluirá, por tanto, en su extremo 3' una secuencia que es capaz de unirse a la segunda molécula de cebador inmovilizada sobre el soporte sólido. Las rondas adicionales de amplificación conducen a la 30 formación de agregados o colonias de moléculas de molde unidas al soporte sólido.

[0132] Las secuencias S1 y S2 de los cebadores de amplificación pueden ser específicas para un ácido nucleico diana en particular que se desea a amplificar, pero en otras realizaciones, las secuencias S1 y S2 pueden ser secuencias de cebado "universales" que permiten la amplificación de cualquier ácido nucleico diana con una 35 secuencia conocida o desconocida que ha sido modificada para permitir la amplificación con los cebadores universales.

[0133] Los moldes adecuados que se van a amplificar con cebadores universales pueden prepararse mediante la modificación de polinucleótidos de doble hebra diana mediante la adición de secuencias adaptadoras conocidas en 40 los extremos 5' y 3' de los polinucleótidos diana que se van a amplificar. Las propias moléculas diana pueden ser cualquier molécula de doble hebra que se desee secuenciar (por ejemplo, fragmentos aleatorios de ADN genómico humano). Las secuencias adaptadoras permiten la amplificación de estas moléculas sobre un soporte sólido para formar agregados usando cebadores directos e inversos con la estructura general descrita anteriormente, en la que las secuencias S1 y S2 son secuencias de un cebador universal.

45

[0134] Los adaptadores son típicamente oligonucleótidos cortos que pueden sintetizarse mediante medios convencionales. Los adaptadores pueden unirse a los extremos 5' y 3' de fragmentos de ácidos nucleicos diana mediante diversos medios (por ejemplo, subclonación, ligación, etc.). Más específicamente, se unen dos secuencias adaptadoras diferentes a una molécula de ácido nucleico diana que se va a amplificar de forma que un adaptador 50 está unido en un extremo de la molécula de ácido nucleico diana y otro adaptador está unido en el otro extremo de la molécula de ácido nucleico diana. El constructo resultante que comprende una secuencia de ácidos nucleicos objetivo flanqueada por adaptadores puede denominarse en la presente memoria "constructo de ácido nucleico molde". Los procedimientos adecuados de preparación de la muestra para su uso en el procedimiento descrito en la presente memoria se detallan en el Número de solicitud US 11/486953.

55

[0135] Los polinucleótidos de doble hebra diana pueden fraccionarse ventajosamente por tamaños antes de la modificación con las secuencias adaptadoras.

[0136] Los adaptadores contienen secuencias que permiten la amplificación de los ácidos nucleicos usando 60 moléculas cebadoras de la amplificación inmovilizadas sobre el soporte sólido. Estas secuencias de los adaptadores pueden denominarse en la presente memoria "secuencias de unión del cebador". Con objeto de que actúen como molde para la amplificación de ácidos nucleicos, una única hebra del constructo de molde debe contener una secuencia que sea complementaria de la secuencia S1 de los cebadores de amplificación directos (de forma que la molécula del cebador directo pueda unirse y cebar la síntesis de una hebra complementaria) y una secuencia que se 65 corresponda con la secuencia S2 de las moléculas del cebador de amplificación inverso (de forma que la molécula

del cebador inverso pueda unirse a la hebra complementaria). Las secuencias de los adaptadores que permiten la hibridación con moléculas de cebadores tendrán típicamente aproximadamente 20-30 nucleótidos de longitud, aunque la invención no se limita a secuencias de esta longitud.

5 **[0137]** La identidad precisa de las secuencias S1 y S2 de los cebadores de amplificación, y por lo tanto, las secuencias cognadas de los adaptadores, habitualmente no son relevantes para la invención, siempre que las moléculas del cebador sean capaces de interactuar con las secuencias de amplificación con objeto de dirigir la amplificación en puente. Los criterios para el diseño de los cebadores son habitualmente bien conocidos por los expertos habituales en la materia.

10

[0138] La amplificación en fase sólida mediante el procedimiento análogo al del documento WO 98/44151 o al del documento WO 00/18957 dará como resultado la producción de una matriz de colonias de productos de amplificación "en puente". Ambas hebras de los productos de amplificación estarán inmovilizadas sobre el soporte sólido en, o cerca de, el extremo 5', derivando esta unión de la unión original de los cebadores de amplificación.

15 Habitualmente, los productos de la amplificación de cada colonia derivarán de la amplificación de una única molécula de molde.

[0139] Los procedimientos de amplificación adecuados para la presente invención incluyen termociclado y amplificaciones isotérmicas. En una realización particular de la invención, se usan las condiciones de amplificación isotérmica, como se describe en el Número de solicitud WO07107710. Las temperaturas más bajas utilizadas en amplificaciones isotérmicas, que simplemente implican ciclos de cambio de tampón para cambiar entre la extensión y las condiciones de desnaturalización, son ventajosos debido a una mayor retención de las moléculas unidas a la superficie, y por lo tanto, los agregados más brillantes (es decir, agregados que tienen una mayor intensidad de fluorescencia después de la incorporación de un nucleótido marcado con fluorescencia).

25

[0140] La utilidad del procedimiento de secuenciación de la invención no se limita a la secuencia de los moldes producidos por una reacción de amplificación, aunque esto se prefiere. El procedimiento se puede aplicar a la secuenciación de moldes de doble hebra inmovilizados sobre un soporte por cualquier otro medio susceptible de ciclos repetidos de hibridación y secuenciación.

30

[0141] La invención se comprenderá adicionalmente con referencia a los siguientes ejemplos experimentales no limitantes.

Ejemplos

35

[0142] Los siguientes son ejemplos de técnicas generales que se pueden aplicar en la realización del procedimiento de la invención.

Ejemplo 1: Preparación de un agregado utilizando la invención detallada

40

Recubrimiento con acrilamida de chips de vidrio

[0143] Los soportes sólidos usados fueron habitualmente chips de vidrio de 8 canales, tales como los proporcionados por Silex Microsystems (Suecia). Sin embargo, las condiciones experimentales y los procedimientos son fácilmente aplicables a otros soportes sólidos.

45

[0144] Los chips se lavaron como sigue: Decon puro durante 30 min, H₂O milliQ durante 30 min, NaOH 1 N durante 15 min, H₂O milliQ durante 30 min, HCl 0,1 N durante 15 min, H₂O milliQ durante 30 min.

50 **[0145]** Preparación de la disolución de polímero:

[0146] Para 10 ml de una mezcla de polimerización al 2%.

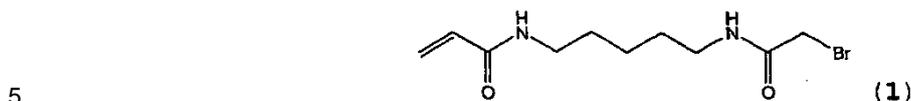
- 10 ml de una disolución al 2% de acrilamida en H₂O milliQ
- 55 - 165 µl de una disolución de 100 mg/ml de N-(5-bromoacetamidilpentil) acrilamida (BRAPA) en DMF (23, 5 mg en 235 µl de DMF)
- 11, 5 µl de TEMED
- 100 µl de una disolución de 50 mg/ml de persulfato potásico en H₂O milliQ (20 mg en 400 µl de H₂O)

60 **[0147]** En primer lugar se desgasificó la disolución de 10 ml de acrilamida con argón durante 15 min. Sucesivamente se añadieron las disoluciones de BRAPA, TEMED y persulfato potásico a la disolución de acrilamida. Después la mezcla se sometió rápidamente a una agitación vorticial y se usó inmediatamente. A continuación, se llevó a cabo la polimerización durante 1 h 30 a temperatura ambiente. A continuación, se lavaron los canales con H₂O milliQ durante 30 min y se llena con tampón fosfato de potasio 0,1 M para su conservación hasta que sea necesario.

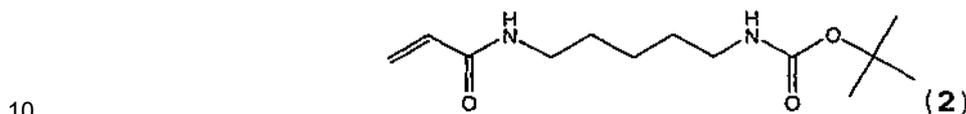
65

Ejemplo 2: Síntesis de *N*-(5-bromoacetamidilpentil) acrilamida (BRAPA) (1)

[0148]



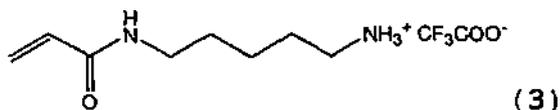
[0149] Se obtuvo ácido *N*-Boc-1,5-diaminopentanotoluensulfónico en Novabiochem. El cloruro de bromoacetilo y el cloruro de acrilóilo se obtuvieron en Fluka. Todos los demás reactivos eran productos de Aldrich.



[0150] A una suspensión agitada de ácido *N*-Boc-1,5-diaminopentanotoluensulfónico (5,2 g, 13,88 mmol) y trietilamina (4,83 ml, 2,5 eq) en THF (120 ml) a 0 °C se añadió cloruro de acrilóilo (1,13 ml, 1 eq) a través de un embudo de goteo con presión igualada durante un periodo de una hora. A continuación, la mezcla de reacción se agitó a temperatura ambiente y se comprobó el progreso de la reacción mediante TLC (éter de petróleo:acetato de etilo 1:1). Después de dos horas, las sales formadas durante la reacción se filtraron y el filtrado se evaporó a sequedad. El residuo se purificó mediante cromatografía ultrarrápida (éter de petróleo puro seguido por un gradiente de acetato de etilo de hasta el 60%) para dar 2,56 g (9,98 mmol, 71%) del producto 2 como un sólido beige. RMN de ¹H (400 MHz, d₆-DMSO) : 1,20-1,22 (m, 2H, CH₂), 1,29-1,43 (m, 13H, tBu, 2 x CH₂), 2,86 (c, 2H, J = 6,8 Hz y 12,9 Hz, CH₂), 3,07 (c, 2H, J = 6,8 Hz y 12,9 Hz, CH₂), 5,53 (dd, 1H, J = 2,3 Hz y 10,1 Hz, CH), 6,05 (dd, 1H, J = 2,3 Hz y 17,2 Hz, CH), 6,20 (dd, 1H, J = 10,1 Hz y 17,2 Hz, CH), 6,77 (t, 1H, J = 5,3 Hz, NH), 8,04 (sa, 1H, NH). Masa (electronebulización+) calculada para C₁₃H₂₄N₂O₃ 256, encontrada 279 (256+Na⁺).

15

20



[0151] Se disolvió el producto 2 (2,56 g, 10 mmol) en ácido trifluoroacético:diclorometano (1:9, 100 ml) y se agitó a temperatura ambiente. El progreso de la reacción se monitorizó mediante TLC (diclorometano:metanol 9:1). Al completarse, la mezcla de reacción se evaporó a sequedad, el residuo se coevaporó tres veces con tolueno y después se purificó mediante cromatografía ultrarrápida (diclorometano puro seguido por un gradiente de metanol de hasta el 20%). El producto 3 se obtuvo como un polvo blanco (2,43 g, 9 mmol, 90%). RMN de ¹H (400 MHz, D₂O) : 1,29-1,40 (m, 2H, CH₂), 1,52 (quint., 2H, J = 7,1 Hz, CH₂), 1,61 (quint., 2H, J = 7,7 Hz, CH₂), 2,92 (t, 2H, J = 7,6 Hz, CH₂), 3,21 (t, 2H, J = 6,8 Hz, CH₂), 5,68 (dd, 1H, J = 1,5 Hz y 10,1 Hz, CH), 6,10 (dd, 1H, J = 1,5 Hz y 17,2 Hz, CH), 6,20 (dd, 1H, J = 10,1 Hz y 17,2 Hz, CH). Masa (electronebulización+) calculada para C₈H₁₆N₂O 156, encontrada 179 (156+Na⁺).

30

[0152] A una suspensión del producto 3 (6,12 g, 22,64 mmol) y trietilamina (6,94 ml, 2,2 eq) en THF (120 ml) se añadió cloruro de bromoacetilo (2,07 ml, 1,1 eq), a través de un embudo de goteo con presión igualada durante un periodo de una hora y a -60 °C (baño de hielo seco e isopropanol en un vaso dewar). A continuación, la mezcla de reacción se agitó a temperatura ambiente hasta el día siguiente y la compleción de la reacción fue comprobada mediante TLC (diclorometano:metanol 9:1) el día siguiente. Las sales formadas durante la reacción se filtraron y la mezcla de reacción se evaporó a sequedad. El residuo se purificó mediante cromatografía (diclorometano puro seguido por un gradiente de metanol de hasta un 5%). Se obtuvieron 3,2 g (11,55 mmol, 51%) del producto 1 (BRAPA) como un polvo blanco. Una recristalización adicional realizada en éter de petróleo:acetato de etilo dio 3 g del producto 1. RMN de ¹H (400 MHz, d₆-DMSO) : 1,21-1,30 (m, 2H, CH₂), 1,34-1,48 (m, 4H, 2 x CH₂), 3,02-3,12 (m, 4H, 2 x CH₂), 3,81 (s, 2H, CH₂), 5,56 (d, 1H, J = 9,85 Hz, CH), 6,07 (d, 1H, J = 16,9 Hz, CH), 6,20 (dd, 1H, J = 10,1 Hz y 16,9 Hz, CH), 8,07 (sa, 1H, NH), 8,27 (sa, 1H, NH). Masa (electronebulización+) calculada para C₁₀H₁₇BrN₂O₂ 276 ó 278, encontrada 279 (278+H⁺), 299 (276+Na⁺).

35

40

45

Ejemplo 3: Injerto de cebadores sobre la superficie de chip recubierto con SFA

50

[0153] Se colocó un chip recubierto con SFA en un termociclador de MJ-Research modificado y se unió a una bomba peristáltica. La mezcla de injerto, que consistía en 0,5 μM de un cebador directo y 0,5 μM de un cebador inverso en tampón de fosfato 10 mM (pH 7,0), se bombeó en los canales del chip a un caudal de 60 μl/min durante 75 s a 20 °C. A continuación, se calentó el termociclador hasta 51,6 °C y el chip se incubó a esta temperatura durante 1 hora. Durante este tiempo, la mezcla de injerto experimentó 18 ciclos de bombeo: la mezcla de injerto se bombeó a 15 μl/min durante 20 s, después la disolución se bombeó hacia atrás y hacia delante (5 s hacia delante a

55

15 $\mu\text{l}/\text{min}$, después 5 s hacia atrás a 15 $\mu\text{l}/\text{min}$) durante 180 s. Después de 18 ciclos de bombeo, el chip se lavó mediante bombeo con 5 x SSC/EDTA 5 mM a 15 $\mu\text{l}/\text{min}$ durante 300 s a 51,6 °C. A continuación, se enfrió el termociclador a 20 °C.

- 5 **[0154]** Los cebadores serán típicamente oligonucleótidos de 5'-fosforotioato que incorporan cualquier secuencia o modificación específica necesaria para la escisión. Sus secuencias y los proveedores variarán según el experimento en el que se vayan a usar, y en este caso eran complementarios de los extremos 5' del dúplex de molde. La secuencia de ADN usada en este proceso fue una única secuencia monomolde de 240 bases, con extremos complementarios a los cebadores injertados. La secuencia completa del dúplex de molde se muestra en la Figura 6.
- 10 El dúplex de ADN se desnaturalizó usando tratamiento con hidróxido de sodio seguido por la dilución a presión como se ha descrito.

- [0155]** Para algunos de los experimentos descritos, los agregados amplificados contenían un conector de diol en uno de los cebadores injertados. Los conectores diol pueden introducirse incluyendo un conector adecuado en uno de los cebadores usados para la amplificación en fase sólida. La síntesis de fosforamidita diol se describe en el Ejemplo 4 a continuación. Los productos que contienen dichos conectores diol pueden escindirse usando peryodato y propanolamina como se describe y los polinucleótidos monohebra resultantes hibridaron como se ha descrito.
- 15

- [0156]** Los cebadores injertados contenían una secuencia de bases de T en el extremo 5' para que actuara como grupo separador para ayudar a la linealización y la hibridación. La síntesis del fosforamidita de diol se detalla a continuación. Las secuencias de los dos cebadores injertados al chip están diseñadas para permitir el tratamiento adecuado para liberar el resto hidroxilo libre en 3' y contienen cada una la base U de la siguiente manera:
- 20

- P5diol: 5' PS-TTTTTTTTTT-diol-AATGATACGGCGACCACCGA (SEC ID N° : 1)
- 25 P7GAU: 5' PS-TTTTTTTTTTCAAGCAGAAGACGGCATACGAGAU (SEC ID N° : 2)

- [0157]** Los mismos cebadores pueden permitir el corte con BstNB1, suponiendo que las muestras a ser amplificadas contienen el sitio de reconocimiento correspondiente de la siguiente manera:

- 30 **5'...GAGTCNNNNT^YN...3'**
3'...CTCAGNNNNN...5'

- [0158]** Para permitir que el corte con la enzima de restricción de corte remoto Mme1; los extremos de los dos cebadores de injerto contenían la secuencia siguiente:
- 35

5'...TCCGAC-3'

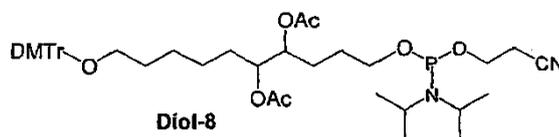
- [0159]** Para permitir que el corte con la enzima de restricción de corte remoto EcoP15; los extremos de los dos cebadores de injerto contenían la secuencia siguiente:
- 40

5'...CAGCAG-3'

- [0160]** El diseño de los cebadores se muestra en las Figuras 5 y 6:

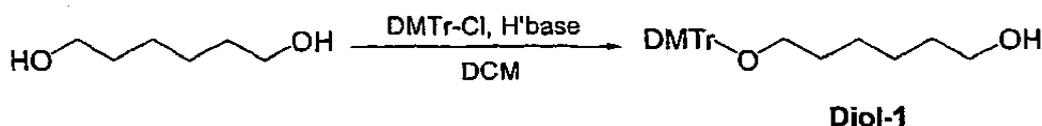
- 45 **Ejemplo 4: Preparación de diol-fosforamidita para el acoplamiento de ADN**

[0161]



50

Etapa 1:



- [0162]** Se disuelve 1,6-hexanodiol (Sigma Aldrich 99%) (14,6 g, 124 mmol), *N,N*-diisopropiletilamina (base de Hünig; Sigma Aldrich; redestilada) (21,6 ml, 124 mmol) en DCM/DMF anhidro (250/50 ml) bajo N₂. La disolución se enfría

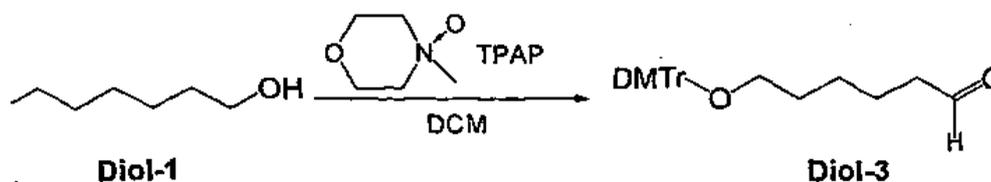
hasta 0 °C y se añade la primera porción de cloruro de 4,4'-dimetoxitritilo (DMTr-Cl; Sigma-Aldrich 95%) (10,5 g, 31 mmol). A continuación, la mezcla de reacción se calienta hasta la temperatura ambiente. Después de agitar durante 1 h, la mezcla de reacción se enfría hasta 0 °C de nuevo y se añade la segunda porción de DMTr-Cl (10,5 g, 31 mmol) y después se deja agitando a temperatura ambiente durante otras 2 horas. El análisis mediante TLC (EtOAc:éter de petróleo 4:6) indica un consumo de aproximadamente un 95% del derivado del material inicial (DMTr-OH). La reacción se concentra a presión reducida y se vierte una disolución acuosa de NaHCO₃ (sat.) (500 ml) en el residuo. La mezcla resultante se extrae con éter de petróleo/EtOAc (2:1) (3 x 1.000 ml). Las capas orgánicas combinadas se secan sobre MgSO₄ y se concentran a vacío. El residuo se coevapora con xileno (2 x 100 ml) para eliminar la DMF. La mezcla de reacción se preabsorbe en gel de sílice y se somete a una cromatografía ultrarrápida usando disolventes que contienen un 1% de Et₃N desde éter de petróleo hasta éter de petróleo/EtOAc (7:3) como eluyente. El rendimiento del aceite amarillo claro es de 16,58 g, 64%, con 7,8 g (17%) adicionales del subproducto bis-tritilado.

[0163] TLC: R_f: 0,35 (diol-1); R_f: 0,7 (subproducto bis-tritilado) (éter de petróleo/EtOAc 6:4).

15

[0164] RMN de ¹H (400 MHz, CDCl₃): δ 1,32-1,44 (m, 4H, 2 x CH₂), 1,54 -1,68 (m, 4H, 2x CH₂), 3,06 (t, J = 6,6 Hz, 2H, CH₂O), 3,62- 3,68 (m, 2H, CH₂OH), 3,81 (s, 6H, 2 x MeO), 6,83- 6,85 (m, 4H, Ph), 7,24-7,35 (m, 7H, Ph), 7,45-7,47 (m, 2H, Ph).

Etapa 2:



20

[0165] A una disolución del diol-1 (16,6 g, 39,5 mmol) en DCM anhidro (200 ml), se añade perrutenato de tetrapropilamonio (TPAP; Sigma Aldrich 97%) (277 mg, 0,79 mmol) bajo una atmósfera de N₂. La disolución se enfría hasta 0°C y se añade N-óxido de N-metilmorfolina (Sigma Aldrich 97%) (2,7 g, 23 mmol). La reacción se calienta hasta la temperatura ambiente. Después de 1 hora, se añaden las otras tres porciones de N-óxido de N-metilmorfolina (3 x 2,0 g, 51,2 mmol) durante un periodo de cuatro horas. La TLC (EtOAc:éter de petróleo 4:6) indica que la reacción sigue hasta su finalización. La reacción se extingue con NaHCO₃ (sat.) acuoso (1000 ml) y se extrajo con CH₂Cl₂ (4 x 1000 ml). Las capas orgánicas combinadas se secan sobre MgSO₄. La disolución se concentra a presión reducida. El diol-3, 9,9 g, 60%, se aísla mediante cromatografía ultrarrápida usando disolventes que contienen un 1% de Et₃N desde éter de petróleo hasta éter de petróleo/EtOAc (6:4) como eluyente, como un aceite amarillo claro.

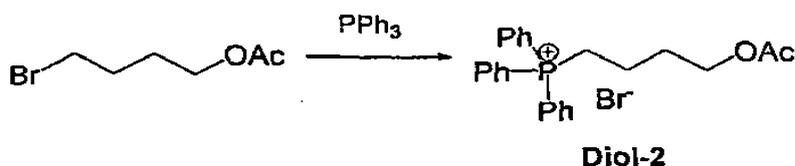
30

[0166] TLC: R_f: 0,7 (éter de petróleo/EtOAc 6:4).

[0167] RMN de ¹H (400MHz, CDCl₃): δ 1,30-1,37 (m, 2H, CH₂), 1,48-1,57 (m, 4H, 2 x CH₂), 2,34 (td, J = 1,7 y 7,4 Hz, 2H, CH₂CHO), 2,97 (s, 2H, CH₂O), 3,72 (s, 6H, 2 x MeO), 6,73-6,76 (m, 4H, Ph), 7,10- 7,26 (m, 7H, Ph), 7,34-7,36 (m, 2H, Ph), 9,67 (t, J = 1,7, 1H, CHO).

35

Etapa 3:



40

[0168] Se calienta una disolución de trifetilfosfina (Sigma-Aldrich 99%, ReagentPlus™) (39,3 g, 150 mmol) y acetato de 4-bromobutano (Sigma-Aldrich) (26 ml, 180 mmol) en tolueno anhidro (300 ml) a reflujo durante 36 horas bajo N₂ en un baño de aceite (140 °C). Durante el reflujo el aceite precipita. La mezcla de reacción se enfría hasta la temperatura ambiente. El análisis mediante TLC (éter de petróleo/EtOAc 7:3) de la disolución de tolueno demostró que todavía había trifetilfosfina (R_f: 0,8). El sobrenadante se decanta en otro matraz de fondo redondo y se concentra hasta un volumen aproximado de 30 ml. La disolución se calienta a reflujo de nuevo durante otras 12 horas. El sobrenadante se decanta. Las porciones de aceite se combinan entre sí, se disuelven en agua (500 ml) y se extraen con EtOAc (2 x 500 ml). Las capas orgánicas combinadas se retroextraen con agua (150 ml). Se

45

combinan dos lotes de capas acuosas, se evaporan a presión reducida. El residuo resultante se coevapora con acetonitrilo (2 x 100 ml) para dar 78,4 g, 95% de rendimiento de un aceite amarillo claro. La RMN indica que el producto era puro y se usa para la siguiente etapa de reacción sin purificación adicional.

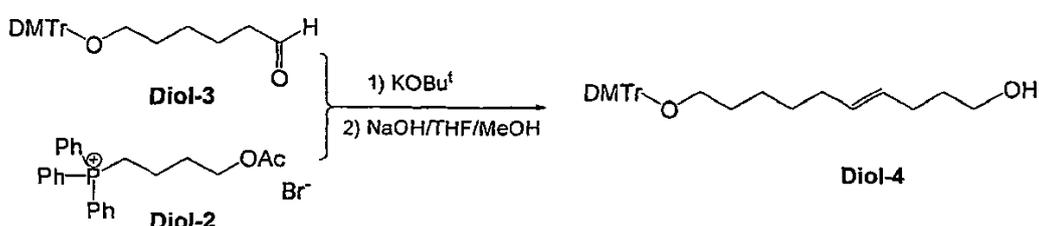
5 [0169] TLC: R_f : 0,0 (éter de petróleo/EtOAc 7:3).

[0170] RMN de ^1H (400 MHz, CDCl_3) : δ 1,63- 1,73 (m, 2H, CH_2), 1,94 (s, 3H, 2 x CH_3), 2,06-2,16 (m, 2H, CH_2), 3,97-4,05 (m, 2H, CH_2P), 4,11 (t, $J = 6,0$, 2H, CH_2O), 7,69-7,95 (m, 15H, Ph) .

10 [0171] RMN de ^{31}P (162 MHz, CDCl_3) : 25,9 ppm.

[0172] Detalles del espectro de masas: CL-EM (Electronebulización positivo) : (M^+) 377.

Etapa 4:



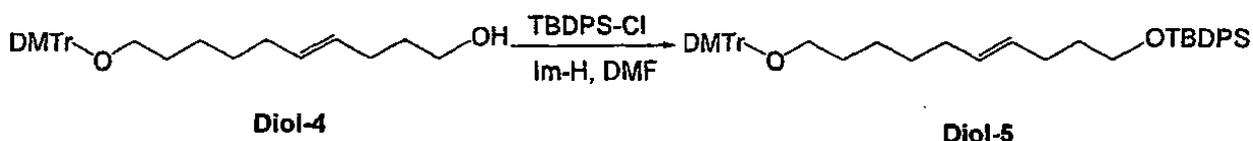
15

[0173] Se pesa el diol-2 (10,34 g, 22,7 mmol) en un matraz de fondo redondo y se disuelve con DCM (20 ml). La disolución se evapora a continuación a presión reducida hasta que da una espuma blanca. A continuación, el matraz se somete a alto vacío durante 24 h. A este matraz se añade THF anhidro (180 ml) bajo N_2 . La suspensión resultante se enfría hasta -78°C con un baño de acetona-hielo seco. Con agitación vigorosa se añade KOBut (3,3 g, 29,5 mmol) bajo N_2 . Lentamente el color de la suspensión se vuelve naranja, y gradualmente precipitan sólidos blancos. A esta suspensión se añade gota a gota una disolución del diol-3 (secado a 60°C a alto vacío durante 1 h antes de la reacción), (9,5 g, 22,7 mmol) en THF (50 ml) durante media hora. A continuación, se retira el baño de acetona-hielo seco. La mezcla de reacción se calienta lentamente hasta la temperatura ambiente y se agita durante otra hora. El color de la mezcla de reacción se vuelve amarillo tras la adición del diol-3. La mezcla de reacción se concentra a presión reducida. El residuo resultante se fracciona entre DCM (800 ml) y NaCl (sat.) acuoso (800 ml). La capa acuosa se extrae con DCM adicional (2 x 800 ml). Las extracciones orgánicas se combinan, se secan sobre MgSO_4 , se filtran y se evaporan a presión reducida para dar un aceite amarillo. El aceite se disuelve en THF/MeOH (125/100 ml) y se enfría hasta 0°C . A esta disolución se añade NaOH (1M en H_2O , 25 ml). Después de dejar la reacción en agitación durante 1 hora, el análisis mediante TLC indica un consumo total del material de partida. La mezcla de reacción se neutraliza con ácido acético (1,5 ml). La mezcla de reacción se concentra a presión reducida. El residuo resultante se fracciona entre DCM (800 ml) y NaHCO_3 (sat.) acuoso (800 ml). La capa acuosa se extrae con DCM adicional (2 x 800 ml). Las extracciones orgánicas se combinaron, se secaron sobre MgSO_4 , se filtran y se evaporan para dar un aceite amarillo claro. El diol-4,6, 45 g, 60% se aísla mediante cromatografía ultrarrápida usando disolventes que contienen un 1% de Et_3N desde éter de petróleo hasta éter de petróleo/EtOAc (5:5) como eluyente, como un aceite amarillo claro.

[0174] TLC: $R_f = 0,45$ (éter de petróleo/EtOAc 6:4).

[0175] RMN de ^1H (400 MHz, CDCl_3) δ 1,24-1,32 (m, 4H, 2 x CH_2), 1,54 -1,57 (m, 4H, 2 x CH_2), 1,93-1,96 (m, 2H, CH_2), 2,02- 2,07 (m, 2H, CH_2), 2,96 (t, $J = 6,6$ Hz, 2H, CH_2O), 3,54-3,59 (m, 2H, CH_2OH), 3,72 (s, 6H, 2 x MeO), 5,29-5,32 (m, 2H, 2 x =CH), 6,73-6,77 (m, 4H, Ph), 7,11-7,27 (m, 7H, Ph), 7,36-7,38 (m, 2H, Ph) .

Etapa 5:



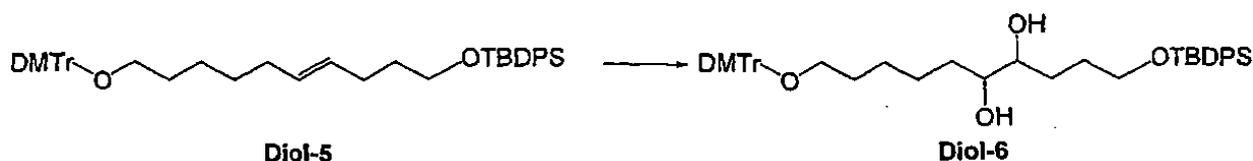
45 [0176] A una disolución del diol-4 (5,68 g, 12 mmol) e imidazol (Sigma Aldrich, 99%), (1,63 g, 24 mmol) en DMF anhidra (100 ml), se añade gota a gota cloruro de t-butildifenilsililo (Sigma Aldrich, 98%), (4,05 ml, 15,6 mmol) bajo una atmósfera de N_2 a temperatura ambiente. La reacción se agita durante 1 hora. La TLC (éter de petróleo/EtOAc

8:2) indica que el material de partida se ha consumido completamente. Se añade una disolución acuosa saturada de NaHCO_3 (500 ml) para extinguir la reacción. La mezcla resultante se extrae con éter de petróleo/EtOAc (2:1) (3 x 500 ml). Las capas orgánicas se combinan, se secan sobre MgSO_4 , se filtran y se evaporan para dar un aceite amarillo. El diol-5, 8,14 g, 95%, se aísla mediante cromatografía ultrarrápida usando disolventes que contienen un 1% de Et_3N desde éter de petróleo hasta éter de petróleo/EtOAc (9:1) como eluyente, como un aceite incoloro.

[0177] TLC: $R_f = 0,7$ (éter de petróleo : EtOAc 8:2) .

[0178] RMN de ^1H (400 MHz, CDCl_3) : δ 0,97 (s, 9H, 3 x Me), 1,19-1,30 (m, 4H, 2 x CH_2), 1,48-1,55 (m, 4H, 2 x CH_2), 1,91-1,95 (m, 2H, CH_2), 2,01-2,06 (m, 2H, CH_2), 2,95 (t, $J = 6,6$ Hz, 2H, CH_2O), 3,58 (t, $J = 6,3$ Hz, 2H, CH_2O), 3,70 (s, 6H, 2 x MeO), 5,24-5,27 (m, 2H, 2 x =CH), 6,72-6,75 (m, 4H, Ph), 7,11-7,37 (m, 15H, Ph), 7,57-7,60 (m, 4H, Ph).

Etapa 6:

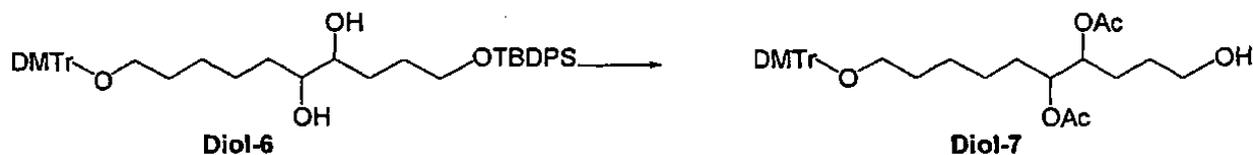


15 [0179] Se agita conjuntamente vigorosamente una mezcla de diol-5 (9,27 g, 13 mmol), AD-mix- α (Sigma Aldrich), (18,2 g), metansulfonamida (Sigma Aldrich, 97%), (1,23 g, 13 mmol), t-BuOH (65 ml) y agua (65 ml) a 55 °C durante 14 h. El análisis mediante TLC (éter de petróleo : EtOAc 6:4) indica aproximadamente un consumo de un 95% del material de partida. La mezcla de reacción se enfría hasta la temperatura ambiente, se trata con sulfito sódico (15,3 g, 12 mmol), después se agita adicionalmente durante 30 min. Se añade a la reacción una disolución acuosa saturada de NaHCO_3 (500 ml). La mezcla resultante se extrae con EtOAc (3 x 500 ml). Las capas orgánicas se combinan, se secan sobre MgSO_4 , se filtran y se evaporan para dar un aceite amarillo. El diol-6, 7,96 g, 82%, se aísla mediante cromatografía ultrarrápida (gel de sílice, Fluka, malla de 70-230) usando disolventes que contienen un 1% de Et_3N desde éter de petróleo hasta éter de petróleo/EtOAc (1:1) como eluyente, como un sólido blanco.

25 [0180] TLC: $R_f = 0,3$ (éter de petróleo : EtOAc 6:4) .

[0181] RMN de ^1H (400 MHz, CDCl_3) δ 1,07 (s, 9 H, 3 x Me), 1,41-1,7 (m, 12 H, 6 x CH_2), 1,94 (d, $J = 4,3$ Hz, 1H, OH), 2,94-2,95 (m, 1H, OH), 3,06 (t, $J = 6,6$ Hz, 2H, CH_2O), 3,61-3,63 (m, 2H, 2 x CHOH), 3,73 (t, $J = 5,6$ Hz, 2H, CH_2O), 3,81 (s, 6H, 2 x MeO), 5,24-5,27 (m, 2H, 2 x =CH), 6,82-6,85 (m, 4H, Ph), 7,21-7,47 (m, 15 H, Ph), 7,57-7,60 (m, 4H, Ph) .

Etapa 7:



[0182] A una disolución del diol-6 (7,96 g, 13 mmol) y DMAP (Sigma-Aldrich ReagentPlus™, 99%) (260 mg, 2,13 mmol) en una mezcla de piridina (15 ml) y DCM (30 ml), se añade anhídrido acético (Fluka 99%), (2,5 ml, 26,68 mmol) a temperatura ambiente. El análisis mediante TLC (éter de petróleo:EtOAc 6:4) indica el consumo total del material de partida después de 1 h. La reacción se extingue con una disolución acuosa saturada de NaHCO_3 (500 ml). Después de 5 min, la mezcla se extrae con DCM (3 x 500 ml). Las capas orgánicas se combinan, se secan sobre MgSO_4 , se filtran y se evaporan. El residuo se coevapora con tolueno (2 x 100 ml). El aceite amarillo resultante se somete a un tapón de gel de sílice (50 g, Fluka, malla de 70-230) para eliminar la DMAP usando eluyentes que contienen un 0,1% de Et_3N desde éter de petróleo hasta éter de petróleo/EtOAc (7:3) (250 ml de cada uno). Las fracciones combinadas de producto se concentran a sequedad. El aceite incoloro resultante se disuelve en THF (100 ml) y se trató con TBAF (Sigma-Aldrich; 5% en agua), (1 M en THF, 15 ml) a 0 °C. La disolución de reacción se calienta lentamente hasta la temperatura ambiente y se agita durante 2 horas. El análisis mediante TLC (éter de petróleo:EtOAc 6:4) indica que la desililación se ha completado. El disolvente volátil (THF) se evapora a presión reducida a baja temperatura. Se añade una disolución acuosa saturada de NaHCO_3 (500 ml) al residuo. La mezcla resultante se extrae con EtOAc (3 x 500 ml). Las capas orgánicas se combinan, se secan sobre MgSO_4 , se filtran y se evaporan para dar un aceite amarillo. El diol-7, 4,2 g, 66%, se aísla mediante cromatografía ultrarrápida

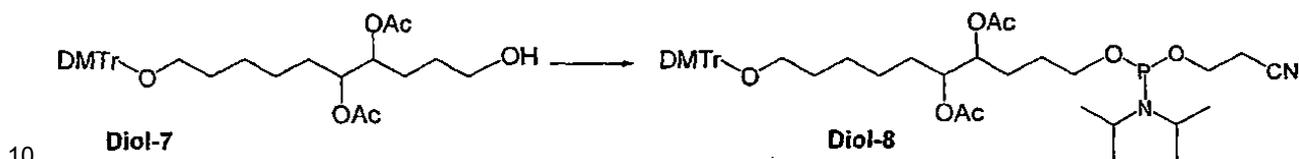
usando disolventes que contienen un 1% de Et₃N desde éter de petróleo hasta éter de petróleo/EtOAc (1:1) como eluyente, como un sólido blanco.

[0183] TLC: R_f = 0,45 (éter de petróleo/EtOAc 1:1) .

5

[0184] RMN de ¹H (400 MHz, CDCl₃) δ 1,29-1,33 (m, 4H, 2 x CH₂), 1,47-1,63 (m, 8 H, 4 x CH₂), 1,99, 2,01 (2s, 6H, 2 MeC(O)), 3,00 (t, J = 6,5 Hz, 2H, CH₂O), 3,50-3,64 (m, 2H, CH₂O), 3,75 (s, 6H, 2 x MeO), 4,92-4,97 (m, 2H, 2 x CHOAc), 6,76-6,80 (m, 4H, Ph), 7,15-7,29 (m, 7 H, Ph), 7,38- 7,40 (m, 2 H, Ph) .

Etapa 8:



[0185] A una disolución del diol-7 (2,08 g, 3,5 mmol) y diisopropiletilamina (Sigma Aldrich), (1,53 ml, 8,75 mmol) en DCM (17 ml), se añade gota a gota *N,N*-diisopropilclorofosforamidita de 2-cianoetilo (1,0 g, 4, 2 mmol) a temperatura ambiente bajo N₂. Después de agitar durante 1 hora, el análisis mediante TLC (éter de petróleo/EtOAc 4:6) indica un consumo total del material de partida. El disolvente (THF) se concentra a presión reducida. El residuo resultante se somete directamente a cromatografía. El diol-8, 2,5 g, 90%, se aísla mediante cromatografía ultrarrápida usando disolventes que contiene un 1% de Et₃N desde éter de petróleo hasta éter de petróleo/EtOAc (1:1) como eluyente, como un jarabe incoloro.

20 [0186] TLC: R_f = 0,55 (éter de petróleo/EtOAc 4:6).

[0187] RMN de ¹H (400 MHz, CDCl₃) δ 1,09, 1,10, 1,11, 1,12 (4 s, 12 H, N(CHMe₂)₂), 1,26-1,31 (m, 4H, 2 x CH₂), 1,45-1,56 (m, 8H, 4 x CH₂), 1,95, 1,969, 1,971, 1,98 (4 s, 6H, 2 MeCO), 2,56 (t, J = 6,5 Hz, 2H, CH₂CN), 2,95 (t, J = 6,5 Hz, 2H, CH₂O), 3,49-3,55 (m, 4 H, CH₂O), 3,72 (s, 6H, 2 x MeO), 4,89-4,92 (m, 2H, 2 x CHOAc), 6,74-6,76 (m, 4H, Ph), 7,13-7,25 (m, 7 H, Ph), 7,34-7,37 (m, 2 H, Ph).

25

[0188] RMN de ³¹P (162 MHz, CDCl₃): 148,67, 148,69 ppm.

Ejemplo 5: Preparación de agregados por amplificación isotérmica

30

Etapa 1: Hibridación y amplificación

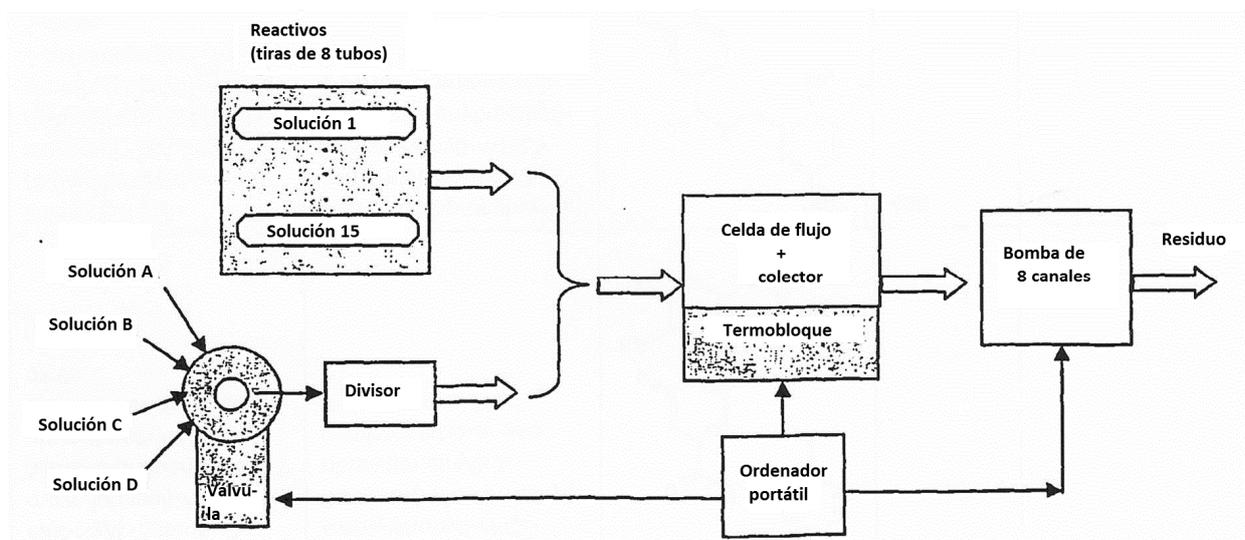
[0189] La secuencia de ADN usada en el proceso de amplificación es una secuencia única monomolde de 240 bases, con extremos complementarios a los cebadores injertados. La secuencia completa de una hebra del dúplex de molde se muestra en la Figura 6. El dúplex de ADN (1 nM) se desnaturaliza usando un tratamiento con hidróxido sódico 0,1 M seguido de una rápida dilución a la "concentración de trabajo" de 0,2-2 pM deseada en "tampón de hibridación" (5 x SSC/Tween al 0,1%).

35

[0190] La amplificación de superficie se llevó a cabo mediante amplificación isotérmica usando un termociclador MJ Research, acoplado a una bomba peristáltica de 8 vías Ismatec IPC ISM931 equipada con tuberías Ismatec (naranja/amarillo, de 0, 51 mm de DI). A continuación se muestra un esquema del instrumento.

40

[0191] El molde monohebra se hibrida con los



5 cebadores injertados inmediatamente antes de la reacción de amplificación, que por lo tanto comienza con una etapa inicial de extensión del cebador más que con la desnaturalización del molde. El procedimiento de hibridación comienza con una etapa de calentamiento en tampón riguroso para asegurar la completa desnaturalización antes de la irritación. Después de la hibridación, que se produce durante una lenta etapa de enfriamiento de 20 min, la celda de flujo se lavó durante 5 minutos con un tampón de lavado (0,3 x SSC/Tween al 0,1%).

10 **[0192]** En la siguiente tabla se detalla un proceso de amplificación típico, detallando los volúmenes de flujo por canal:

1. Hibridación del molde y 1ª extensión

15 **[0193]**

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
1	Bombeo de la premezcla de hibridación	20	120	60	120
2	Bombeo de la mezcla de hibridación	98,5	300	15	75
3	Eliminar las burbujas	98,5	10	100	16,7
4	Detener el flujo y mantener la T	98,5	30	estática	0
5	Enfriamiento lento	98,5-40,2	19,5 min	estática	0
6	Bombeo del tampón de lavado	40,2	300	15	75
7	Bombeo de la premezcla de amplificación	40,2	200	15	50
8	Bombeo de la mezcla de amplificación	40,2	75	60	75
9	Primera extensión	74	90	estática	0
10	Enfriar a temperatura ambiente	20	0	estática	0

2. Amplificación isotérmica

[0194]

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
(1) Esta secuencia 35 veces	Bombeo de formamida	60	75	60	75
	Bombeo de la premezcla de amplificación	60	75	60	75
	Bombeo de la mezcla Bst	60	95	60	95
	Detener el flujo y mantener la T	60	180	Estática	0
2	Bombeo del tampón de lavado	60	120	60	120

5

[0195] Premezcla de hibridación (tampón) = 5 x SSC/Tween al 0,1%

[0196] Mezcla de hibridación = muestra de hidróxido de ADN 0,1 M, diluido en premezcla de hibridación

10

[0197] Tampón de lavado = 0, 3 x SSC/Tween al 0, 1%

[0198] Premezcla de amplificación = betaína 2 M, Tris 20 mM, sulfato amónico 10 mM, sulfato magnésico 2 mM, Triton al 0,1%, DMSO al 1, 3%, pH 8, 8.

15

[0199] Mezcla de amplificación = betaína 2 M, Tris 20 mM, sulfato amónico 10 mM, sulfato magnésico 2 mM, Triton al 0, 1%, DMSO al 1,3%, pH 8,8 más mezcla de dNTP 200 µM y 25 unidades/ml de polimerasa Taq (Producto NEB ref. M0273L).

20 **[0200]** Mezcla de Bst = Mezcla de amplificación = betaína 2 M, Tris 20 mM, sulfato amónico 10 mM, sulfato magnésico 2 mM, Triton al 0, 1%, DMSO al 1,3%, pH 8,8 más mezcla de dNTP 200 µM y 80 unidades/ml de polimerasa Bst (Producto NEB ref. M0275L).

Etapa 2: Bloqueo de los grupos 3'-OH extensibles

25

[0201] Para preparar la premezcla de bloqueo se mezclaron 1.530 µl de agua y 170 µl de tampón de bloqueo 10 X (tampón NEB 4; número de producto B7004S) hasta un volumen final de 1.700 µl. Para preparar la mezcla de bloqueo se mezclaron 1.065,13 µl de premezcla de bloqueo, 21,12 µl de mezcla de ddNTP 125 µM y 13, 75 µl de transferasa terminal de TdT (NEB; parte nº M0252S) hasta un volumen final de 1.100 µl.

30

[0202] Para bloquear el ácido nucleico dentro de los agregados formados en los canales de las celdas de flujo, el componente informático de la instrumentación bombeó el apropiado tampón de bloqueo a través de la celda de flujo, y controló la temperatura según se muestra en las formas de realización ejemplares, a continuación.

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
1	Bombeo de la premezcla de bloqueo	20	200	15	50
2	Bombeo de la mezcla de bloqueo	37,7	300	15	75
3	Detener el flujo y mantener la T	37,7	20	estática	0
4	Bombeo cíclico de la mezcla de bombeo y esperar	37,7	8 x (20+180)	15/estática	45
5	Bombeo del tampón de lavado	20	300	15	75

35

Ejemplo 6: Tratamiento de agregados para obtener un grupo hidroxilo en 3'

Procedimiento a: USER™

40 **[0203]** Los agregados amplificados fueron tratados en la plataforma MJ a las siguientes condiciones; flujo con

premezcla USER (KCl 10 mM, Tris-HCl 20 mM (pH 8,8), (NH₄)₂SO₄ 10 mM, MgSO₄ 2 mM, 0,1% de Triton X-100) durante 5 min a 38 °C; flujo con tampón USER (premezcla USER más 10 unidades/ml de mezcla de enzimas USER (NEB cat # M5505) durante 5 minutos y se incubó durante 45 min a 38 °C; lavar con tampón de lavado durante 5 minutos.

5

[0204] Las hebras escindidas se trataron con exonucleasa de T7 (100 unidades/ml en tampón NEB 4 (1x NEB4 = acetato de potasio 50 mM, Tris-acetato 20 mM, acetato de magnesio 10 mM, DTT 1 mM pH 7,9) durante 30 min a 30 °C y enjuagado con tampón de lavado.

10 **[0205]** Las hebras se extendieron por ligación del cebador de ligación SBS 5:

3'-GTAGCTGAGCCAAGTCGTCCTTACGGCTCTGGCT-PO₄-5' (SEC ID N° 3)

15 **[0206]** El cebador (1 µM) en la premezcla de hibridación se hizo fluir en la célula a 60 °C y se dejó hibridar durante 5 minutos. La célula se lavó con tampón de ligación de T4 y se enfrió a 30 °C. Se añadió ADN ligasa de T4 y la reacción se mantuvo durante 30 minutos antes de lavar con tampón de lavado. El extremo 3' del cebador SBS 5 ligado hibridado está disponible para posteriores reacciones de extensión.

Procedimiento b: Nt.BstNBI

20

[0207] Los grupos amplificados se trataron en la plataforma MJ con las siguientes condiciones; las celdas de flujo se lavaron con tampón NEB 3 a 55 °C durante 5 min a 15 µl/min y después se trató con 50 µl de Nt.BstNBI (1 X tampón NEB 3; conc. final de la enzima 1.000 unidades/ml). La célula se incubó durante 1 hora a 55 °C y después se bajó a 20 °C, se enjuagó con tampón NEB 3 (5 min a 15 µl/min), a continuación, con tampón de lavado (5 min a 15 µl/min).

25

[0208] Las hebras escindidas se trataron con exonucleasa de T7 (100 unidades/ml en tampón NEB 4 (1x NEB4 = acetato de potasio 50 mM, Tris-acetato 20 mM, acetato de magnesio 10 mM, DTT 1 mM pH 7,9) durante 30 min a 30 °C y se enjuagó con tampón de lavado.

30 **Ejemplo 7: Secuenciación a partir de un cebador inmovilizado**

[0209] La secuenciación de los agregados del protocolo ilustrativo anterior se llevó a cabo usando nucleótidos modificados preparados según se describe en la solicitud de patente internacional WO 2004/018493 y se marcaron con cuatro fluoróforos espectralmente diferentes, como se describe en la solicitud US número 60/801270; presentada el 18 de mayo de 2006. La secuenciación de los agregados se describe con más detalle en la publicación de patente internacional W006064199.

35

[0210] Se usó una enzima mutante de la polimerasa 9°N (una variante exo que incluye la triple mutación L408Y/Y409A/P410V y C223S) para las etapas de incorporación de nucleótidos.

40

[0211] La mezcla de incorporación, el tampón de incorporación (Tris-HCl 50 mM pH 8,0, MgSO₄ 6 mM, EDTA 1 mM, Tween-20 al 0,05% (v/v), NaCl 50 mM) más YAV exo-C223S 110 nM y 1 µM de cada uno de los cuatro nucleótidos marcados modificados, se aplicó a los moldes agregados y se calentó a 45°C.

45 **[0212]** Los moldes se mantuvieron a 45 °C durante 30 min, se enfriaron hasta 20 °C y se lavaron con tampón de incorporación, después con 5 x SSC/Tween 20 al 0,05%. A continuación, los moldes se expusieron a tampón de imagen (Tris 100 mM pH 7,0, NaCl 30 mM, Tween 20 al 0,05%, ascorbato sódico 50 mM, recién disueltos). Los moldes se escanearon a 4 colores a temperatura ambiente. Los moldes se expusieron entonces a ciclos de secuenciación de Escisión e Incorporación como sigue:

50

Escisión: Cebador con tampón de Escisión (Tris 0,1 M pH 7,4, NaCl 0,1 M y Tween 20 al 0,05%). 125 µl/canal; 60 µl/min

Calentar hasta 60°C.

55

Tratar los agregados con mezcla de Escisión (TCEP 100 mM en tampón de Escisión). 75 µl/canal; 60 µl/min.

60

Esperar un total de 15 min además de bombear mezcla de escisión nueva, 25 µl/canal; 60 µl/min, cada 4 min.

Enfriar hasta 20°C.

Lavar con tampón Enzimológico.

65

Lavar con 5 x SSC/Tween 20 al 0,05%.

Cebar con tampón de imagen.

5 Escanear a 4 colores a temperatura ambiente.

Incorporación: Cebar con tampón de Incorporación. 125 µl/canal; 60 µl/min. Calentar hasta 60 °C.

Tratar con mezcla de Incorporación. 75 µl/canal; 60 µl/min

10

Esperar un total de 15 min además de bombear mezcla de incorporación nueva 25 µl/canal; 60 µl/min, cada 4 min. Enfriar hasta 20°C.

Lavar con tampón de Incorporación. 75 µl/canal; 60 µl/min.

15

Lavar con 5 x SSC/Tween 20 al 0,05%. 75 µl/canal; 60 µl/min. Cebar con tampón de imagen, 100 µl/canal; 60 µl/min

Escanear a 4 colores a temperatura ambiente.

20

Repetir el proceso de Incorporación y Escisión durante tantos ciclos como sea necesario.

[0213] Los nucleótidos incorporados se detectaron usando un aparato de imagen fluorescente de CCD basado en reflexión interna total descrito en "Sistemas y dispositivos para la secuencia por análisis de síntesis", USSN 25 60/788,248, presentada el 31 de marzo de 2006.

[0214] Una imagen representativa que muestra dos ciclos de incorporación se muestra en la Figura 7. El primer ciclo de la incorporación (arriba a la derecha) muestra una imagen con 3271 agregados que han incorporado un trifosfato de G de acuerdo con la secuencia monomolde. El segundo ciclo (abajo a la derecha), muestra una imagen del segundo ciclo donde 3235 agregados han incorporado un trifosfato de A. La imagen principal muestra la imagen colocalizada, donde 3058 de los agregados son comunes a ambas imágenes, lo que demuestra que los agregados pueden pasar a través de ciclos de secuenciación a partir del cebador inmovilizado.

35 Ejemplo 8: Extensión con nucleótidos no marcados

[0215] Después de la etapa de desbloqueo final del protocolo de secuenciación, la matriz se sometió a una ronda de extensión con cuatro dNTP y polimerasa Bst, de forma análoga a un ciclo de extensión en el proceso de generación de agregados. La matriz se trató como sigue:

Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
Bombeo de la premezcla de amplificación	60	75	60	75
Bombeo de la mezcla Bst	60	95	60	95
Detener el flujo y mantener la T	60	180	Estática	0
Bombeo del tampón de lavado	60	120	60	120

40

Ejemplo 9: Tratamiento para escindir la región central de un agregado

[0216] Los agregados fueron tratados con Mme1, seguido de la polimerasa T4 para eliminar los salientes dinucleótidos y la ligasa T4 para unir los extremos lisos, romos en una secuencia de 36 pares de bases. Protocolos idénticos se han descrito en solución en construcciones de vectores circulares, (Nature Methods, 2, 105 - 111 (2005)) y los procedimientos se repitieron en los agregados amplificados, utilizando condiciones de tampón y las temperaturas, según lo recomendado para la enzima apropiada.

50 Ejemplo 10: linealización y la hibridación de un cebador de secuenciación

50

[0217] Para preparar la linealización, mezclar 1429 µl de agua, 64 mg de peryodato de sodio, 1500 µl de formamida, 60 µl de Tris 1 M pH 8 y 6011,4 µl de 3-aminopropanol hasta un volumen final de 3 ml. El peryodato se mezcla primero con el agua mientras que el Tris se mezcla con la formamida. Las dos soluciones se mezclan juntas y el 3-aminopropanol se añade a esa mezcla.

55

[0218] Para linealizar el ácido nucleico dentro de los agregados formados dentro de los canales de la celda de flujo, se hace fluir 300 µl por canal de mezcla linealización a 15 µl/min a 20 °C, seguido por 75 µl de agua a la misma velocidad de flujo.

[0219] Para preparar la mezcla de cebadores, se mezcla 895,5 µl de tampón de hibridación y 4,5 µl de cebador de secuenciación (100 mM) hasta un volumen final de 900 µl. La secuencia del cebador de secuenciación utilizado en esta reacción era:

5

5'-ACACTCTTTCCCTACACGACGCTCTTCCGATC (SEC ID N° 4)

[0220] Para desnaturalizar el ácido nucleico dentro de los agregados y para hibridar el cebador de secuenciación, los siguientes reactivos se hacen fluir a través de la célula:

10

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
1	Bombeo de NaOH 0,1 M	20	300	15	75
2	Bombeo de TE	20	300	15	75
3	Bombeo de la mezcla de cebadores	20	300	15	75
4	Mantener a 60 °C	60	300	0	0
5	Bombeo del tampón de lavado	40,2	300	15	75

Ejemplo 11: Secuenciación de un cebador no inmovilizado

[0221] El proceso de secuenciación se llevó a cabo exactamente de la misma manera como se describe en el Ejemplo 7. La segunda lectura se puede alinear con la primera, en el caso de dos lecturas, o el proceso de secuenciación se llevó a cabo durante 36 ciclos en el caso de tratamiento con Mme1. Esta única lectura generó 18 bases de información para cada extremo del fragmento original.

[0222] Mientras que la invención anterior se ha descrito con cierto detalle para fines de claridad y comprensión, será evidente para un experto en la materia a partir de una lectura de esta divulgación que se pueden hacer varios cambios en forma y detalle sin apartarse del verdadero alcance de la invención. Por ejemplo, todas las técnicas y aparatos descritos anteriormente pueden utilizarse en varias combinaciones.

Ejemplo 12: Lecturas por pares de una biblioteca de fragmentos

25

[0223] Los siguientes detalles experimentales describen la exposición completa de una realización de la invención como se describe anteriormente.

[0224] La biblioteca se fabricó utilizando ADN de BAC humano purificado (inserto de cromosoma humano de 140k clonado en un vector de pTARBAC). El ADN se preparó en primer lugar para la ligación a los adaptadores de horquilla por: fragmentación del ADN mediante nebulización, reparación final de los extremos de ADN para que sean extremos romos fosforilados y, a continuación, la adición de un único nucleótido 'A' en los extremos 3' de los fragmentos de ADN. La reacción de ligación se realizó con el ADN fragmentado preparado y adaptadores preformado por hibridación de 'Oligo A' y 'Oligo B' (secuencias dadas a continuación). El producto de la reacción se aisló/purificó a partir de adaptador no ligado por electroforesis en gel. Por último, el producto de la reacción de ligación se sometió a ciclos de PCR para amplificar selectivamente el producto ligado que contenía ADN genómico con adaptador en ambos extremos de los fragmentos.

Materiales y métodos

40

Etapa 1) Nebulización

[0225] Materiales:

- 45 • 0,5 µg/µl de BAC con ADN humano (inserto de cromosoma humano de 140k clonado en un vector de pTARBAC)
- Tampón de nebulización (53,1 ml de glicerol, 42,1 ml de agua, 3,7 ml de Tris HCl 1 M pH 7,5, 1,1 ml de 0,5 M EDTA)
- TE
- Nebulizadores (Invitrogen, K7025-05)
- 50 • Columnas del kit de purificación por PCR (Qiagen, 28104)

Procedimiento:

[0226] Se mezclaron 10 µl (5 mg) de ADN de BAC con 40 µl de TE y 700 µl de tampón de nebulización. Las soluciones de ADN enfriadas cada una fueron fragmentados en un nebulizador en hielo durante 6 minutos bajo 32 libras por pulgada cuadrada (psi) de presión. Los volúmenes recuperados se purificaron cada uno con una columna del kit de purificación de PCR Qiagen y se eluyeron en 30 µl de EB.

Etapa 2) Reparación de extremos

[0227] Materiales:

- 5
- ADN nebulizado (de la Etapa 1)
 - Agua
 - Tampón de ADN ligasa T4 con ATP 10 mM (10x) (NEB, B0202S)
 - Mezcla de dNTP (10 mM cada uno) (NEB, N0447S)
- 10
- ADN polimerasa T4 (3U/μl) (NEB, M0203L)
 - Fragmento grande de ADN Pol I de E. coli (Klenow) (5U/μl) - (NEB, M0210S)
 - Polinucleótido quinasa T4 (10U/μl) (NEB, M0201L)
 - Columnas del kit de purificación para PCR (Qiagen, 28104)

15 Procedimiento:

[0228] La mezcla para la reparación de extremos se dispuso de la siguiente manera:

ADN nebulizado	30 μl
Agua	45 μl
Tampón de ADN ligasa T4 con ATP 10 mM	10 μl
dNTPs	4 μl
ADN pol T4	5 μl
ADN pol Klenow	1 μl
PNK T4	<u>5 μl</u>
	100 μl total

[0229] La reacción se incubó durante 30 minutos a temperatura ambiente. El ADN se purificó en una columna de Qiagen, eluyendo en 30μl de EB.

Etapa 3) Reacción de adición de la cola de A

[0230] Materiales:

- 25
- ADN con extremo reparado (de la Etapa 2)
 - Agua
 - Tampón NEB 2 (10 veces) (NEB, B7002S)
 - dATP (1 mM) (Amersham-Pharmacia, 272050)
- 30
- Fragmento Klenow (3' a 5' exo menos) (5U/μl) (NEB, M0212B)
 - Bloque caliente o aparato de PCR
 - Columna del kit de purificación de PCR MinElute (Qiagen, 28004)

Procedimiento:

35 **[0231]** La mezcla de reacción se dispuso de la siguiente manera:

ADN reparado final	30 μl
Agua	2 μl
Tampón NEB 2	5 μl
dATP	10 μl
Fragmento Klenow (3 'a 5' exo menos)	<u>3 μl</u>
	50 μl total

[0232] La reacción se incubó durante 30 min a 37 °C, a continuación, el ADN se purificó en una columna Qiagen MinElute, eluyendo en 10 μl de EB.

Etapa 4) Adaptadores hibridados

[0233] Materiales:

- 45
- Oligo A: 5'ACACTCTTTCCCTACACGACGCTCTTCCGATC_xT (x = enlace fosforotioato) (SEC ID N° 5)
 - Oligo B: 5'fosfato-GATCGGAAGAGCGGTTCAGCAGGAATGCCGAG (SEC ID N° 6)
 - Tris/50 mM/NaCl 50 mM pH 7,0
 - Aparato de PCR
- 50

Procedimiento:

[0234] Los oligos se mezclaron hasta una concentración final de 15 mM cada uno, en Tris 10 mM/NaCl 10 mM pH 7,0. Las hebras del adaptador se hibridaron en un aparato de PCR programado de la siguiente manera: rampa en 0,5 °C/seg a 97,5 °C; Mantener a 97,5 °C durante 150 seg; entonces una etapa de 97,5 °C durante 2 seg con una caída de temperatura de 0,1 °C/ ciclo durante 775 ciclos.

Etapa 5) Ligación

[0235] Materiales:

- 10
- ADN genómico con cola de A (de la Etapa 3)
 - Tampón de ligasa rápida (2x) (NEB, B2200S)
 - Adaptador de hibridación (15 µM) (de 4).
 - Ligasa rápida (1U/µl) (NEB, M2200L)
- 15 • Columnas del kit de purificación de PCR (Qiagen, 28104)

Procedimiento:

[0236] La mezcla de reacción se dispuso de la siguiente manera:

20

ADN genómico con cola de A	10 µl
Tampón de ligasa rápida	25 µl
Adaptador hibridado	10 µl
Ligasa rápida	<u>5 µl</u>
	50 µl total

[0237] La reacción se incubó durante 15 min a temperatura ambiente, a continuación, el ADN se purificó en una columna Qiagen, eluyendo en 30 µl de tampón de elución (EB).

25 Etapa 6) Purificación en gel

[0238] Materiales:

- 30
- Reacción de ligación (de la Etapa 5)
 - Agarosa (Biorad, 161 a 3107)
 - TAE (50x)
 - Agua destilada
 - Bromuro de etidio (Sigma, E1510)
 - Tampón de carga (4x) (Tris pH 8 50 mM, EDTA 40 mM, sacarosa al 40% p/v)
- 35 • Escalera de bajo peso molecular (NEB, N3233L)
- Bandejas de gel y tanque. Unidad de electroforesis
 - Transiluminador lector oscuro (Clare Chemical Research, D195M)
 - Columnas del kit de extracción en gel (Qiagen, 28704)

40 Procedimiento:

[0239] Toda la muestra de la reacción de ligación purificada se cargó en un carril de un gel de agarosa al 2% que contiene bromuro de etidio y se procesó a 120 V durante 60 min. El gel se visualizó a continuación en una caja de "luz blanca" y se cortaron los fragmentos de 120 pb a 170 pb y se purificaron con una columna de extracción en gel, 45 eluyendo en 30µl de tampón de elución (EB).

Etapa 7) Tratamiento con exonucleasa I de los cebadores de PCR

[0240] Materiales:

- 50
- Exonucleasa I (E. coli) (20 U/µl) (NEB, M0293S)
 - Tampón de reacción de exonucleasa I (10x) (NEB, M0293S)
 - Agua
 - Cebadores de ADN con un fosforotioato en la posición n-1
- 55 • Columnas de Bio-Rad P6 (Bio-Rad, 732-6.221)

Procedimiento:

[0241] Los cebadores de ADN con un fosforotioato en la posición n-1 (5 x 85 µl de cada cebador (aproximadamente 25 µM)) se dividieron en partes alícuotas en tubos Eppendorf. Se añadieron 10 µl de 10 X tampón de reacción de

exonucleasa I y 5 µl de exonucleasa I a cada tubo. Cada tubo Eppendorf se colocó en una gradilla y se almacenó en un horno a 37 °C durante 16 horas. Después de 16 horas, los tubos se colocaron en un bloque caliente fijado en 80 °C durante 2 minutos. Posteriormente, se pasaron las soluciones de los tubos Eppendorf a través de columnas de Bio-Rad P6 y se centrifugaron en una centrifuga a 2000 rpm durante 2 minutos. Se añadió a las columnas 20 µl extra de H₂O y se volvieron a centrifugar. Las soluciones filtradas se colocaron en un SpeedVac® y se evaporaron hasta que cada uno tenía 20 µl y se combinaron las fracciones. Las fracciones reunidas se inyectaron en un sistema de HPLC de fase inversa y se recogió el pico principal. Las fracciones recogidas se evaporaron hasta sequedad en un SpeedVac®, se añadieron 50 µl de agua y la fracción fue sometida de nuevo a evaporación hasta sequedad. Se disolvieron los sedimentos resultantes en 50 µl de agua, se agruparon y se hizo una medición UV para determinar la concentración del oligonucleótido.

Etapa 8) PCR

[0242] Materiales:

- 15
- ADN purificado en gel (de la etapa 6)
 - Agua
 - Phusion master mix (2x) (NEB, F-531L)
 - Exonucleasa tratada con cebador universal de PCR 1 (25 µm):
- 20 5' AATGATACGGCGACCACCGAGATCTACTCTTTCCCTACACGACGCTCTTCCGA TCxT 3', donde x = enlace fosforotioato (de la Etapa 7) (SEC ID N° 7)
- Exonucleasa tratada con cebador universal de PCR 2 (25 µm):
- 5' CAAGCAGAAGACGGCATAACGAGATCGGTCTCGGCATTCTGCTGAACCGCTCTTC CGATCxT, donde x = enlace fosforotioato (de la Etapa 7) (SEC ID N° 8)
- 25
- Aparato de PCR
 - Columnas del kit de purificación de PCR (Qiagen, # 28104)

Procedimiento:

30 **[0243]** La reacción de PCR se preparó como sigue:

ADN en BAC purificado en gel	1 µl
Phusion mastermix	25 µl
Cebador de PCR universal 1	1 µl
Cebador de PCR universal 2	1 µl
Agua	<u>22 µl</u>
	50 µl total

[0244] El termociclado se llevó a cabo en una máquina de PCR en las siguientes condiciones:

- 35
- 30 segundos a 98 °C
 - [10 seg a 98 °C, 30 seg a 65 °C, 30 seg a 72 °C] 18 ciclos
 - 5 min a 72 °C
 - Mantener a 4 °C
- 40 **[0245]** Los productos de PCR se purificaron en una columna Qiagen, eluyendo en EB 30 µl. Las bibliotecas de ADN resultantes estaban listas para la amplificación en una plataforma de amplificación en superficie.

Validación de bibliotecas por secuenciación Sanger convencional

- 45 **[0246]** Cuatro (4) µl de las bibliotecas se clonaron en un vector plásmido (Zero Blunt TOPO kit de clonación de PCR, Invitrogen # K2800-20) y sembró en agar, de acuerdo con las instrucciones del fabricante. Las colonias se recogieron, se mini-prepararon y los fragmentos clonados se secuenciaron por secuenciación de Sanger convencional.
- 50 16 clones de la biblioteca BAC

[247]

1 (204pb Inserto: E.coli 85pb (SEC ID N° 9)

AATGATACGGCGACCACGGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCTACTGATTTCAATTGCAGCCAAAGGCAAACCTTTGGCTGCATCGTTTTA
CAGTCGCCATAAGCCTTTCCTCTGTAAACCGCCTTCTGAGATCGGAAGAGC
GGTTCAGCAGGAATGCCGAGACCGATCTCGTATGCCGTCTTCTGCTTG

2 (214pb) Inserto: BAC 95pb (SEC ID N° 10)

AATGATACGGCGACCACGGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCTATCAATATTGTGAAAATGACCATACTGCCAAAAAAAAAACTACAAATT
CAATGCAATTTTCATCAAAATACCATCATCATTCTTCACAATATTGATAGATC
GGAAGAGCGGTTTCAGCAGGAATGCCGAGACCGATCTCGTATGCCGTCTTCT
GCTTG

3 (215pb) Inserto: BAC 96pb (SEC ID N° 11)

AATGATACGGGGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTC
CGATCTCTCACTCCTGGCAGAGGGACGTGTGACTAGCCATGGGCCCTAG
GTCTCCAGTTCCTGGGTAGCTTGTATTTTTGAACATCTCCTGTATATTAGTTA
GATCGGAAGAGCGGTTGAGCAGGAATGCCGAGACCGATCTCGTATGCCGT
CTTCTGCTTG

4 (147 pb) Inserto: BAC 28pb (SEC ID N° 12)

AATGATACGGGGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTC
CGATCTAGTGTAGTTGAGATCTGCCTTAGCAGCAAGATCGGAAGAGCGGTT
GAGCAGGAATGCCGAGACCGATCTCGTATGCCGTCTTCTGCTTG

5 (183 pb) Inserto BAC 64pb (SEC ID N° 13)

AATGATACGGGGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTC
CGATCTAACACATTTCAAAGTTTGGGGCCCTCCTCCTCCCCAAAAACAAC
CACAAAAACAACAAAAAGATCGGAAGAGCGGTTGAGCAGGAATGCCGAG
ACCGATCTCGTATGCCGTCTTCTGCTTG

6 (170 pb) Inserto BAC 59pb (SEC ID N° 14)

GGCGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTCCGATCTG
AATGCCTTTTATAGCATTAAATTTTCCCTAAGTATAATTACCAAATAAAAATTG
TATAAGATCGGAAGAGCGGTTGAGCAGGAATGCCGAGACCGATCTCGTATG
CCGTCTTCTGCTTG

7 (180 pb) Inserto BAC 61pb (SEC ID n° 15)

AATGATACGGGGACCACCGAGATCTACACTCTTCCCTACACGACGCTCTTC
CGATCTTGGGCCCGGGAGGAGTTTGCCGGGGAGGAGTGGGTTTGAATCG
GGTTAAAGGAAAGAGAAGATCGGAAGAGCGGTTGAGCAGGAATGCCGAG
ACCGATCTCGTATGCCGTCTTCTGCTTG

8 (190 pb) Inserto: BAC 73pb (SEC ID N° 16)

TGATACGGGGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC
GATCTAAGATCTATTTCAAATGGACTGTAGATCTAAGTATAAAAGGTAAGAGA
ATAATTATTCTAGAAAGTAAATGTAAGATCGGAAGAGCGGTTGAGCAGGAAT
GGCGAGACCGATCTCGTATGCCGTCCTCTGCTTG

9 (192 pb) Inserto: BAC 74 pb (SEC ID N° 17)

AATGATACGGGGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC
GATCTGGGAGGCCAAGGTGGGTGGATCACCTGAGATCAGGAGTTCGAGAC
CAGCTGGCCAACATGATGAAACTCTGTCTAGATCGGAAGAGCGGTTGAGCA
GGAATGCCGAGACCGATCTCGTATGCCGTCCTCTGCTTG

10 (185 pb) Inserto: BAC 66pb (SEC ID N° 18)

AATGATACGGGGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC
CGATCTTGACCATTGTAACCATTAATGTAGACTGCAATGATATGCACTATTTA
CAACCTTTTTAAGACTCTAGATCGGAAGAGCGGTTGAGCAGGAATGCCGA
GACCGATCTCGTATGCCGTCCTCTGCTTG

11 (199 pb) Inserto: BAC 80pb (SEC ID N° 19)

AATGATACGGGGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC
CGATCTCTTTGAAGAGCTGGCAGTAGAAGATAAACAGGCTGGGGAAGAAGA
GAAAGTGCTCAAGGAGAAGGAGCAGCAGCAGCAGC
AGATCGGAAGAGCGGTTGAGCAGGAATGCCGAGACCGATCTCGTATGCCGTC
CTCTGCTTG

12 (212 pb) Inserto: BAC 93pb (SEC ID N° 20)

ATGATACGGGGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCC
GATCTAGTATTCAACAAGTCTGTCTTTTCCAAGTGTCTTTAAAGACCAGAAAT
ACCTGTTTTTAACACACAGGGTTGCAAATTCAGAGGAGATTGGCAGATCGG
AAGAGCGGTTGAGCAGGAATGCCGAGACCGATCTCGTATGCCGTCCTCTGCTG
CTC

13 (247pb) Inserto: E.coli 128pb (SEC ID Nº 21)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCTCTTGAGATGAGTGATGACGGCGCGCTGGAAGTTGCTCGTCGCGCT
CGCGGTACGCCGCGCATTGCCAACCGTCTGCTGCGTCGAGTGCGTGATTTC
GCCGAAGTGAAGCACGATGGCACCATCTCAAGAGATCGGAAGAGCGGTTC
GCAGGAATGCCGAGACCGATCTCGTATGCCGCTCTCTGCTTC

14 (202pb) Inserto: BAC 83 pb (SEC ID Nº 22)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCTGGGGTTGGTGAACCCAGATGCCTCCCAGGATTGGTGGGCCCTG
TGGCACTTGTACCTGCTGTTGCTGTTGCTGCTGCTGCTGAGATCGGAAGAG
CGGTTCAGCAGGAATGCCGAGACCGATCTCGTATGCCGCTCTCTGCTTC

15 (166pb) Inserto: BAC 47pb (SEC ID Nº 23)

AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTC
CGATCTATGATAAGGAGCAGGTTTACAGATCATAAGTGCAAAGCGGGCGA
GAAGATCGGAAGAGCGGTTCAGCAGGAATGCCGAGACCGATCTCGTATGC
CGTCTCTGCTTC

16 (147pb) Inserto: BAC 31pb (SEC ID Nº 24)

GATACGGCGAACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCGG
ATCTCTGATACTGTTGTAACCACCCAATTGGTTCAAGATCGGAAGAGCGGTTC
CAGCAGGAATGCCGAGACCGATCTCGTATGCCGCTCTCTGCTTC

[0248] Estos resultados confirman que el procedimiento de preparación de la biblioteca produce una biblioteca de moldes de ADN 'secuenciables'. Cada biblioteca contenía una pluralidad de insertos genómicos, cada uno de los cuales estaba flanqueado por los dos adaptadores

5 (AATGATACGGCGACCACCGAGATCTACACTCTTTCCCTACACGACGCTCTTCCGATCT SEC ID N° 25 y AGATCGAAGAGCGGTTTCAGCAGGAATGCCGAGACCGATCTCGTATGCCGTCTTCTGCTT SEC ID N° 26), necesario para la formación de agregados y secuenciación de SBS. El inserto de ADN de cada una de las bibliotecas se alineó a la referencia del BAC, con una pequeña cantidad de contaminación por E. coli.

10 **[0249]** Los agregados se prepararon a partir de la biblioteca anterior como se ha descrito en los Ejemplos 1-5 anteriores.

[0250] Los cebadores injertados a la superficie tenían la secuencia siguiente:

15 P5diol: 5' PS-TTTTTTTTTT-diol-AATGATACGGCGACCACCGA (SEC ID N° 1)
 P7-TU: 5' TTTTTTTTTTUCAAGCAGAAGACGGCATAACGA-OH (SEC ID N° 27)
 P5-fosfato: 5' PS-TTTTTTTTTTAATGATACGGCGACCACCGA-PO₄ (SEC ID N° 28)

[0251] Los tres cebadores se mezclaron a una concentración 1:1:1 y se utilizaron a una concentración 0,5 µM de cada cebador.

[0252] Los agregados se linealizaron utilizando el tratamiento de peryodato tal como se describe en el ejemplo 10 y se bloquearon con transferasa terminal como se describe en el ejemplo 5, etapa 2.

25 **[0253]** Los agregados fueron tratados de hibridar un primer cebador de secuenciación, como se describió anteriormente. La secuencia del cebador de secuenciación fue

5'-ACACTCTTTCCCTACACGACGCTCTTCCGATC (SEC ID N° 4)

30 **[0254]** Para desnaturalizar el ácido nucleico dentro de los agregados y para hibridar el cebador de secuenciación, los siguientes reactivos se hacen fluir a través de la celda como se muestra en la Tabla 4:

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
1	Bombeo de NaOH 0,1 M	20	300	15	75
2	Bombeo de TE	20	300	15	75
3	Bombeo de la mezcla de cebadores	20	300	15	75
4	Mantener a 60 °C	60	900	0	0
5	Bombeo del tampón de lavado	40,2	300	15	75

35 La secuenciación del fragmento diana se realizó utilizando los reactivos y procedimientos descritos anteriormente en el ejemplo 7.

Ejemplo 12a: Obtención de una segunda lectura

Etapa 1: Desbloqueo de los grupos fosfato con polinucleótido quinasa de T4

40 **[0255]** Los agregados fueron tratados como sigue: Se hizo fluir 1X tampón de reacción de intercambio (imidazol-HCl 50 mM, pH 6,4; MgCl₂ 12 mM; 2-mercaptoetanol 1 mM; ADP 70 µM) a 15 µl/canal/min durante 5 min a 20 °C. Esto fue seguido por la solución de PNK (0,2 U/µl de PNK en 1X tampón de reacción de intercambio) a 15 µl/canal/min durante 5 min a 20 °C. La celda de flujo se calentó a 38 °C durante 30 minutos y se dejó enfriar. Los canales se lavaron con tampón de lavado durante 5 minutos, a continuación, 5XSSC durante 5 minutos.

Etapa 2: Resíntesis de agregados.

50 **[0256]** Los agregados se trataron hasta 15 ciclos de amplificación isotérmica tal como se describe en el ejemplo 5; etapa 1

Etapa 3: Linealización con USER

[0257] Los agregados fueron tratados como se describe en el ejemplo 6 para escindir el resto de uracilo en el

cebador P7. Los agregados se bloquearon con transferasa terminal como se ha descrito en el ejemplo 5, etapa 2.

[0258] Los agregados fueron tratados para hibridar un primer cebador de secuenciación, como se describió anteriormente. La secuencia del cebador de secuenciación fue

5 5'-CGGTCTCGGCATTCTGCTGAACCGCTCTTCCGATC (SEC ID N° 29)

[0259] Para desnaturalizar el ácido nucleico dentro de los agregados y para hibridar el cebador de secuenciación, los siguientes reactivos se hacen fluir a través de la celda como se muestra en la Tabla 4:

10

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
1	Bombeo de NaOH 0,1 M	20	300	15	75
2	Bombeo de TE	20	300	15	75
3	Bombeo de la mezcla de cebadores	20	300	15	75
4	Mantener a 60 °C	60	900	0	0
5	Bombeo del tampón de lavado	40,2	300	15	75

La secuenciación del fragmento diana se realizó utilizando los reactivos y procedimientos descritos anteriormente en el ejemplo 7.

15 **[0260]** Los datos obtenidos a partir de la primera y segunda lecturas se muestran en la figura 8. Ambas lecturas se alinean claramente contra la secuencia de BAC, donde sólo un 4% de las lecturas derivan de la E. coli que contaminaba la muestra original. El tamaño medio de fragmento de los insertos fue de 80 pares de bases, por lo que para las dos lecturas de 25 pares de bases, la distancia entre los extremos de las lecturas fue de 30 pares de bases en promedio. La mayoría de los grupos de la primera lectura (8208/8572) (96%) también dio una buena lectura para la segunda lectura.

20

Ejemplo 13: Preparación y secuenciación del extremo emparejado usando dos cebadores inmovilizados modificados con uracilo y 8-oxo G, respectivamente.

25 Etapa 1) Injerto de cebadores sobre la superficie del chip recubierto con SFA

[0261] Un chip recubierto con SFA se coloca en un termociclador MJ-Research y se conecta a una bomba peristáltica. La mezcla de injerto que consiste en 0,5 µM de un cebador directo y 0,5 µM de un cebador inverso en tampón fosfato 10 mM (pH 7,0) se bombea en los canales del chip a un caudal de 60 µl/minuto durante 75 s a 20 °C.

30 El termociclador se calienta a continuación a 51,6 °C y el chip se incubó a esta temperatura durante 1 hora. Durante este tiempo, la mezcla de injerto se somete a 18 tiempos de bombeo: la mezcla de injerto se bombea a 15 µl/minuto durante 20 s, a continuación, la solución se bombea hacia atrás y adelante (5 s hacia adelante a 15 µl/minuto, a continuación, 5 s hacia atrás a 15 µl/minuto) durante 180 s. Después de 18 ciclos de bombeo, el chip se lava mediante el bombeo en 5xSSC/5mM EDTA a 15 µl/minutos durante 300 s a 51,6 °C. El termociclador se enfría a

35

[0262] Los cebadores son habitualmente oligonucleótidos 5'-fosforotioato que incorporan cualquier secuencia o modificaciones específicas que se requieren para la escisión. Sus secuencias y proveedores varían de acuerdo con el experimento para el que se van a utilizar y en este caso eran complementarios a los extremos 5' del dúplex molde.

40 La secuencia de ADN usada en este proceso era el conjunto de las dos bibliotecas, que tienen extremos complementarios con los cebadores injertados. La mezcla de la biblioteca se desnaturalizó usando un tratamiento con hidróxido de sodio seguido por la dilución a presión como se ha descrito.

[0263] Para algunos de los experimentos descritos en detalle, los agregados amplificados contenían un conector de diol en uno de los cebadores injertados. Los conectores de diol se pueden introducir mediante la inclusión de un conector adecuado en uno de los cebadores utilizados para la amplificación en fase sólida. La síntesis de la diol fosforamidita se describe en el Ejemplo 4 a continuación. Los productos que contienen esos conectores de diol pueden escindirise usando peryodato y propanolamina como se describe y los polinucleótidos monohebra resultantes se hibridaron como se ha descrito.

50

[0264] Los cebadores injertados contienen una secuencia de bases de T en el extremo 5' para actuar como un grupo separador para facilitar la linealización y la hibridación. Las secuencias de los tres cebadores injertados al chip son las siguientes:

Oligo **A**: 5'-PS-TTTTTTTTTTAAATGATACGGCGACCACCGAUCTACAC-3', donde **U** = 2-desoxiuridina (SEC ID N° 30);

5 Oligo **B**: 5'-PS-TTTTTTTTTTCAAGCAGAAGACGGCATACGAG**Goxo**AT-3', donde **Goxo** = 8-oxoguanina (SEC ID N° 31).

Etapa 2) Preparación de los agregados por amplificación isotérmica:

10 **[0265]** La secuencia de ADN utilizada en el proceso de amplificación es la mezcla de las dos bibliotecas preparadas en el Ejemplo 1, que tienen extremos complementarios a los de los cebadores injertados. El dúplex de ADN (1 nM) se desnatura mediante tratamiento con hidróxido de sodio 0,1 M seguido por dilución a presión hasta la 'concentración de trabajo' 0,2-2 pM deseada en 'tampón de hibridación' (5 x SSC / Tween al 0,1%).

15 **[0266]** La amplificación sobre la superficie se llevó a cabo mediante amplificación isotérmica usando una estación de agregados disponibles comercialmente Solexa/Illumina como se describe en el documento PCT/US/2007/014649. La estación de agregados es esencialmente una placa calefactora y un sistema de fluidos para el suministro controlado de los reactivos a una celda de flujo.

20 **[0267]** El molde de una hebra (desnaturalizado como se ha indicado anteriormente) se hibrida con los cebadores injertados inmediatamente antes de la reacción de amplificación, que de este modo comienza con una etapa inicial de extensión del cebador en lugar de desnaturalización del molde. El procedimiento de hibridación se inicia con una etapa de calentamiento en un tampón estricto para asegurar la completa desnaturalización antes de la hibridación. Después de la hibridación, que se produce durante una etapa de enfriamiento lento de 20 minutos, la celda de flujo se lavó durante 5 minutos con un tampón de lavado (0,3 x SSC/Tween al 0,1%).

25 **[0268]** Durante la hibridación del molde y extensión del cebador, se emplean habitualmente una serie de soluciones/tampones, por ejemplo, una solución que comprende las muestras de ADN, un tampón de hibridación (5 x SSC/ Tween al 0,1%), un tampón de lavado (0,3 x SSC/ Tween al 0,1%), una solución de hidróxido de sodio 2M, un tampón de agregado (Tris 200 mM, sulfato de amonio 100 mM, sulfato de magnesio 20 mM, 1% de Triton, 1,3% de DMSO, pH 8,8); un aditivo de amplificación (betaína 5 M), ADN polimerasa y mezcla de dNTP 10 mM.

30 **[0269]** Para preparar las mezclas de hibridación, se enfría previamente un tubo con la tira de muestra de 0,2 ml y tampón de hibridación. Usando tubo(s) Eppendorf de 1,7 ml, se diluyó el molde(s) de ADN hasta 1 nM en tampón EB (Qiagen). Se añadió 1 µl de NaOH 2 M a 19 µl de molde, se agitó brevemente con vórtice y se incubó durante 5 minutos a temperatura ambiente para desnaturalizar el molde de ADN en hebras sencillas. El ADN desnaturalizado se diluyó a la concentración de trabajo (0,2-2 pM) en tampón de hibridación pre-enfriado (por ejemplo, para 1 ml de mezcla de hibridación 1 pM, se diluyó 1 µl de ADN desnaturalizado en 1 ml de tampón de hibridación pre-enfriado). El volumen requerido depende del número de canales utilizados, opcionalmente se utiliza por lo menos 120 µl de mezcla de hibridación. Por lo tanto, 1 ml de mezcla de hibridación es suficiente para 8 canales. Las muestras se agitaron con vórtice brevemente, se centrifugaron y se añadieron alícuotas a las tiras de los tubos de 0,2 ml previamente enfriados (sin burbujas en la parte inferior de los tubos) y se usaron inmediatamente.

40 **[0270]** Para preparar la premezcla de amplificación (de volumen suficiente para la primera extensión y 35 ciclos de amplificación isotérmica), se mezclan 35 ml de H₂O (milliQ), 7 ml de tampón de agregado (Tris 200 mM, sulfato de amonio 100 mM, sulfato de magnesio 20 mM, 1% de Triton, 1,3% de DMSO, pH 8,8) y 28 ml de aditivo amplificación (disolución de betaína 5 M) hasta lograr un volumen final de 70 ml.

50 **[0271]** Para preparar la primera extensión se mezcló mezcla de Taq, 780 µl de premezcla de amplificación, 16 µl de dNTPs 10 mM y 4 µl de ADN polimerasa Taq hasta un volumen final de 800 µl.

[0272] Un proceso de amplificación típico se detalla en la tabla siguiente (Tabla 1), que detalla los volúmenes de flujo por canal, controlados automáticamente por el componente informatizado de la invención.

Tabla 1. Hibridación del molde y primera extensión

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
1	Bombeo de la premezcla de hibridación	20	120	60	120
2	Bomba de la mezcla de hibridación	96	300	15	75
3	Eliminación de burbujas	96	6	100	10
4	Detener flujo y mantener T	96	30	estático	0

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
5	Enfriamiento lento	96-40	1120	estático	0
6	Bombeo de tampón de lavado	40	300	15	75
7	Bombeo de la premezcla de amplificación	40	280	15	70
8	Bombeo de la mezcla de amplificación	40	95	60	95
9	Primera extensión	74	90	estático	0
10	Enfriar a temperatura ambiente	20	0	estático	0

Amplificación isotérmica a 60 °C usando formamida como desnaturizante

[0273] El ADN copiado se puede amplificar isotérmicamente en agregados a 60 °C utilizando formamida como desnaturizante. La amplificación isotérmica (incluyendo tanto el control de la temperatura como el control de reactivo) es supervisado por el componente informatizado. La Tabla 2 muestra ejemplos de controles de secuencia de comandos. Después de la amplificación isotérmica y de una etapa de lavado opcional, el ácido nucleico de los agregados está listo para ser linealizado (ver más abajo).

10

Tabla 2. Amplificación isotérmica

Etapa	Descripción	T (°C)	Tiempo (seg)	Caudal (µl/min)	V bombeado (µl/min)
(1) Esta secuencia 35 veces	Bombeo de formamida	60	56	30	28
	Bombeo de la premezcla de amplificación	60	56	30	28
	Bombeo de la mezcla Bst	60	72	30	36
2	Bombeo del tampón de lavado	60	280	30	140
3	Bombeo del tampón de almacenamiento	20	380	15	95

Tampón de lavado = 0,3 x SSC/Tween al 0,1%

15 **[0274]** Premezcla de amplificación = betaína 2 M, Tris 20 mM, sulfato de amonio 10 mM, sulfato de magnesio 2 mM, Triton 0,1%, DMSO 1,3%, pH 8,8

[0275] Mezcla Bst = betaína 2 M, Tris 20 mM, sulfato de amonio 10 mM, sulfato de magnesio 2 mM, Triton 0,1%, DMSO 1,3%, pH 8,8, más dNTPs 200 µM y 80 unidades/ml de polimerasa Bst (producto NEB ref M0275L)

20 Tampón de almacenamiento = 5X SSC.

Etapa 3) Preparación de agregados para la primera lectura de secuenciación

[0276] La preparación para la lectura uno se realizó en la estación de agregados Illumina. Todos los volúmenes utilizados en el protocolo fueron 95 µl por carril a menos que se indique lo contrario. La linealización de los oligonucleótidos inmovilizados sobre la superficie de tipo A se logró mediante la incubación con la mezcla de enzimas USER (cóctel de uracilo ADN glicosilasa y endonucleasa VIII, NEB # M5505, 10 U/ml, KCl 10 mM, Tris 20 mM, pH 8,8, (NH₄)₂SO₄ 10 mM, MgSO₄ 2 mM, Triton X-100 al 0,1%, 37 °C, 30 minutos). Todos los extremos 3'-OH expuestos de ADN, ya sea del molde extendido o de oligonucleótidos sobre la superficie no extendidos se bloquearon por terminación de la cadena didesoxi usando un cóctel de transferasa terminal (0,25 U/µl) y una polimerasa modificada (SBS polimerasa como se describe a continuación) (0,015 mg/ml, ddNTP 100 µM, Tris 50 mM, NaCl 50 mM, MgSO₄ 6 mM, EDTA 1 mM, Tween 20 al 0,05%). El bloqueo se logró en un protocolo de dos etapas, incubación inicial a 37 °C durante 30 minutos seguido de una rampa de 60 °C e incubación de la celda de flujo durante otros 15 minutos). Los agregados linealizados y bloqueados se lavaron con 0,3X SSC y tampón de almacenamiento antes de la desnaturización con NaOH 0,1 N. Los agregados desnaturizados se neutralizaron con tampón TE (Tris 10 mM, pH 8,0, EDTA 1 mM) y se lavaron con tampón de almacenamiento. 1 El cebador de secuenciación específico de la lectura 1 (5'-ACACTCTTTCCCTACACGACGCTCTTCCGATC-3', 0,5 µM en tampón

de hibridación (SEC ID N° 4)) se hibridó con los agregados mediante incubación a 60 °C durante 15 minutos, seguido de un lavado de 0,3X SSC a 40 °C (velocidad de rampa de 1 °C/seg). La celda de flujo finalmente se enjuagó con tampón de almacenamiento (a 20 °C). Las celdas de flujo procesadas se transfirieron al Analizador del Genoma Illumina para la lectura de secuenciación 1.

5

Etapa 4) Secuenciación del fragmento diana

[0277] La secuenciación de los agregados del protocolo ilustrativo anterior se llevó a cabo utilizando nucleótidos modificados preparados tal como se describe en la solicitud de patente internacional WO 2004/018493 y se marcaron con cuatro fluoróforos espectralmente distintos, como se describe en el Número de solicitud PCT PCT/GB2007/001770. La secuenciación de los agregados se describe con más detalle en la patente WO06064199.

[0278] Se usó una enzima mutante de la polimerasa 9°N (una variante exo que incluye la triple mutación L408Y/Y409A/P410V y C223S) (SBS polimerasa) para las etapas de incorporación de nucleótidos.

15

[0279] Todos los procesos se llevaron a cabo como se describe en el manual de operación Analizador del Genoma de Illumina. La celda de flujo se montó en el analizador, se cebó con los reactivos de secuenciación: posición # 1 = mezcla de incorporación (mezcla 1 µM de dNTP, 0,015 µg/ml de polimerasa SBS, Tris 50 mM, pH 9,0, NaCl 50 mM, MgSO₄ 6mM, EDTA 1 mM, Tween 20 al 0,05%); posición # 2 = de repuesto (sólo agua Milli-Q); posición # 3 = mezcla de exploración (Tris 100 mM, pH 7,0, ascorbato de sodio 50 mM); posición # 4 = lavado de alta sal (5x SSC, Tween 20 al 0,05%); posición # 5 = tampón de incorporación (Tris 50 mM, pH 9,0, NaCl 50 mM, EDTA 1 mM, Tween 20 al 0,05%); posición # 6 = mezcla de escisión (TCEP 100 mM, Tris 100 mM, pH 9,0, NaCl 100 mM, ascorbato de sodio 50 mM, Tween 20 al 0,05%); la posición # 7 = tampón de escisión (Tris 100 mM, pH 9,0, NaCl 100 mM, Tween 20 al 0,05%); posición # 8 = repuesto (lecturas individuales) o conectada a la salida del módulo PE (experimentos de lectura por pares). Las celdas de flujo se secuenciaron usando recetas estándar de secuenciación, ya sea experimentos de 27 o 37 ciclos. Los datos fueron analizados utilizando el análisis de tuberías estándar.

Incorporación: Cebear con tampón de Incorporación. 125 µl/canal; 60 µl/min. Calentar hasta 60 °C.

Tratar con mezcla de Incorporación. 75 µl/canal; 60 µl/min

Esperar un total de 15 min además de bombear mezcla de incorporación nueva, 25 µl/canal; 60 µl/min, cada 4 min.

30

Enfriar hasta 20°C.

Lavar con tampón de Incorporación. 75 µl/canal; 60 µl/min.

Lavar con 5 x SSC/Tween 20 al 0,05%. 75 µl/canal; 60 µl/min. Cebear con tampón de imagen, 100 µl/canal; 60 µl/min.

Escanear a 4 colores a temperatura ambiente.

Escisión: Cebear con tampón de Escisión (Tris 0,1 M pH 7,4, NaCl 0,1 M y Tween 20 al 0,05%). 125 µl/canal; 60

35

µl/min
Calentar hasta 60°C.

Tratar los agregados con mezcla de Escisión (TCEP 100 mM en tampón de Escisión). 75 µl/canal; 60 µl/min.

Enfriar hasta 20°C.

Lavar con tampón Enzimológico.

40

Lavar con 5 x SSC/Tween 20 al 0,05%.

[0280] Repetir el proceso de constitución y escisión tantos ciclos como sea necesario.

[0281] Se detectaron nucleótidos incorporados utilizando el analizador del genoma Illumina, una cámara CCD fluorescente basada en reflexión total interna descrita en "Sistemas y dispositivos para la Secuenciación por análisis de síntesis", USSN 60/788,248, presentada el 31 de marzo de 2006 y la solicitud PCT correspondiente PCT/US07/07991 presentada el 30 de marzo de 2007.

Etapa 5) Preparación de agregados para la segunda lectura

50

Tras la secuenciación con éxito de la lectura 1 en Analizador del Genoma, las celdas de flujo se dejaron montadas y se prepararon para la lectura 2 *in situ*, utilizando el módulo final emparejado Illumina. El control de la temperatura se logró utilizando el Analizador de Peltier del Genoma. Todos los caudales fueron de 60 µl/min y 75 µl por carril a menos que se indique lo contrario. Los agregados se desnaturalizaron con NaOH 0,1 M para eliminar el cebador de

secuenciación extendida de la lectura 1. Los agregados fueron 3'-desfosforilados utilizando polinucleótido quinasa de T4 (Invitrogen # 18004-010, 200 U/ml, imidazol 50 mM pH 6,4, MgCl₂ 12 mM, 70 µM ADP, 2-mercaptoetanol 1 mM, 37 °C, 30 minutos), antes de la re-síntesis de la hebra A lo que se consiguió utilizando 15 ciclos de amplificación isotérmica a 60 °C (los mismos reactivos y condiciones que se describen para la creación del agregado, salvo que se llevó a cabo a 30 µl/min). Los agregados se llevaron antes y después de la resíntesis con 0,3X SSC (150 µl y 245

µl, respectivamente). La linealización de la hebra B de los agregados resintetizados se logró mediante la escisión de la 8-oxoguanina del oligo tip B utilizando Fpg (formamidopirimidina ADN glicosilasa, NEB # M0240, 80 U/ml, Bis Tris propano 10 mM pH 7,0, MgCl₂ 10 mM, ditiotreitól 1 mM, 37 °C, 30 minutos). El bloqueo se realizó como se describe para lectura 1 utilizando el mismo cóctel de bloqueo. Los agrupados linealizados y bloqueados se desnaturalizaron antes de la hibridación del cebador de secuenciación específica 2 (5'-CGGTCTCGGCATTCCTGCTGAACCGCTCTCCGATC-3', 0,5 µM en tampón de hibridación (SEC ID N° 29)) como

65

se describe para la lectura 1. La lectura 2 de las celdas de flujo procesadas se secuenciaron posteriormente en el Analizador del Genoma Illumina.

[0283] Aunque la invención anterior se ha descrito con cierto detalle para fines de claridad y comprensión, será evidente para un experto en la materia a partir de una lectura de esta divulgación que se pueden hacer varios cambios en forma y detalle sin apartarse del verdadero alcance de la invención. Por ejemplo, todas las técnicas y aparatos descritos anteriormente pueden utilizarse en varias combinaciones.

LISTADO DE SECUENCIAS

10

[0284]

<110> Solexa Limited

15

<120> Procedimiento para la secuenciación de un molde polinucleotídico

<130> P88738WO00

20

<160> 31

<170> PatentIn versión 3.3

25

<210> 1

<211> 30

<212> ADN

<213> Secuencia artificial

30

<220>

<223> Cebador oligonucleotídico

<220>

<221> misc_feature

<222> (10)..(11)

<223> El enlace entre el nucleótido 10 y el nucleótido 11 es diol

35

<400> 1

ttttttttt aatgatacgg cgaccaccga 30

40

<210> 2

<211> 34

<212> ADN

<213> Secuencia artificial

45

<220>

<223> Cebador oligonucleotídico

<220>

<221> misc_feature

<222> (34)..(34)

<223> "n" en la posición 34 representa un desoxinucleótido que contiene la base uracilo

50

<400> 2

ttttttttt caagcagaag acggcatacg agan 34

55

<210> 3

<211> 34

<212> ADN

<213> Secuencia artificial

60

<220>

<223> Cebador oligonucleotídico

<400> 3

tcggtctcgg cattcctgct gaaccgagtc gatg 34

65

45

<210> 4
 <211> 32
 <212> ADN
 <213> Secuencia artificial
 5
 <220>
 <223> Cebador oligonucleotídico
 <400> 4
 10 acactctttc cctacacgac gctctccga tc 32
 <210> 5
 <211> 33
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> Cebador oligonucleotídico
 20 <220>
 <221> misc_feature
 <222> (32)..(33)
 <223> La unión entre el nucleótido 32 y el nucleótido 33 es fosforotioato
 <
 25 <400> 5
 acactctttc cctacacgac gctctccga tct 33
 <210> 6
 <211> 32
 30 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Cebador oligonucleotídico
 35 <400> 6
 gatcgaaga gcggttcagc aggaatgccg ag 32
 <210> 7
 <211> 58
 <212> ADN
 40 <213> Secuencia artificial
 <220>
 <223> Cebador oligonucleotídico
 45 <220>
 <221> misc_feature
 <222> (57)..(58)
 50 <223> La unión entre los nucleótidos 57 y 58 es fosforotioato
 <400> 7
 aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatct 58
 55 <210> 8
 <211> 61
 <212> ADN
 <213> Secuencia artificial
 60 <220>
 <223> Cebador oligonucleotídico
 <220>
 <221> misc_feature
 65 <222> (60)..(61)

ES 2 424 155 T3

<223> La unión entre los nucleótidos 60 y 61 es fosforotioato

<400> 8
caagcagaag acggcatacg agatcggctc cggcattcct gctgaaccgc tcttccgac 60
t 61

5

<210> 9
 <211> 204
 <212> ADN
 <213> Secuencia artificial

10

<220>
 <223> Clon procedente de la biblioteca BAC

15

<400> 9
aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctac 60
tgatttcatt gcagcacaag gcaaaacttg gctgcatcgt ttacagtcg ccataagcct 120
ttcctctggt aaaccgcctt ctgagatcgg aagagcgggt cagcaggaat gccgagaccg 180
atctcgtatg ccgtcttctg cttg 204

20

<210> 10
 <211> 214
 <212> ADN
 <213> Secuencia artificial

25

<220>
 <223> Clon procedente de la biblioteca BAC
 <400> 10
aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctat 60
caatattgtg aaaatgacca tactgcaaaa aaaaaactac aaattcaatg caattttcat 120
caaaatacca tcatcattct tcacaatatt gatagatcgg aagagcgggt cagcaggaat 180
gccgagaccg atctcgtatg ccgtcttctg cttg 214

30

<210> 11
 <211> 215
 <212> ADN
 <213> Secuencia artificial

35

<220>
 <223> Clon procedente de la biblioteca BAC
 <400> 11
aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctct 60
cactcctggc agagggacgt gtgactagcc atgggcccct aggtctccag ttctgggta 120
gcttgtattt ttgaacatct cctgtatatt agttagatcg gaagagcgggt tcagcagga 180
tgccgagacc gatctcgtat gccgtcttct gcttg 215

40

<210> 12

ES 2 424 155 T3

<211> 147
 <212> ADN
 <213> Secuencia artificial

5

<220>
 <223> Clon procedente de la biblioteca BAC

<400> 12

aatgatacgg cgaccaccga gatctacact cttccctac acgacgctct tccgatctag 60
 tgtagttgag atctgcctta gcagcaagat cggaagagcg gttcagcagg aatgccgaga 120
 cccgatctcgt atgccgtctt ctgcttg 147

10

<210> 13
 <211> 183
 <212> ADN
 <213> Secuencia artificial

15

<220>
 <223> Clon procedente de la biblioteca BAC

20

<400> 13

aatgatacgg cgaccaccga gatctacact cttccctac acgacgctct tccgatctaa 60
 cacatttcaa agtttggggc cctcctcctc cccaaaaaac aaaccacaaa aaacaaacaa 120
 aaagatcggg agagcgggtc agcaggaatg ccgagaccga tctcgtatgc cgtctttctgc 180
 ttg 183

25

<210> 14
 <211> 170
 <212> ADN
 <213> Secuencia artificial

30

<220>
 <223> Clon procedente de la biblioteca BAC

<400> 14

ggcgaccacc gagatctaca ctctttccct acacgacgct cttccgatct gaatgccttt 60
 tatagcattt aatttttccct aagtataatt accaaataaa aattgtataa gatcgggaaga 120
 gcggttcagc aggaatgccg agaccgatct cgtatgccgt cttctgcttg 170

35

<210> 15
 <211> 180
 <212> ADN
 <213> Secuencia artificial

40

<220>
 <223> Clon procedente de la biblioteca BAC

<400> 15

ES 2 424 155 T3

	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttg	60
	ggcccgggag gagtttgccg gggaggagtg ggtttggaat cggggttaaa ggaaagagaa	120
	gatcgggaaga gcggttcagc aggaatgccg agaccgatct cgtatgccgt cttctgcttg	180
5	<210> 16 <211> 190 <212> ADN <213> Secuencia artificial	
10	<220> <223> Clon procedente de la biblioteca BAC <400> 16	
	tgatacggcg accaccgaga tctacactct ttcctacac gacgctcttc cgatctaaga	60
	tctatttcaa atggactgta gatctaagta taaaaggtaa gagaataatt attctagaaa	120
	gtaaagttaa gatcgggaaga gcggttcagc aggaatgccg agaccgatct cgtatgccgt	180
	cttctgcttg	190
15	<210> 17 <211> 192 <212> ADN <213> Secuencia artificial	
20	<220> <223> Clon procedente de la biblioteca BAC <400> 17	
	aatgatacgg cgaccaccag atctacactc tttccctaca cgacgctctt ccgatctggg	60
	aggccaaggt gggtgatca cctgagatca ggagttcgag accagctggc caacatgatg	120
	aaactctgtc tagatcggaa gagcggttca gcaggaatgc cgagaccgat ctcgtatgcc	180
25	gtcttctgct tg	192
30	<210> 18 <211> 185 <212> ADN <213> Secuencia artificial	
35	<220> <223> Clon procedente de la biblioteca BAC <400> 18	
	aatgatacgg cgaccaccga gatctacact ctttcctac acgacgctct tccgatcttg	60
	accattgtaa ccattaatgt agactgcaat gatatgcact atttacaacc ttttttaaga	120
	ctctagatcg gaagagcggg tcagcaggaa tgccgagacc gatctcgtat gccgtcttct	180
	gcttg	185
40	<210> 19 <211> 199	

ES 2 424 155 T3

<212> ADN
<213> Secuencia artificial

5 <220>
<223> Clon procedente de la biblioteca BAC

<400> 19

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctct      60
ttgaagagct ggcagtagaa gataaacagg ctggggaaga agagaaagtg ctcaaggaga      120
aggagcagca gcagcagcag atcggaagag cggttcagca ggaatgccga gaccgatctc      180
gtatgccgtc ttctgcttg                                                    199

```

10 <210> 20
<211> 212
<212> ADN
<213> Secuencia artificial

15 <220>
<223> Clon procedente de la biblioteca BAC

<400> 20

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctag      60
tattcaacaa gtctgtcttt tccaagtgtc tttaaagacc agaaatacct gtttttaaca      120
cacagggttg caaaattcag aggagattgg cagatcggaa gagcggttca gcaggaatgc      180
cgagaccgat ctggtatgcc gtcttctgct tg                                    212

```

25 <210> 21
<211> 247
<212> ADN
<213> Secuencia artificial

30 <220>
<223> Clon procedente de la biblioteca BAC

<400> 21

```

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctct      60
tgagatgagt gatgacggcg cgctggaagt tgetctctgc gctcgcggtc cgccgcgcat      120
tgccaaccgt ctgctgcgtc gagtgcgtga tttcgcgaa gtgaagcacg atggcaccat      180
ctcaagagat cggagagcgc gttcagcagg aatgccgaga ccgatctcgt atgccgtctt      240

```

ctgcttg 247

35 <210> 22
<211> 202
<212> ADN
<213> Secuencia artificial

40 <220>
<223> Clon procedente de la biblioteca BAC

ES 2 424 155 T3

<400> 22

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctgg 60
 ggttggtgga acccagatgc ctcccaggat tgggtggccc tgtggcactt gtacctgctg 120
 ttgctggtgc tgctgctgct gagatcggaa gagcggttca gcaggaatgc cgagaccgat 180
 ctcgatatgcc gtcttctgct tg 202

5 <210> 23
 <211> 166
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Clon procedente de la biblioteca BAC
 <400> 23

aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatctat 60
 gataaggagc aggtttacag atcataagtg caaaagcggg cgagaagatc ggaagagcgg 120
 ttcagcagga atgccgagac cgatctcgtg tgccgtcttc tgcttg 166

15 <210> 24
 <211> 147
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Clon procedente de la biblioteca BAC

20 <400> 24

gatacggcga ccaccgagat ctacactctt tccctacacg acgctcttcc gatctctgat 60
 actgttgtaa ccacccaatt ggttcaagat cgggaagagcg gttcagcagg aatgccgaga 120
 ccgatctcgt atgccgtctt ctgcttg 147

25 <210> 25
 <211> 58
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> Secuencia adaptadora oligonucleotídica

<400> 25
 aatgatacgg cgaccaccga gatctacact ctttccctac acgacgctct tccgatct 58

30 <210> 26
 <211> 61
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Secuencia adaptadora oligonucleotídica

<400> 26
 agatcggaag agcggttcag caggaatgcc gagaccgatc tcgtagccg tcttctgctt 60

g 61

5 <210> 27
 <211> 31
 <212> ADN
 <213> Secuencia artificial

10 <220>
 <223> Cebador oligonucleotídico

15 <220>
 <221> misc_feature
 <222> (10)..(10)
 <223> "n" en la posición 10 representa un desoxirribonucleótido con la base uracilo

<400> 27
 ttttttttn caagcagaag acggcatacg a 31

20 <210> 28
 <211> 30
 <212> ADN
 <213> Secuencia artificial

25 <220>
 <223> Cebador oligonucleotídico

<400> 28
 tttttttttt aatgatacgg cgaccaccga 30

30 <210> 29
 <211> 36
 <212> ADN
 <213> Secuencia artificial

35 <220>
 <223> Cebador oligonucleotídico

40 <400> 29
 cggctctggc attcctgctg aaccgctctt cggatc 36

45 <210> 30
 <211> 37
 <212> ADN
 <213> Secuencia artificial

50 <220>
 <221> misc_feature
 <222> (31)..(31)
 <223> "n" en la posición 31 representa 2-desoxiuridina

55 <400> 30
 tttttttttt aatgatacgg cgaccaccga nctacac 37

60 <210> 31
 <211> 34
 <212> ADN
 <213> Secuencia artificial

65 <220>
 <223> Cebador oligonucleotídico

<220>

<221> misc_feature

<222> (32)..(32)

<223> "n" en la posición 32 representa 8-oxoguanina

5

<400> 31

tttttttt caagcagaag acggcatacg anat

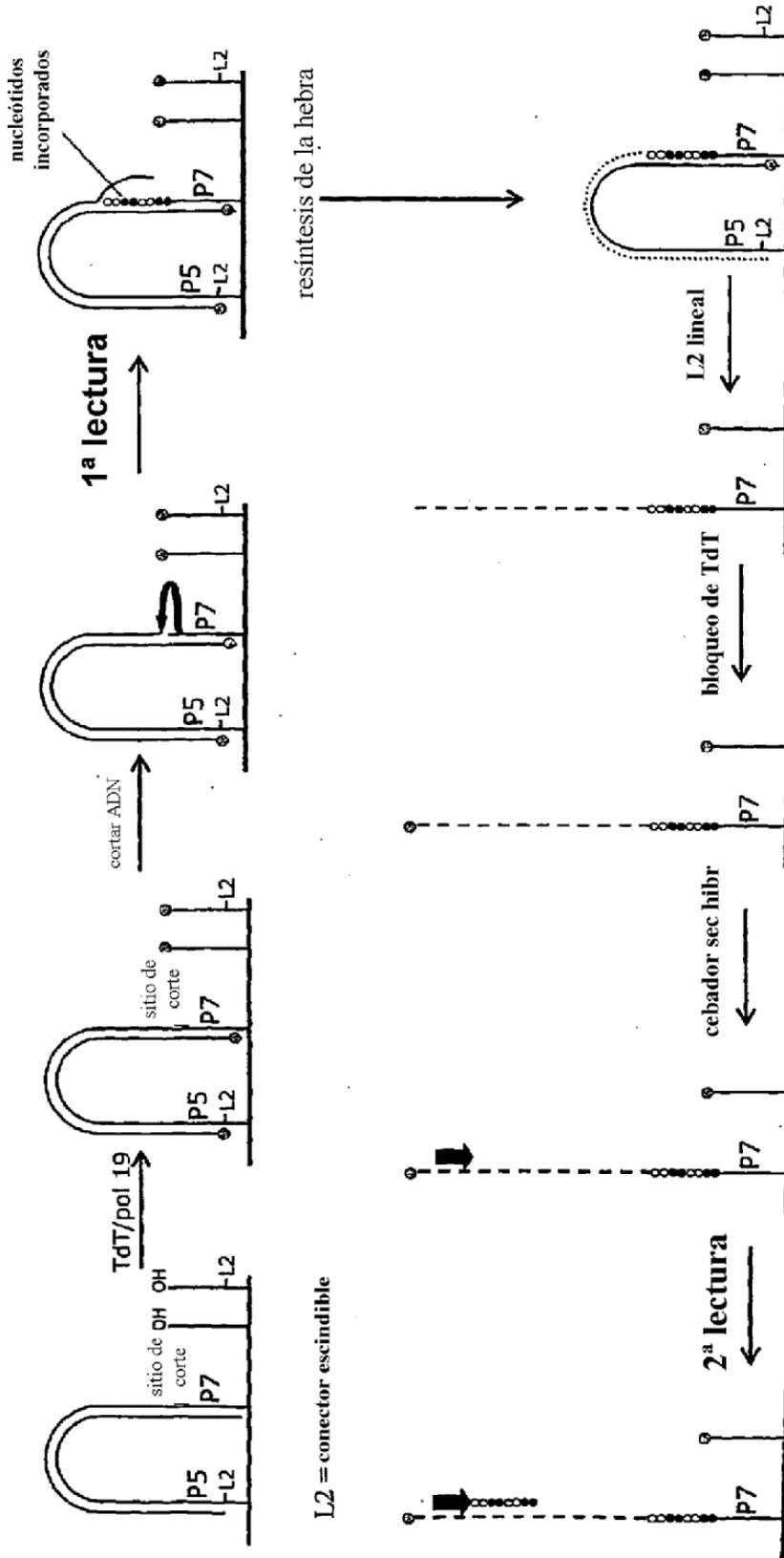
34

REIVINDICACIONES

1. Un procedimiento para la secuenciación por pares de una primera y segunda regiones de un polinucleótido de doble hebra diana, en el que dichas primera y segunda regiones están en el mismo polinucleótido de doble hebra diana, comprendiendo el procedimiento:
- 5
- (a) proporcionar un soporte sólido que tiene inmovilizado sobre él una pluralidad de polinucleótidos molde de doble hebra formados cada uno a partir de una primera y segunda hebras molde complementarias unidas al soporte sólido en sus extremos 5' y múltiples copias de uno o más cebadores inmovilizados en el extremo 5' capaces de hibridar con el extremo 3' de la primera hebra del molde;
- 10
- (b) eliminar selectivamente las segundas hebras del molde de la pluralidad de polinucleótidos molde de doble hebra para permitir la hibridación de las primeras hebras molde con cebadores inmovilizados en el extremo 5';
- (c) llevar a cabo una primera lectura de secuenciación para determinar la secuencia de una primera región del polinucleótido molde mediante una técnica de secuenciación por síntesis o mediante una técnica de secuenciación por ligación;
- 15
- (d) llevar a cabo una reacción de extensión para extender uno o más de los cebadores inmovilizados para copiar la primera hebra del molde para generar una segunda hebra del molde inmovilizada;
- (e) eliminar selectivamente las primeras hebras del molde de la pluralidad de polinucleótidos molde de doble hebra para permitir la hibridación de un cebador de secuenciación con las hebras del molde generadas en la etapa (d);
- 20
- (f) llevar a cabo una segunda lectura de secuenciación para determinar la secuencia de una segunda región del polinucleótido molde, mediante una técnica de secuenciación por síntesis o mediante una técnica de secuenciación por ligación, en el que determinar las secuencias de la primera y segunda regiones del polinucleótido diana logra la secuenciación por pares de dichas primera y segunda regiones de dicho polinucleótido de doble hebra diana.
- 25
2. El procedimiento de acuerdo con la reivindicación 1, en el que las lecturas de secuenciación se realizan utilizando nucleótidos u oligonucleótidos marcados.
- 30
3. El procedimiento de acuerdo con la reivindicación 1, en el que los moldes se inmovilizan en un único soporte sólido plano o en una pluralidad de microesferas.
4. El procedimiento de acuerdo con las reivindicaciones 1 a 3, en el que la eliminación selectiva en las etapas 1 (b) o 1 (e), implica el corte de los polinucleótidos molde de doble hebra inmovilizados con una endonucleasa.
- 35
5. El procedimiento de acuerdo con las reivindicaciones 1 a 3, en el que la eliminación selectiva en las etapas 1 (b) o 1 (e), implica la formación y escisión de un sitio abásico en los polinucleótidos molde de doble hebra inmovilizados.
- 40
6. El procedimiento de acuerdo con la reivindicación 5, en el que dicho sitio abásico se genera a partir de una base uracilo o una base 8-oxo-guanina.
7. El procedimiento de acuerdo con la reivindicación 6, en el que el uracilo se elimina mediante tratamiento con uracilo ADN glicosilasa (UDG) y Endonucleasa ADN glicosilasa-liasa VIII.
- 45
8. El procedimiento de acuerdo con la reivindicación 1, que comprende además la etapa adicional previa a la etapa 1 (c) de convertir el molde inmovilizado en monocatenario.
9. El procedimiento de acuerdo con las reivindicaciones 1 a 3, en el que la primera hebra del molde se une a través de un conector diol que se escinde por tratamiento con un agente de escisión química que comprende perodato y el dúplex se desnaturaliza a continuación.
- 50
10. El procedimiento de acuerdo con la reivindicación 1, en el que la etapa 1 (d) se repite a través de múltiples ciclos de extensión y desnaturalización.
- 55
11. El procedimiento de acuerdo con la reivindicación 1, en el que al menos uno de los cebadores inmovilizados está bloqueado en el extremo 3' y el bloque se elimina antes de la etapa 1 (d).
12. El procedimiento de acuerdo con la reivindicación 11, en el que el bloque es un grupo fosfato y la superficie se trata con una fosfatasa para eliminar el bloque.
- 60
13. El procedimiento de acuerdo con la reivindicación 1, en el que el polinucleótido de doble hebra diana comprende un marcador de secuencia nucleotídica específica para identificar la fuente del polinucleótido.
- 65
14. Uso del procedimiento de cualquiera de las reivindicaciones 1 a 13 para obtener dos lecturas unidas o

emparejadas de información de la secuenciación de cada molde de doble hebra en una matriz agregada.

Metodología de secuenciación de extremos emparejados utilizando una enzima de corte, SBS y resíntesis de la hebra cortada



L2 = conector escindible

Figura 1

Secuenciación de extremos emparejados: procedimiento de síntesis por linealización con USER

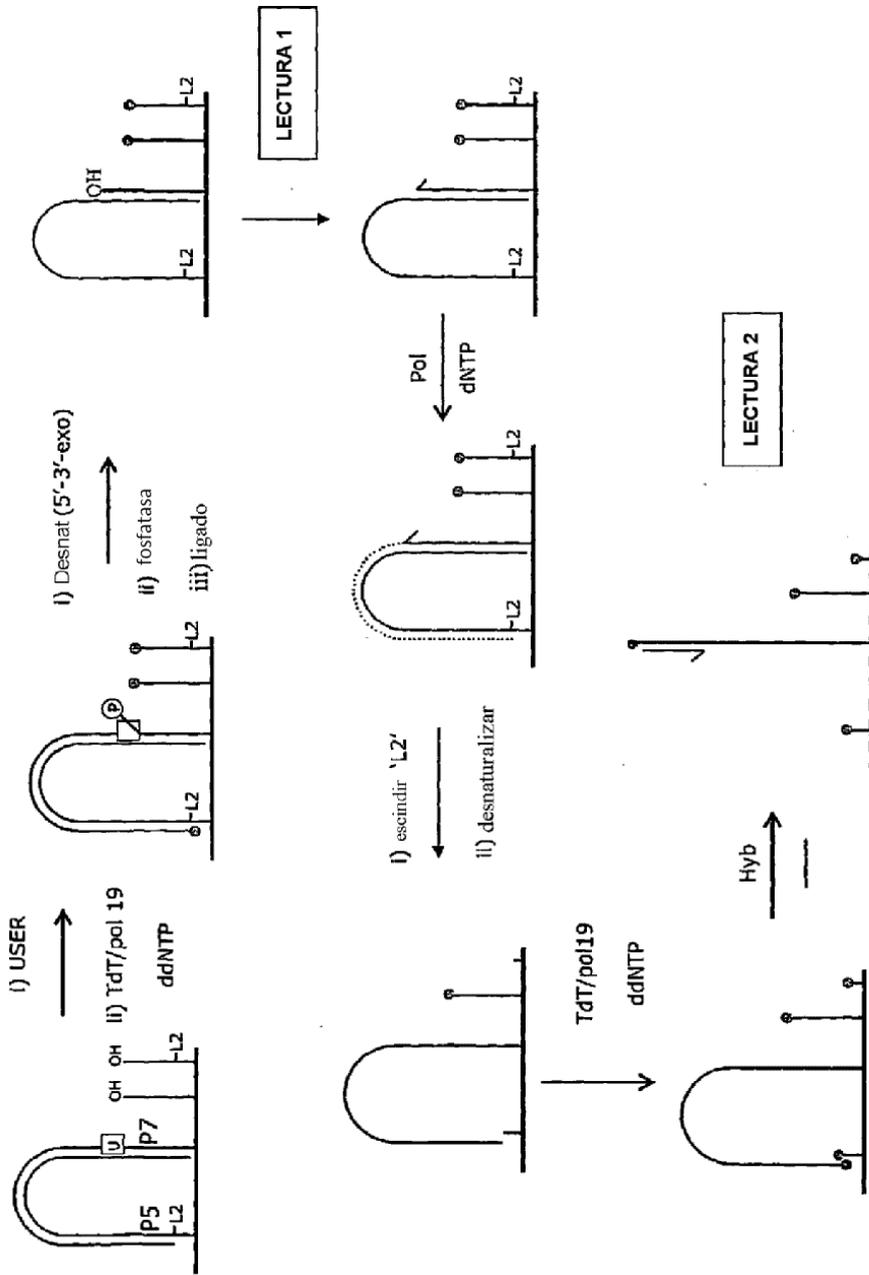


Figura 2

Secuenciación de extremos emparejados: ligación de extremos

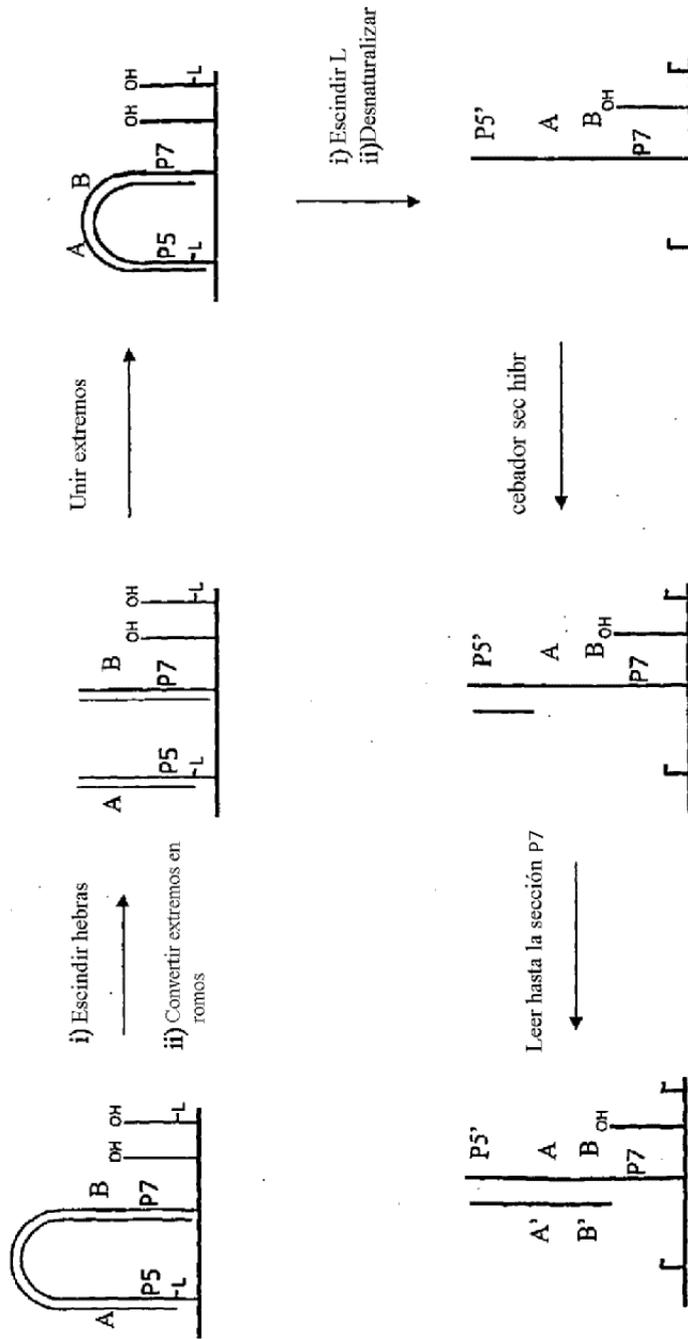


Figura 3

Secuenciación de extremos emparejados: amplificación de agregados entre lecturas

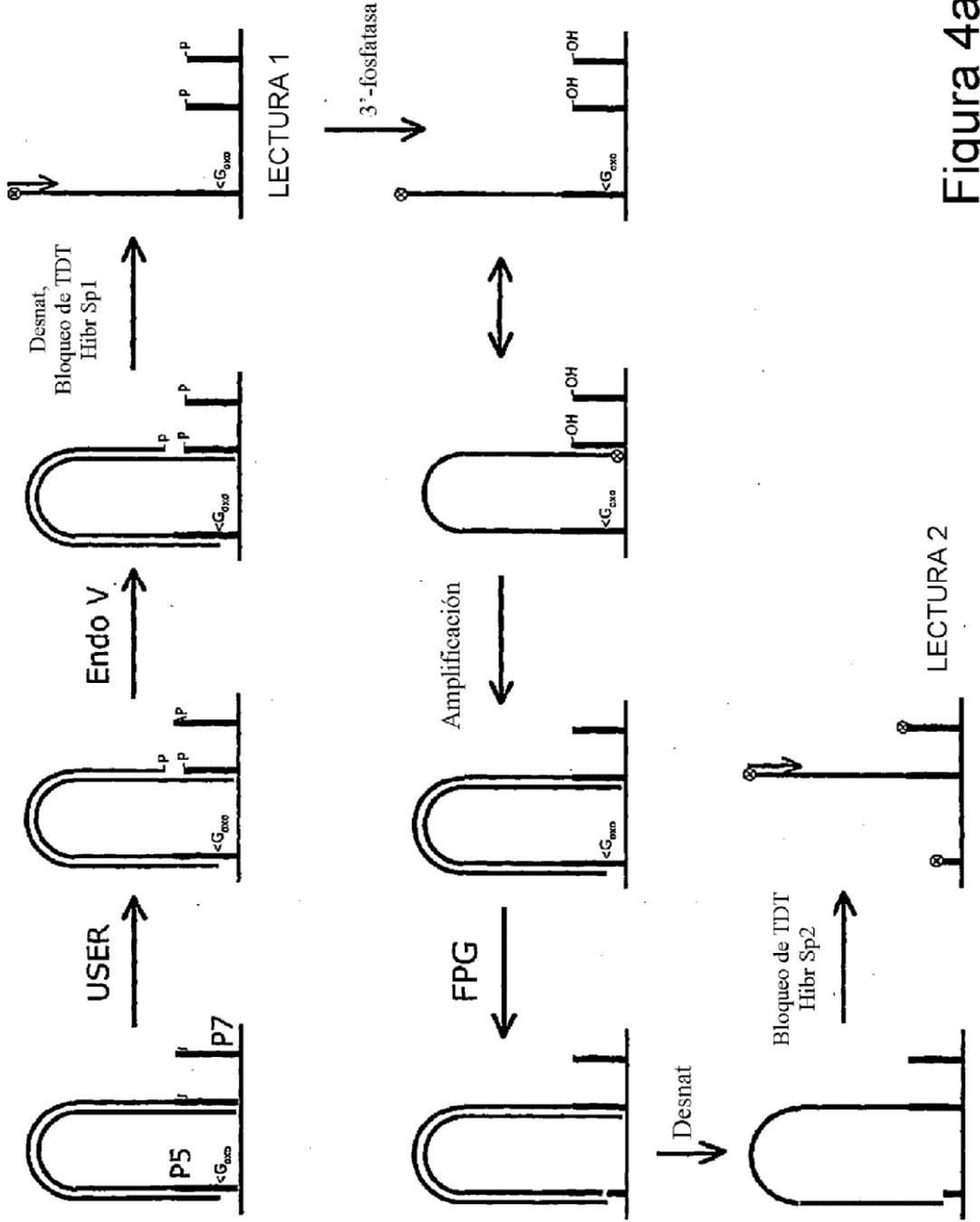


Figura 4a

Secuenciación de extremos emparejados: amplificación de agregados entre las lecturas usando 3 cebadores de injerto

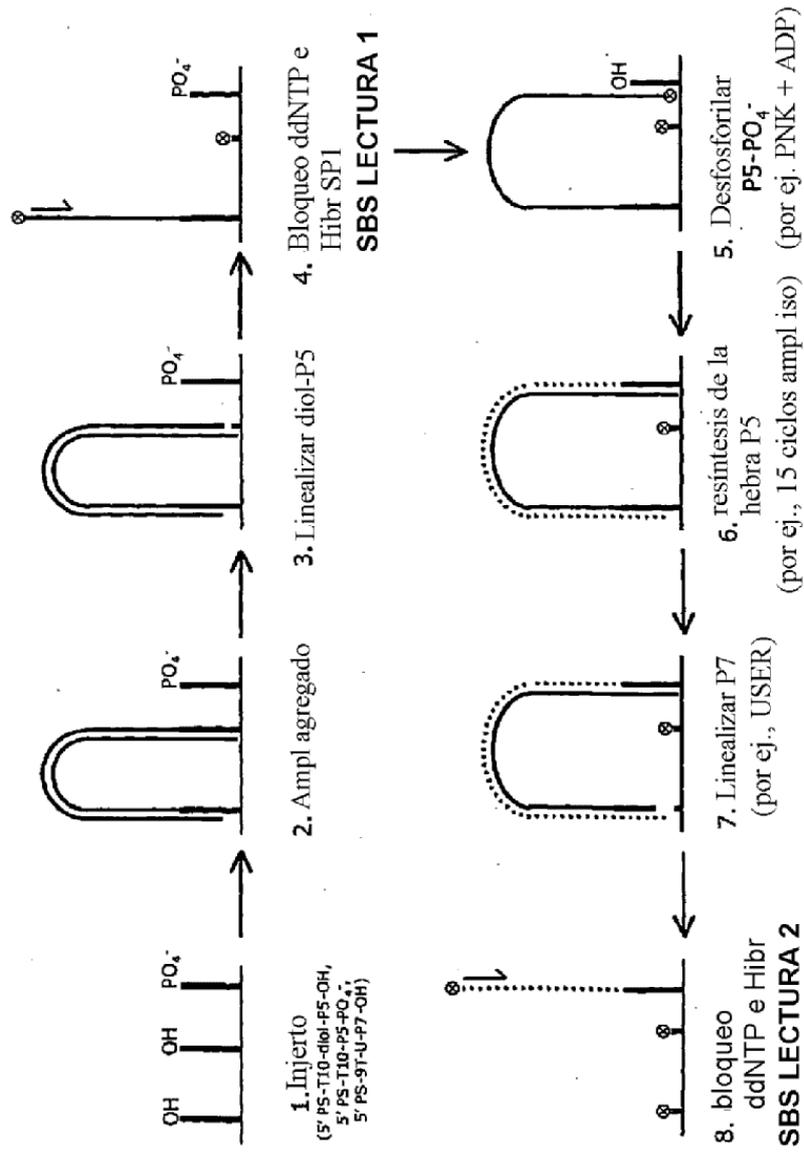


Figura 4b

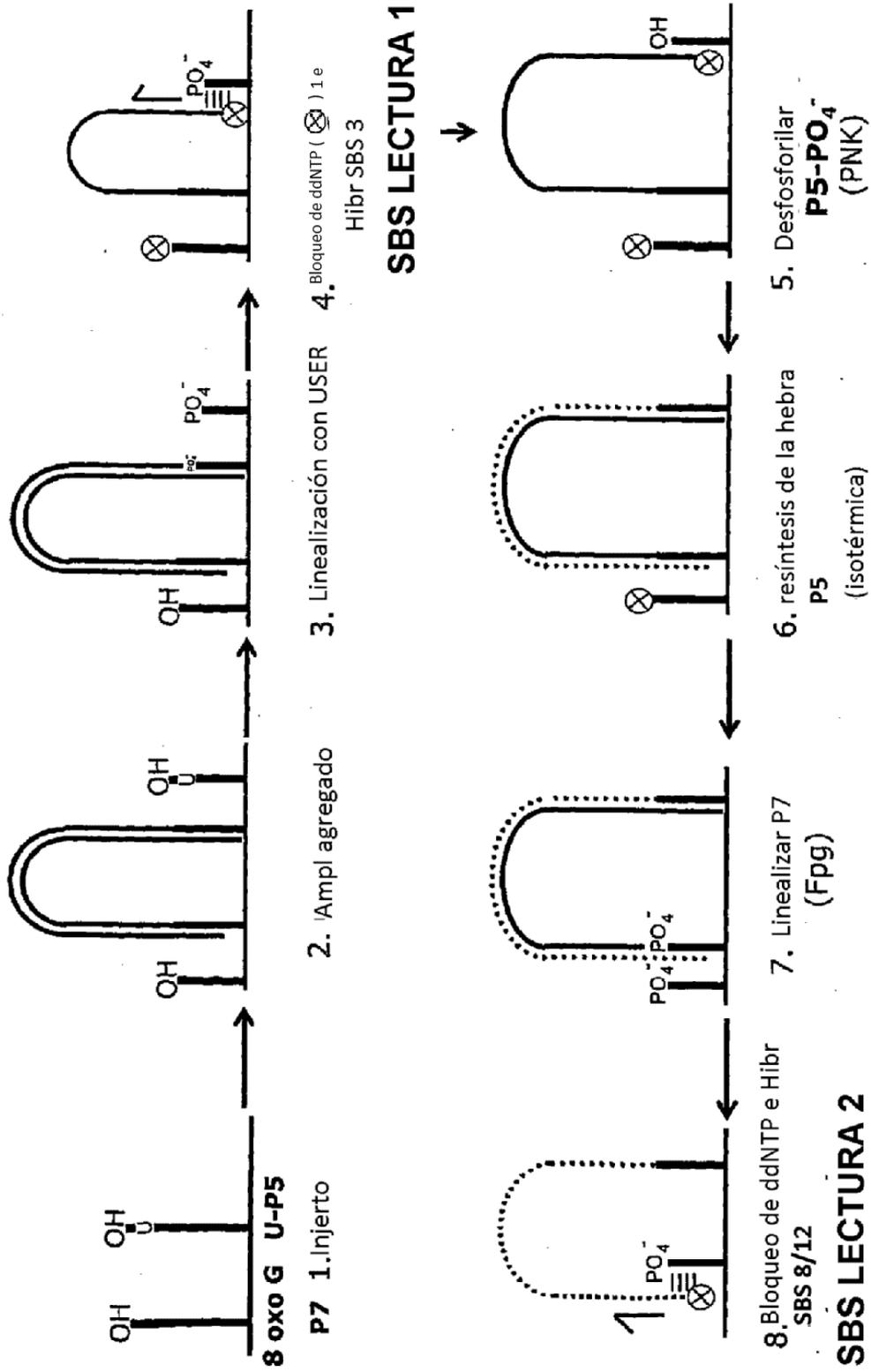
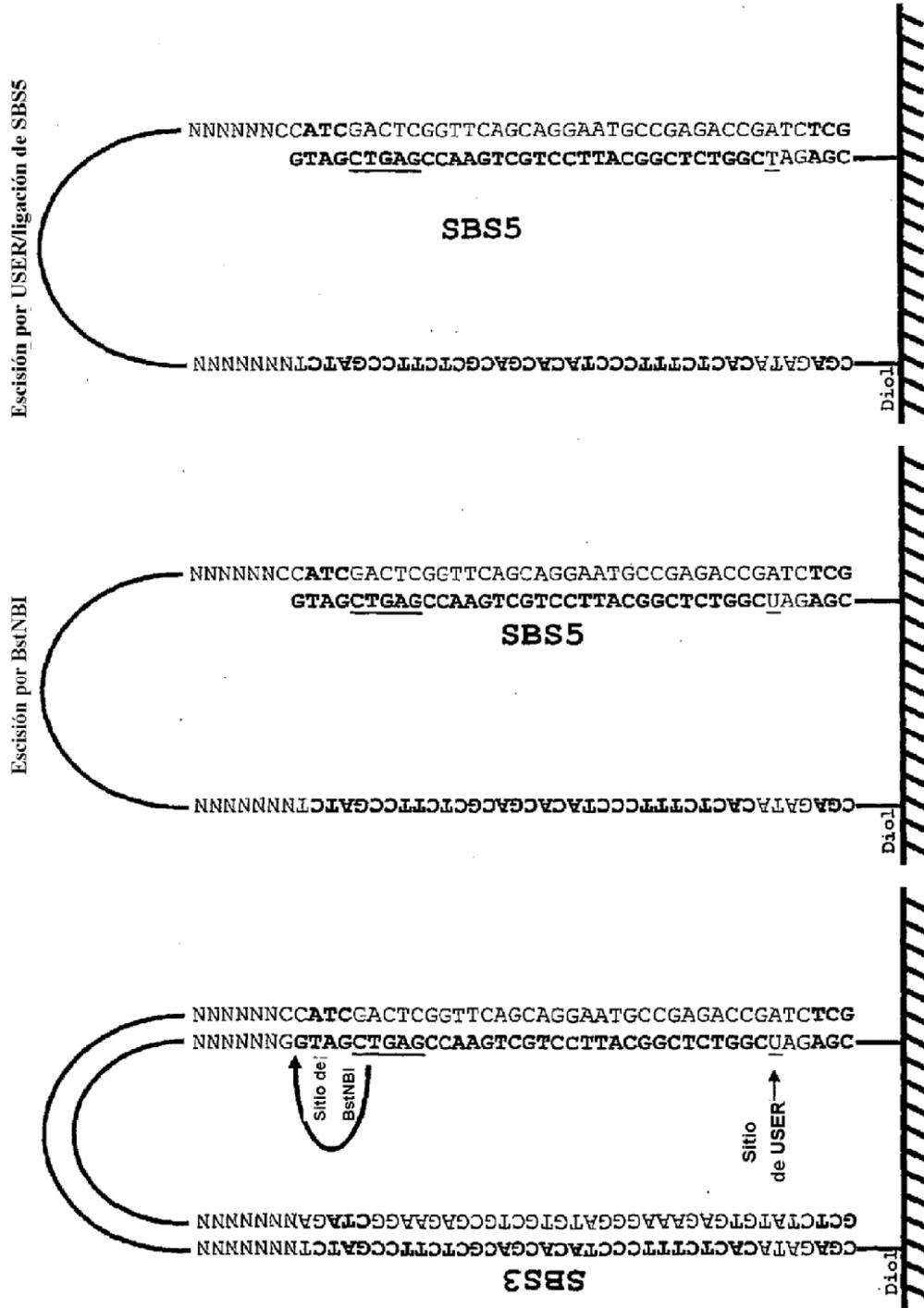


Figura 4c

Figura 5

Estrategias clave para los agregados



Secuenciación de extremos emparejados: constructo de ADN para el crecimiento del agregado

TTTTTTTTTAAATGATACGGCGACCACCGAGATACACTCTTTCCGTTACAGGACGGCTCTCCGATCTCAG
TGGAATGCAITGGCTGCGAGCTGGGGCCGACAGCGCTTCTCCGGGTATGAACAGTTCCGCCCTACGAC
GGCAAGGATATCTCACCCCTGAATGATCCATCGACTCGGTTACAGCAGGAATGCCGAGACCCGATCTCGTA
TGCCGCTTCTGCTTGAAAAAAAA

TTTTTTTTTAAATGATACGGCGACCACCGAG = 10T-P5

ACACTCTTTCCCTACAGGACCGCTCTTCCGATC = SBS3

GACTC = complemento inverso de la secuencia de reconocimiento para N.BstNB I

ATCTCGTATGCCGGTCTTCTTGCTTGAAAAAAAA = complemento inverso de 10T-P7-GAU

CATCGACTCGGTTACAGCAGGAATGCCGAGACCCGA = complemento inverso de SBS5
 cebador de ligación

Figura 6

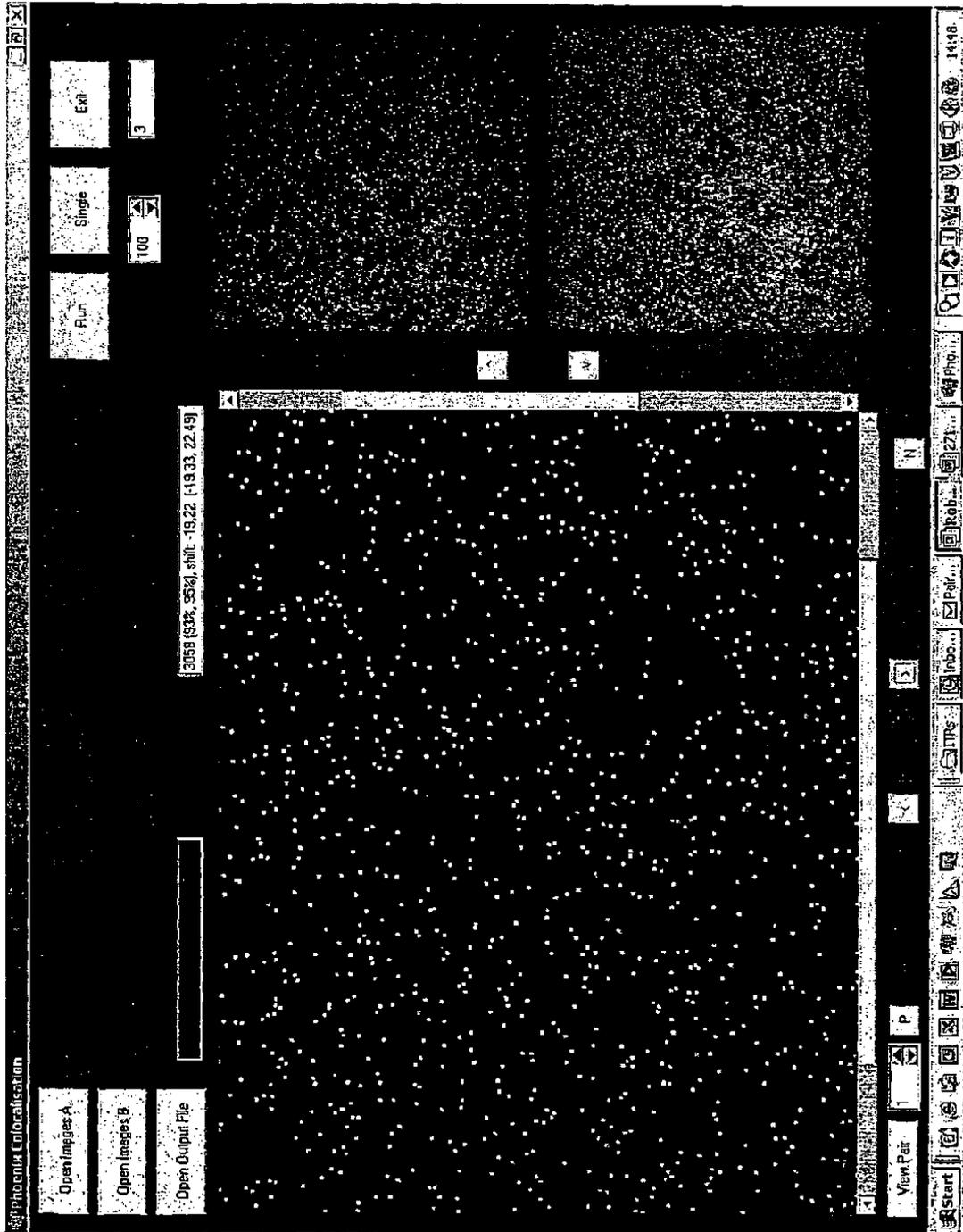


Figura 7: Dos ciclos de SBS utilizando cebador de secuenciación inmovilizado

Comparación entre dos lecturas emparejadas de una biblioteca de fragmentos BAC

BAC CT413 FC 3433 Carril 3	Lectura 1 (25 ciclos)	Lectura 2 (25 ciclos)
Promedio # / bloque	8572	8208
% PF Agregados	77,5	77,8
Prom. % alin. (PF)	95	95
Prom. tasa de error (Phagealign)	3,24	4,41
Prom. tasa de error (ELAND)	0,31	1,12
"Phasing"	0,69	0,85
"Prephasing"	0,21	0,34

Contaminación por *E. Coli* ↓

Figura 8

Lecturas emparejadas con dos enzimas de corte y resíntesis de la hebra

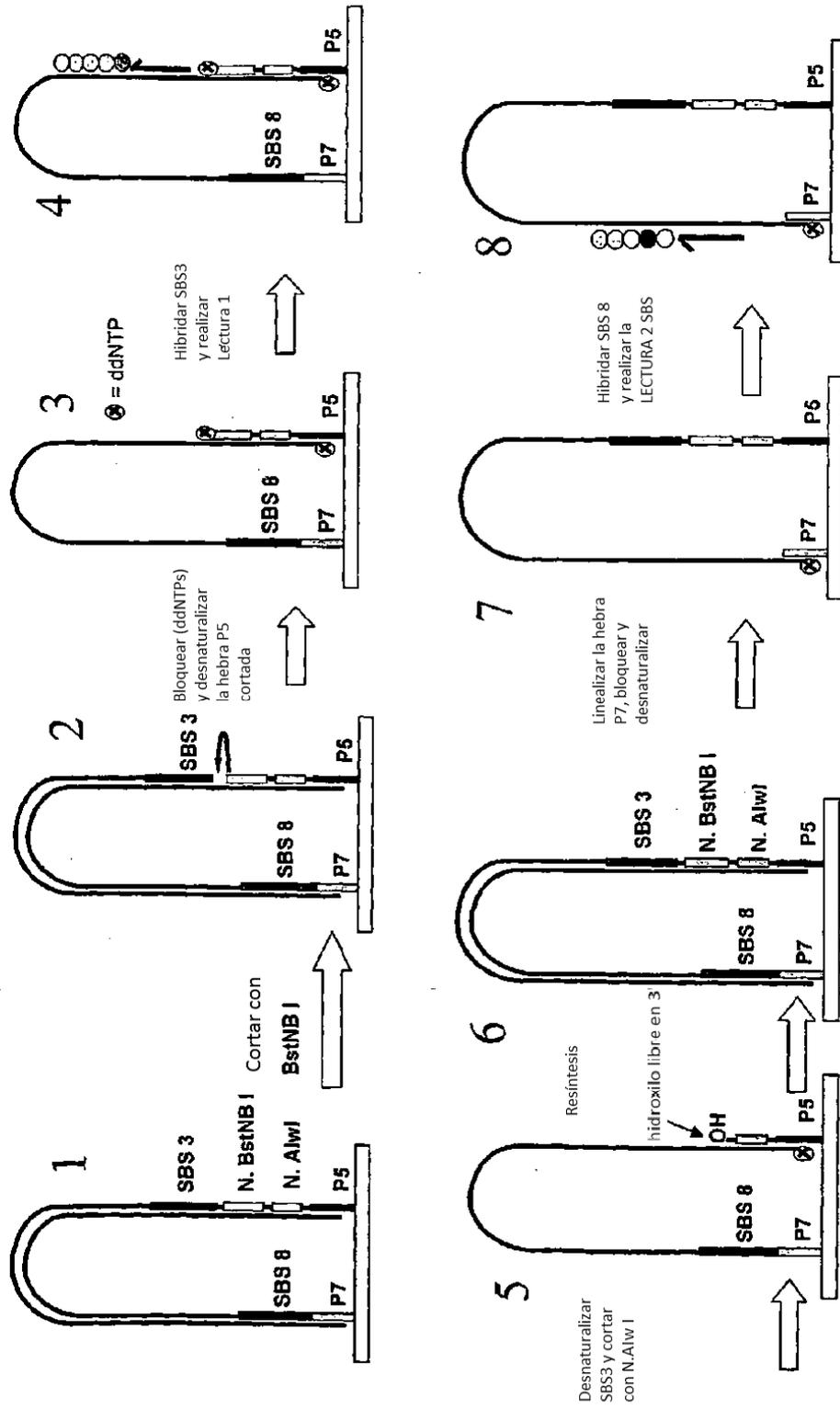
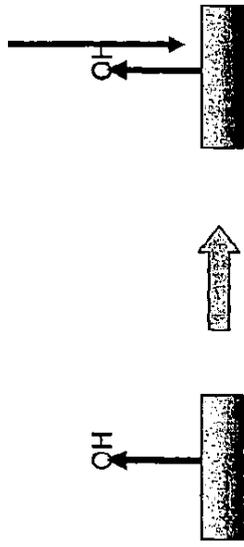


Fig 9

5

2 Hibridación del oligonucleótido complementario inverso

1 Oligo inmovilizado en la superficie



3 Extensión de cebador inmovilizado con dNTPs y ADN polimerasa

4 Deshibridar oligonucleótido molde

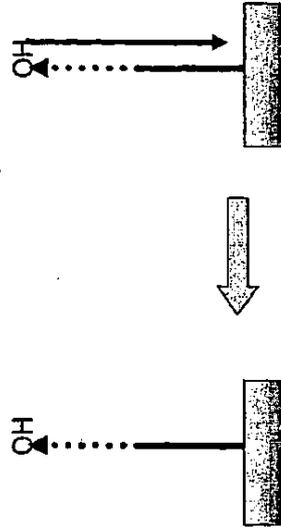


Fig 10

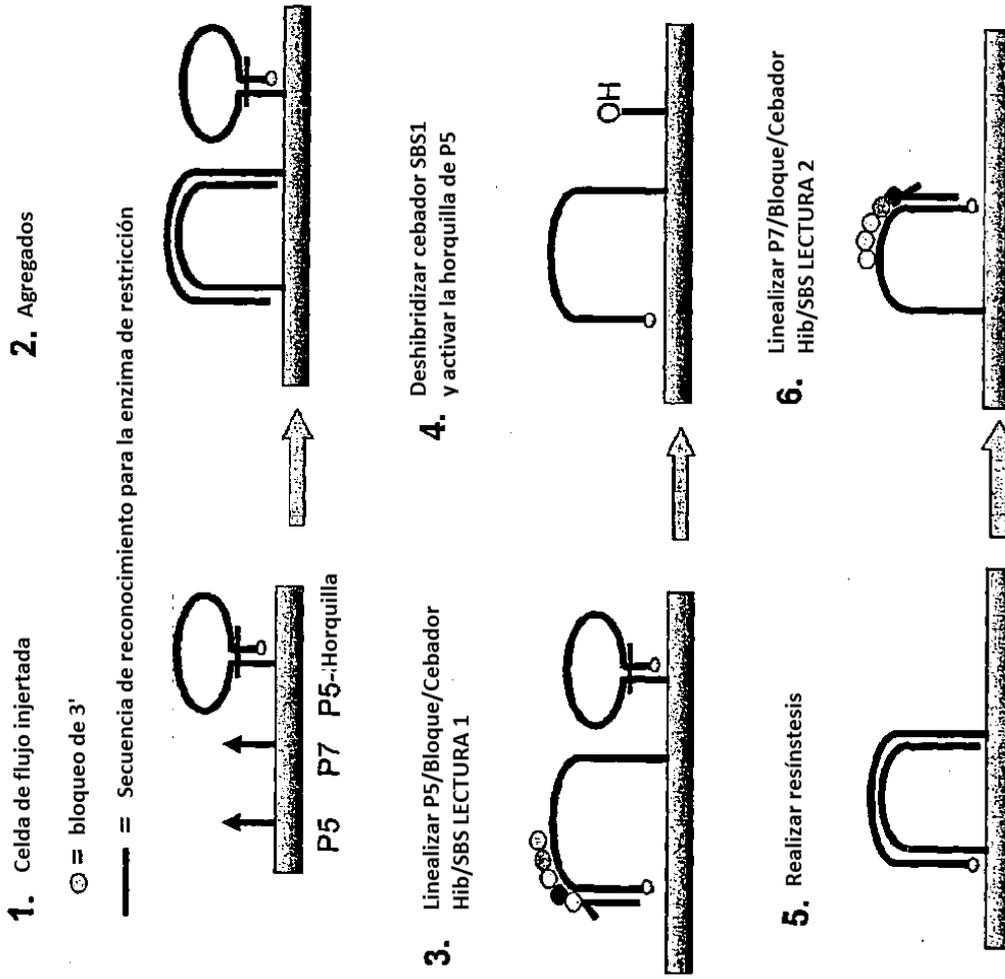


Fig 11

Secuenciación de extremos emparejados y el marcador index

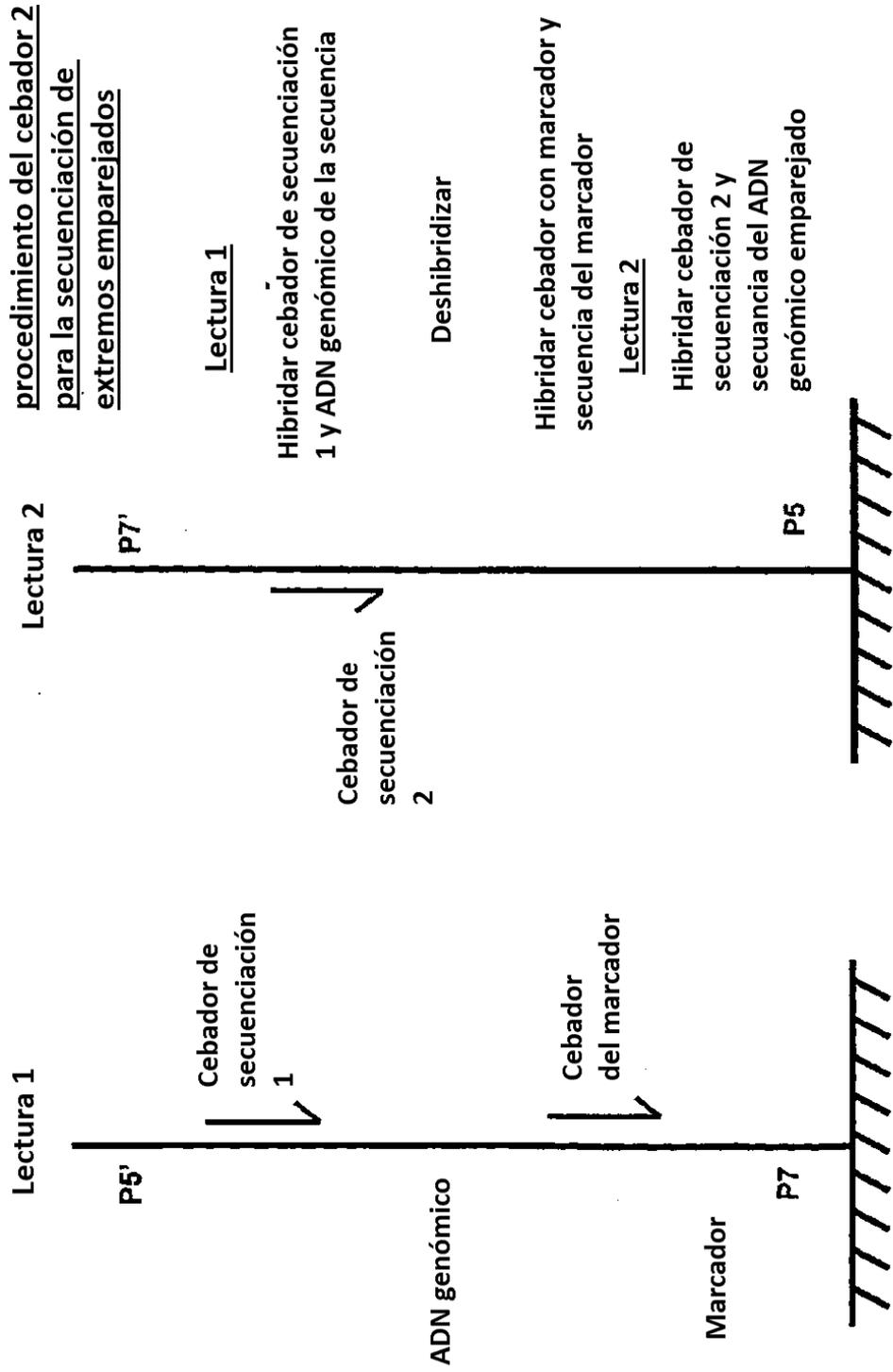


Fig 12

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

- US 20060024681 A [0009]
- US 20060292611 A [0009]
- WO 06110855 A [0009]
- WO 06135342 A [0009]
- WO 03074734 A [0009]
- WO 07010252 A [0009]
- WO 07091077 A [0009]
- WO 00179553 A [0009]
- DE 10051564 A1 [0010]
- WO 9844151 A [0011] [0109] [0131] [0138]
- WO 0018957 A [0011] [0109] [0130] [0138]
- WO 9844152 A [0011]
- WO 07010251 A [0031] [0059]
- US 6306597 B [0046]
- WO 06084132 A [0046]
- GB 2007001770 W [0051] [0104] [0277]
- US 2007007991 W [0054]
- WO 04018497 A [0103]
- US 07007991 W [0105]
- WO 07052006 A [0108]
- WO 0206456 A [0109]
- WO 07107710 A [0109] [0139]
- WO 05068656 A [0110]
- US 60899221 B [0110]
- WO 05065814 A [0127]
- US 11486953 B [0134]
- WO 2004018493 A [0209] [0277]
- US 60801270 B [0209]
- WO 06064199 A [0209] [0277]
- US SN60788248 P [0213] [0281]
- US 2007014649 W [0266]
- US 0707991 W [0281]

Literatura diferente de patentes citada en la descripción

- *Nucleic Acids Research*, 1999, vol. 27 (24), e34 [0022]
- *Science*, 2005, vol. 309 (5741), 1728-1732 [0022]
- *Nature*, 2005, vol. 437, 376-380 [0022]
- CHEN et al. *Biotechniques*, 2002, vol. 32, 518-520 [0073]
- KOMIYAMA et al. *Chem. Commun.*, 1999, 1443-1451 [0073]
- SHIDA. *Nucleic Acids Research*, 1996, vol. 24, 4572-4576 [0084]
- ZHAO et al. *Nucleic Acids Research*, 2001, vol. 29 (4), 955-959 [0127]
- PIRRUNG et al. *Langmuir*, 2000, vol. 16, 2185-2191 [0127]
- *Nature Methods*, 2005, vol. 2, 105-111 [0216]