



19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 353 059**

51 Int. Cl.:  
**C12Q 1/68** (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **04727930 .2**

96 Fecha de presentación : **16.04.2004**

97 Número de publicación de la solicitud: **1613773**

97 Fecha de publicación de la solicitud: **11.01.2006**

54 Título: **Procedimiento para la caracterización de polinucleótidos.**

30 Prioridad: **16.04.2003 GB 0308851**

45 Fecha de publicación de la mención BOPI:  
**25.02.2011**

45 Fecha de la publicación del folleto de la patente:  
**25.02.2011**

73 Titular/es: **LINGVITAE AS.**  
**Trimveien 6, A562**  
**0372 Oslo, NO**

72 Inventor/es: **Lexow, Preben y**  
**Ragnhildstveit, Erlend**

74 Agente: **No consta**

ES 2 353 059 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 caracterización de polinucleótidos.

**5 Campo de la invención**

La presente invención se refiere a procedimientos para convertir un polinucleótido diana en cadenas de información legible, para la determinación posterior de la secuencia del polinucleótido diana.

**10 Antecedentes de la invención**

Los avances en el estudio de las moléculas han sido impulsados, en parte, por la mejora de las tecnologías utilizadas para caracterizar las moléculas o las reacciones biológicas de las mismas. En particular, el estudio de los ácidos nucleicos ADN y ARN se ha beneficiado de las tecnologías en desarrollo utilizadas para el análisis de secuencias y del estudio de los sucesos de hibridación.

El procedimiento principal utilizando en general para la secuenciación de ADN a gran escala es el procedimiento de terminación de cadena. Este procedimiento fue desarrollado por primera vez por Sanger y Coulson (Sanger *et al.*, Proc. Natl. Acad. Sci. USA 74:5463-5467, 1977) y se basa en la utilización de derivados dideoxi de los cuatro nucleótidos que se incorporan en la cadena polinucleótida naciente en una reacción de polimerasa. Tras la incorporación, los derivados dideoxi terminan la reacción de polimerasa y seguidamente los productos se separan mediante electroforesis en gel y se analizan para revelar la posición en la que se ha incorporado en la cadena el derivado dideoxi particular.

Aunque dicho procedimiento ha sido utilizado ampliamente y produce resultados fiables, se reconoce que resulta lento, laborioso y caro.

El documento US-A-5.302.509 da a conocer un procedimiento para secuenciar un polinucleótido inmovilizado sobre un soporte sólido. El procedimiento se basa en la incorporación de bases A, G, C y T 3'-bloqueadas que presentan un marcaje fluorescente diferente al del polinucleótido inmovilizado, en presencia de ADN polimerasa. La polimerasa incorpora una base complementaria al polinucleótido diana, pero el grupo 3'-bloqueante impide adiciones posteriores. El marcaje de la base incorporada seguidamente puede determinarse y eliminarse el grupo bloqueante mediante corte químico, permitiendo que se produzca la polimerización posterior. Sin embargo, la necesidad de eliminar los grupos bloqueantes de esta manera requiere mucho tiempo y debe realizarse con elevada eficiencia.

El documento WO-A-00/39333 describe un procedimiento para secuenciar un polinucleótido mediante la conversión de la secuencia de un polinucleótido diana en un segundo polinucleótido que presenta una secuencia definida e información posicional contenidas en el mismo. La información de secuencia de la diana se dice que se encuentra "magnificada" en el segundo polinucleótido, permitiendo una mayor facilidad para distinguir entre las bases individuales de la molécula diana. Esto se consigue utilizando "etiquetas magnificadoras", que son secuencias de ácidos nucleicos predeterminadas. Cada una de las bases adenina, citosina, guanina y timina en la molécula diana se encuentra representada por una etiqueta magnificadora individual, convirtiendo la secuencia diana original en una secuencia magnificada. A continuación, las técnicas convencionales pueden utilizarse para determinar el orden de las etiquetas magnificadoras, determinando de esta manera la secuencia específica del polinucleótido diana.

En un procedimiento de secuenciación preferente, cada etiqueta magnificadora comprende un marcaje, por ejemplo un marcaje fluorescente, que seguidamente puede identificarse y utilizarse para caracterizar la etiqueta magnificadora.

Aunque el procedimiento dado a conocer en la presente publicación de patente presenta muchas ventajas, todavía existe una necesidad de procedimientos mejoradas para secuenciar los polinucleótidos diana.

**50 Sumario de la invención**

Según un primer aspecto de la presente invención, un procedimiento para identificar características específicas de un polinucleótido diana presente en una muestra comprende las etapas siguientes:

- 55 i) unir a un extremo de cada polinucleótido diana en la muestra una secuencia polinucleótida de señal que es específica de la característica a estudio,
- 60 ii) poner en contacto los polinucleótidos diana con una molécula que interacciona con el polinucleótido diana en el caso de que la característica se encuentre presente,
- iii) separar aquellos polinucleótidos diana que interaccionan respecto de aquellos que no interaccionan,
- iv) opcionalmente repetir las etapas (i) a (iii), y
- 65 v) identificar qué secuencias de señal se encuentran presentes en los polinucleótidos diana separados, y en qué orden, determinando de esta manera las características de cada polinucleótido diana.

## ES 2 353 059 T3

El procedimiento de la presente invención permite determinar múltiples características de un polinucleótido diana de una manera simple que puede automatizarse fácilmente. En la forma de realización preferida, el procedimiento se basa en la incorporación de dos moléculas polinucleótidas en aquellas dianas que presentan la característica específica, permitiendo llevar a cabo una reacción de amplificación, incrementado selectivamente de esta manera el número de polinucleótidos diana con la característica deseada y eliminando efectivamente aquellas dianas que no presentan dicha característica. El procedimiento resulta particularmente adecuado para determinar la secuencia de un polinucleótido diana.

Según un segundo aspecto de la invención, un procedimiento para determinar la secuencia de un polinucleótido diana comprende las etapas siguientes:

- i) tratar una muestra de un polinucleótido bicatenario diana para crear extremos protuberantes 5' y 3', cada uno con un número de bases definido,
- ii) dividir la muestra y poner en contacto cada muestra dividida con una secuencia polinucleótida bicatenaria de señal y con un polinucleótido bicatenario adaptador, representando cada secuencia de señal una secuencia polinucleótida específica presente en el extremo 5'-protuberante y que comprende un extremo protuberante que permite la hibridación y la ligación con el extremo 3' del polinucleótido diana, y comprendiendo cada adaptador un extremo protuberante cuya secuencia es complementaria al de la secuencia representada por la secuencia de señal,
- iii) llevar a cabo la reacción de polimerasa en la muestra o muestras utilizando cebadores que se hibridan con las secuencias de señal y de adaptador, en la que el producto de la reacción de polimerasa comprende un sitio de restricción que permite cortar y separar el adaptador y la parte del polinucleótido diana que formó el extremo 5'-protuberante para formar un nuevo extremo 5'-protuberante, repitiendo opcionalmente las etapas (i) a (iii) utilizando enzimas de restricción para crear los extremos protuberantes, y
- iv) identificar qué secuencias de señal se encuentran presentes en los productos amplificados, y en qué orden, determinando de esta manera la secuencia del polinucleótido diana.

### Descripción de los dibujos

La invención se describe haciendo referencia a las figuras adjuntas, en las que:

la figura 1 es una representación esquemática de una secuencia de señal, que ilustra la conversión de parte de un polinucleótido diana en una secuencia de señal definida,

la figura 2 es una representación esquemática de un experimento de secuenciación de un polinucleótido que se ha llevado a cabo utilizando la presente invención,

la figura 3 es una representación esquemática de la utilización de diferentes compartimientos con secuencias de señal específicas asociadas en un procedimiento de conversión,

la figura 4 muestra los resultados de la conversión, llevando a cabo tres ciclos para formar una serie de secuencias de señal que representan la secuencia de bases de un polinucleótido diana,

la figura 5 es una representación esquemática de un experimento de secuenciación de un polinucleótido, que se ha llevado a cabo para determinar una secuencia de tres bases por ciclo, y

la figura 6 muestra los resultados de la conversión, que demuestra la incorporación de secuencias de señal específicas en un polinucleótido diana.

### Descripción de la invención

El procedimiento de la presente invención es una adaptación del procedimiento de conversión dado a conocer en el documento WO-A-00/39333. En resumen, la presente invención permite identificar características específicas de un polinucleótido diana basándose en la incorporación de secuencias de señal específicas. Las secuencias de señal se incorporan en la totalidad de los polinucleótidos diana presentes en una muestra, aunque sólo las que comprenden la característica específica se separan y caracterizan.

El procedimiento resulta particularmente adecuado para determinar la secuencia de un polinucleótido diana. Mediante la utilización del procedimiento, puede convertirse un número definido de secuencias en un extremo del polinucleótido diana en una secuencia de señal definida en el otro extremo del polinucleótido diana. En una serie de etapas, el extremo "secuenciado" de la diana se corta, mientras que el extremo de secuencia de señal se construye, generando una serie definida de secuencias de señal que puede determinarse mediante una etapa de identificación posterior.

## ES 2 353 059 T3

El término “polinucleótido” es bien conocido de la técnica y se utiliza para referirse a una serie de moléculas unidas de ácidos nucleicos, por ejemplo ADN o ARN. Los miméticos de ácidos nucleicos, por ejemplo APN, LNA (ácidos nucleicos bloqueados) y 2'-O-metil-ARN, también se encuentran comprendidos dentro del alcance de la invención.

5 Las referencias en la presente memoria a las bases A, T(U), G y C, se refieren a las bases nucleótidas adenina, timina (uracilo), guanina y citosina, tal como se apreciará en la técnica. El uracilo sustituye a la timina en el caso de que el polinucleótido sea ARN, o puede introducirse en ADN utilizando dUTP, nuevamente tal como es bien conocido de la técnica.

10 En la presente memoria se hace referencia a los extremos 5' y 3' del polinucleótido diana. En determinadas formas de realización, el polinucleótido diana es bicatenario, y por lo tanto cada extremo comprende extremos tanto 5' como 3'. Sin embargo, se hace referencia a los extremos 5' y 3' refiriéndose a la cadena de la diana que se está estudiando o secuenciando. También resultará evidente que, donde la memoria se refiere al extremo 5' como el extremo que se  
15 secuencia, resultará posible modificar el procedimiento de manera que es el extremo 3' que se secuencia, es decir las secuencias de señal y adaptadores se ligan a extremos opuestos a aquellos indicados en la presente memoria.

En su aspecto más amplio, la invención permite la identificación de características específicas del polinucleótido diana. Una característica puede ser la secuencia del polinucleótido. Alternativamente, una característica puede ser la capacidad del polinucleótido diana de unirse a otras moléculas, por ejemplo a proteínas ligantes de ADN, o de  
20 hibridarse a polinucleótidos complementarios. El procedimiento también puede utilizarse para identificar y caracterizar sitios de enzima de restricción presentes dentro de la secuencia del polinucleótido diana.

Cada característica se identifica llevando a cabo una reacción con una molécula que interacciona con un polinucleótido diana que presenta la característica y que es discriminante respecto de los polinucleótidos que no presentan la  
25 característica. Los polinucleótidos diana que interaccionan con la molécula se marcan con una secuencia de señal que es específica para la característica y que permite determinar la característica en una etapa posterior de lectura.

La presente invención se basa en la incorporación de una secuencia de señal en todos los polinucleótidos diana en una muestra, representando la secuencia de señal una característica específica a estudio. Los polinucleótidos diana  
30 seguidamente se ponen en contacto con una molécula y aquellos que interaccionan se separan. De esta manera, resulta posible identificar la característica la identificar la secuencia de señal.

La separación de aquellas moléculas diana que interaccionan con la molécula puede llevarse a cabo de diversas maneras. Por ejemplo, la molécula puede contener un grupo de unión que permite la separación del complejo diana  
35 de polinucleótido/molécula. Entre los grupos de unión adecuados se incluyen anticuerpos y biotina/estreptavidina. Alternativamente, puede llevarse a cabo una separación por tamaños, debido a que resultará posible distinguir el complejo de diana/molécula de los componentes no acomplexados.

La forma de realización preferida de la presente invención se basa en la incorporación de una secuencia polinucleótida de señal y de una secuencia polinucleótida adaptadora en el polinucleótido diana, lo que permite llevar a cabo una reacción de amplificación para amplificar aquellos polinucleótidos diana que han reaccionado con una molécula  
40 específica. Los productos resultantes de la reacción de polimerasa pueden caracterizarse adicionalmente mediante la repetición de las etapas de reacción de la diana con una molécula y de incorporación de una secuencia de señal adicional que presenta la característica a estudio.

El procedimiento de la invención preferentemente se lleva a cabo utilizando múltiples compartimientos de reacción, uno por cada característica putativa a estudio. Cada compartimiento de reacción presenta una secuencia de señal asociada específica para dicho compartimiento. La muestra que contiene los polinucleótidos diana se divide entre los  
45 comportamientos de reacción y se une una secuencia de señal, por ejemplo mediante ligación, a un extremo de cada polinucleótido diana. La secuencia de señal se une a la totalidad de los polinucleótidos diana en el compartimiento, con independencia de si la diana presenta la característica especificada. La etapa siguiente es realizar un ensayo para la característica mediante la reacción de cada diana con una molécula que interacciona específicamente con aquellas  
50 dianas que comprenden la característica. Por lo tanto, la molécula puede ser una proteína ligante de polinucleótido, enzima u otro polinucleótido. Seguidamente, se separan aquellas dianas que interaccionan con las moléculas. En la forma de realización preferida, la molécula comprende una secuencia adaptadora. Por lo tanto, al interaccionar la diana con la molécula, la molécula (que comprende el adaptador) se une o se liga a la diana. A continuación, se amplifican aquellas dianas que comprenden tanto una secuencia de señal como un adaptador.

La reacción de amplificación puede llevarse a cabo utilizando cualquier procedimiento adecuado. El procedimiento  
60 preferente es utilizar una reacción convencional de polimerasa bajo condiciones conocidas de la técnica. Se apreciará que también pueden llevarse a cabo otras reacciones de amplificación, incluyendo la PCR isotérmica y la amplificación en bacterias.

Una forma de realización preferida es la utilización del procedimiento para secuencia el polinucleótido diana. Esto se lleva a cabo mediante la identificación de un número definido de bases, por ejemplo tres o cuatro bases, en  
65 un extremo del polinucleótido diana, seguido de la conversión de dicha información en una secuencia de señal en el extremo opuesto, permitiendo de esta manera que se produzcan ciclos adicionales de conversión, sin la eliminación de secuencias de señal incorporadas en ciclos anteriores.

## ES 2 353 059 T3

El procedimiento de secuenciación puede llevarse a cabo mediante la obtención del polinucleótido diana y el tratamiento de éste con una serie de enzimas de restricción para formar regiones protuberantes definidas en cada extremo del polinucleótido diana, lo que permite la incorporación de una secuencia de señal y un adaptador durante cada etapa de conversión.

5 Una "secuencia de señal" es un polinucleótido monocatenario o bicatenario que comprende "unidades" diferentes de secuencia de ácidos nucleicos. Cada característica, por ejemplo cada una de las bases A, T(U), G y C en la diana se encuentra representada por una unidad diferente y predefinida, o por una combinación única de unidades. Cada unidad preferentemente comprende dos o más bases nucleótidas, preferentemente entre 2 y 50 bases, más preferentemente entre 2 y 20 bases, y todavía más preferentemente entre 4 y 10 bases, por ejemplo 6 bases. Cada unidad contiene por lo menos dos bases diferentes. En una forma de realización preferida, cada unidad contiene tres bases diferentes. El diseño de las unidades permite distinguir las diferentes unidades durante una etapa de "lectura", por ejemplo que implique la incorporación de nucleótidos marcados detectablemente en una reacción de polimerización, o la hibridación de oligonucleótidos complementarios. Por ejemplo, en el caso de que la característica a estudiar sea la secuencia de la diana, cada base de la diana se encontrará representada por una serie específica de bases dentro de una unidad.

En una forma de realización preferida, cada secuencia de señal comprende dos unidades de secuencia diferente que representan la totalidad de las cuatro bases en la diana. Según esta realización, las dos unidades pueden utilizarse como sistema binario: una unidad representa "0" y la otra representa "1". Cada base de la diana se caracteriza mediante una combinación de las dos unidades. Por ejemplo, la adenina puede representarse como "0" + "0", la citosina como "0" + "1", la guanina como "1" + "0" y la timina como "1" + "1", tal como se muestra en la figura 1. Resulta necesario distinguir entre las unidades, de manera que pueda incorporarse una señal de "parada" en cada unidad. También resulta preferente utilizar unidades diferentes que representen "1" y "0", dependiendo de si la base en el polinucleótido diana (molde) se encuentra en una posición impar o par.

Lo expuesto anteriormente se demuestra de la manera siguiente:

secuencia de molde en posición impar:

"0": ATTTTAT(CC)  
"1": GTTTTGT(CC)

35 Secuencia de molde en posición par:

"0": ACCCCAC(TT)  
"1": GCCCCGC(TT)

Las secuencias de señal adecuadas también se describen en el documento WO-A-00/39333.

45 El "adaptador" es un polinucleótido monocatenario o bicatenario. El propósito del adaptador es proporcionar la secuencia necesaria para que resulte posible la amplificación. En una forma de realización preferida, el adaptador es un polinucleótido bicatenario que comprende un extremo protuberante definido que es capaz de hibridarse y de ligarse a una región complementaria en el polinucleótido diana. La región complementaria habitualmente se encuentra en el extremo de la diana, enfrente de la región a la que se incorpora la secuencia de señal. El adaptador proporciona el sitio de unión de cebador necesario en la reacción de amplificación. En una forma de realización preferida, el adaptador se diseña para hibridarse a aquella parte del polinucleótido diana que se está secuenciando, es decir, la parte 5'-protuberante del polinucleótido diana. En el caso de que el extremo 5'-protuberante de la diana presente 4 bases, el extremo protuberante complementario en el adaptador también presentará 4 bases. Esto permite que se produzca la ligación para formar un polinucleótido de doble cadena. Dado que el extremo 5'-protuberante de la diana presenta una secuencia desconocida, habitualmente resultará necesario utilizar una combinación de adaptadores la totalidad de los cuales presente permutaciones de la secuencia de (por ejemplo) 4 bases. Los adaptadores de diferente secuencia se utilizan en reacciones separadas, tal como se comenta posteriormente.

60 En cada adaptador, habitualmente se incorpora un sitio de reconocimiento de enzima de restricción, permitiendo su corte durante una etapa posterior. Esto puede utilizarse para crear un nuevo extremo 5'-protuberante.

El procedimiento de secuenciación de la forma de realización preferida comprende las etapas generales siguientes:

### *Digestión*

65 En primer lugar, se trata una muestra del polinucleótido diana con uno o más enzimas de restricción, resultando en un extremo protuberante en la diana con un número definido de bases, por ejemplo 3 bases, que se utiliza para la ligación de secuencias de señal posteriormente en el procedimiento (la región protuberante). También se crea una región

## ES 2 353 059 T3

protuberante en el extremo opuesto del polinucleótido diana (el extremo 5'-protuberante). Es esta región protuberante la que debe identificarse (representarse) con la secuencia de señal durante el procedimiento de conversión y que debe ligarse al adaptador.

### 5 *Ligación*

La muestra con el polinucleótido diana se divide uniformemente en compartimientos de reacción que representan todas las permutaciones de la región protuberante que debe secuenciarse. Por ejemplo, en el caso de que la región protuberante que debe secuenciarse presente 4 bases, deben utilizarse 256 compartimientos de reacción. En el primer compartimiento, el polinucleótido diana se pone en contacto con una secuencia de señal que representa, por ejemplo, AAAA (en el caso de que la región protuberante que debe secuenciarse presente 4 bases), la secuencia de señal que se liga al extremo 3'-protuberante en la totalidad de los polinucleótidos diana. También se añade un polinucleótido adaptador con la región protuberante TTTT, aunque ésta únicamente se ligará específicamente a los polinucleótidos diana que contengan una región protuberante AAAA, dejando los polinucleótidos con una secuencia protuberante diferente con una región protuberante no ligada. Se lleva a cabo el mismo procedimiento en los demás compartimientos, representando cada uno una combinación diferente de las posibles secuencias de 4 bases. En cada caso, el adaptador será el complemento de la secuencia representada por la secuencia de señal.

Aunque resulta preferible que la secuencia de señal represente todas las bases de la región protuberante, resulta posible secuenciar menos bases que las presentes en la región protuberante. Por ejemplo, en el caso de que la región protuberante comprenda 4 bases, puede utilizarse una secuencia de señal que represente 3 de ellas. Esto se ilustra en los Ejemplos dados a conocer posteriormente.

### 25 *Amplificación*

Tras la realización de las etapas de ligación, pueden agruparse las muestras y llevarse a cabo una reacción de polimerasa. La reacción de polimerasa se lleva a cabo utilizando cebadores que presentan como diana los extremos de la secuencia de señal y del adaptador, y por lo tanto únicamente se amplificarán exponencialmente aquellas moléculas que se han ligado con éxito tanto a la secuencia de señal como a la secuencia adaptadora, mientras que aquellas moléculas que contienen únicamente una secuencia de señal resultarán eliminadas como consecuencia de una amplificación lineal. El resultado será una población de fragmentos polinucleótidos convertidos, en la que la región protuberante de 4 bases que se generó inicialmente en un extremo del polinucleótido diana ha sido sustituida por una secuencia de señal que representa las 4 bases presentes en el otro extremo del polinucleótido diana.

En una forma de realización preferida, la reacción de polimerasa se lleva a cabo utilizando metil-dCTP, que garantiza que los sitios nativos de enzima de restricción permanezcan inactivos. Las moléculas adaptadoras incorporan sitios de enzima de restricción adicionales, permitiendo de esta manera que se lleven a cabo ciclos adicionales de conversión. Los cebadores utilizados en la reacción de polimerasa son específicos de secuencias presentes dentro de la secuencia de señal y del adaptador, garantizando de esta manera que la etapa de amplificación se produce únicamente en el caso de que tanto la secuencia de señal como el adaptador se encuentren ligados al polinucleótido diana.

También resulta preferible que los sitios de unión de cebador sean diferentes en diferentes conjuntos de secuencia de señal y adaptador, es decir, la secuencia de señal y el adaptador utilizados en un ciclo de conversión comprenderán sitios de unión de cebador diferentes de aquellos utilizados en un ciclo anterior. Esto garantiza que se amplifica la secuencia correcta.

Las regiones protuberantes pueden producirse por medios diferentes, aunque la realización preferente es la utilización de enzimas de restricción, por ejemplo, enzimas de restricción de clase II. Estas enzimas no muestran ninguna especificidad para la secuencia que se corta y por lo tanto pueden generar regiones protuberantes con todo tipo de composición de bases. El sitio de unión de la enzima de restricción puede localizarse de manera que se forme una región protuberante en el interior del polinucleótido diana mismo. En la práctica, resulta preferible seleccionar enzimas que generen regiones protuberantes de 3 ó 4 pares de bases.

55 Las endonucleasas de restricción de clase II son conocidas y se identifican en el documento WO-A-00/39333.

En una forma de realización, en primer lugar el polinucleótido diana se trata con un polinucleótido adaptador que incorpora una secuencia de reconocimiento para la enzima de restricción BbvI, y posteriormente el polinucleótido se trata con la enzima de restricción para generar una región protuberante específica en el extremo 5'. Los sitios de reconocimiento para BbvI se inactivan tratando en primer lugar el polinucleótido diana mediante metilación. Tras la creación de la primera región protuberante, las regiones protuberantes posteriores en el extremo 5' preferentemente se crean utilizando la enzima de restricción *Sfa*NI. Lo anterior se lleva a cabo ligando una secuencia de adaptador específica en el extremo 5' que incorpora la secuencia de reconocimiento de *Sfa*NI.

65 La región protuberante inicial en el extremo 3' preferentemente se crea mediante ligación de un polinucleótido que incorpora una región protuberante definida. La región protuberante definida permite que una primera secuencia de señal se hibride y se ligue al extremo 3'. En otra forma de realización, en primer lugar el polinucleótido diana se trata con un polinucleótido adaptador que incorpora una secuencia de reconocimiento para las enzimas de restricción

## ES 2 353 059 T3

MmeI y *EaeI* (o su esquizómero *Eam1104I*). A continuación, el polinucleótido se trata con la enzima de restricción MmeI para extraer 20 pares de bases de la molécula diana y asociarlos a un polinucleótido adaptador. Un segundo polinucleótido adaptador, que contiene una secuencia de reconocimiento para la enzima de restricción *SfaNI*, seguidamente se liga con el polinucleótido diana asociado al adaptador utilizando la región protuberante generada con MmeI. Se lleva a cabo una amplificación por PCR para amplificar los fragmentos que contienen el polinucleótido diana de 20 pb flanqueado por dos polinucleótidos adaptadores. Para proteger sitios internos de la digestión posterior con *EaeI* y *SfaNI*, se lleva a cabo la PCR con dCTP metilada en lugar de dCTP normal en la mezcla de dNTP. La amplificación de polinucleótidos diana flanqueados por regiones protuberantes generadas con MmeI (es decir productos secundarios de la digestión) pueden eliminarse mediante la utilización de segundos adaptadores no fosforilados. La región protuberante específica en el extremo 5' se genera mediante la digestión del producto de PCR con la enzima de restricción *SfaNI* (o equivalentes). La región protuberante definida en el extremo 3' se crea mediante la digestión del producto de PCR con la enzima de restricción *EaeI* (o *Eam1104I*). La secuencia de señal incorpora un sitio de enzima de restricción, que preferentemente es para la enzima de restricción *EaeI*. Los ciclos posteriores de corte e incorporación se producen utilizando las enzimas de restricción *EaeI* y *SfaNI*, tal como se muestra en la figura 2.

El ciclo de conversión se ilustra adicionalmente en la figura 3, que muestra el destino de cinco fragmentos diferentes en tres compartimientos de reacción seleccionados arbitrariamente de entre los 256 compartimientos utilizados en el procedimiento. En primer lugar, los fragmentos se distribuyen uniformemente en los 256 compartimientos. En el compartimiento 56, en donde la región protuberante "TCTA" debe identificarse, se introduce una secuencia de señal que representa dicha composición de bases y se liga al extremo 3'-protuberante de 3 bases situado en el extremo izquierdo del polinucleótido diana. Esta ligación se produce con independencia de la composición del extremo 5'-protuberante de 4 bases situado en el lado derecho del polinucleótido. A continuación, se introduce y se liga un adaptador específico, que únicamente se liga a fragmentos que presentan regiones 3'-TCTA 5'-protuberantes. La ligación incorrecta del adaptador puede reducirse llevando a cabo la ligación en presencia de adaptadores "bloqueantes". Los adaptadores bloqueantes presentan una secuencia de región protuberante diferente a la del adaptador específico y por lo tanto se ligan a polinucleótidos diana que presentan una región protuberante no complementaria al adaptador específico. A continuación, una reacción de amplificación final amplifica selectivamente los fragmentos que se han ligado al adaptador específico y por lo tanto elimina los demás fragmentos, que se han asociado a una secuencia de señal incorrecta. Se lleva a cabo el mismo procedimiento en los compartimientos 141 y 194, con la excepción de que la composición de las secuencias de señal y de los adaptadores se ajusta de acuerdo con las regiones protuberantes que deben identificarse en los pocillos (3'-ATAG-5' y 3'-CAAT-5').

Los polinucleótidos diana de los 256 compartimientos preferentemente se agrupan en un tubo de reacción común antes de que tenga lugar la reacción de amplificación, evitando de esta manera el inconveniente de llevar a cabo 256 reacciones de amplificación separadas.

Tras la realización de ciclos de conversión posteriores, se reduce la secuencia en el extremo 5' del polinucleótido diana y las secuencias de señal introducidas en el extremo 3' del polinucleótido diana se incrementan de una manera definida, correspondiente al polinucleótido diana ahora "secuenciado". La determinación del tipo y orden de las secuencias de señal puede llevarse a cabo utilizando procedimientos dados a conocer en el documento WO-A-00/39333 o en la solicitud copendiente de patente internacional presentada en nombre de LingVitae y que reivindica prioridad de la patente GB 0308852.3.

La reacción de amplificación puede producir artefactos no deseados y pueden adoptarse diversas medidas para reducir o eliminar los mismos. Por ejemplo, resulta preferible que el sitio de reconocimiento de cebador en diferentes señales de secuencia (y opcionalmente también diferentes adaptadores) sea diferente, con el fin de reducir la probabilidad de que los cebadores se unan a secuencias de señal internas y amplifiquen fragmentos de la diana convertida, de mayor tamaño.

También resulta preferible diseñar las secuencias de señal de manera que secuencias de señal consecutivas o posteriores comprendan secuencias de bases diferentes. Lo anterior garantiza que los cebadores introducidos para hibridarse a la secuencia de señal terminal no se hibriden a secuencias de señal internas, lo que resultaría en la producción de fragmentos durante la etapa de amplificación.

La digestión incompleta por parte de una o más de las enzimas de restricción también puede ser una fuente de artefactos. La utilización de enzimas de restricción que presenten elevadas especificidad y actividad reducirá este problema.

Los artefactos también pueden eliminarse mediante purificación en gel, o mediante inmovilización de los productos previamente a la digestión con la enzima de restricción; aquellos productos no digeridos por completo permanecerán inmovilizados y pueden separarse de aquellos productos que han experimentado una digestión completa.

La eliminación del exceso de adaptadores y dianas previamente a la amplificación también reducirá la producción de artefactos.

Los Ejemplos siguientes ilustran la invención.

## ES 2 353 059 T3

### Ejemplo 1

#### *Preparación inicial utilizando BbvI y sin restricción por tamaño*

5 Antes (o después) de la fragmentación del ADN, el polinucleótido diana se metila con el fin de inactivar los sitios de reconocimiento nativos. Mediante la utilización de BbvI como la enzima de restricción, se utiliza M.Bbv (Mega-  
base Research Products) para inactivar sus sitios diana. El polinucleótido diana se fragmenta aleatoriamente en trozos  
de 500 a 1.000 pares de bases mediante rotura mecánica utilizando un dispositivo HydroShear (GeneMachines). En  
combinación con la sonicación, pueden generarse fragmentos incluso más cortos. Para incrementar la eficiencia de  
10 la etapa posterior de ligación de adaptador de extremos romos, los polinucleótidos fragmentados se tratan opcional-  
mente con ADN polimerasa para rellenar o para eliminar regiones protuberantes utilizando su función exonuclea-  
sa. Con este fin comúnmente se utilizan ADN polimerasa de T4 y fragmento Klenow grande solos o en combina-  
ción.

15 Se ligan en extremo romo dos tipos de polinucleótidos adaptadores a los fragmentos del polinucleótido diana. Uno  
incorpora un sitio BbvI y el otro incorpora una región protuberante para la ligación posterior a una secuencia de señal.  
Cada adaptador presenta una longitud de entre 10 y 20 pares de bases y se añade en un exceso molar de 10X a los  
polinucleótidos diana. Los adaptadores se fosforilan excepto una de las cadenas del adaptador BbvI. Además, la cadena  
opuesta de este adaptador porta un grupo 3'-desoxi para evitar la concatamerización durante la ligación. La ligación  
20 se lleva a cabo bajo condiciones optimizadas para la ligación de extremos romos siguiendo las recomendaciones del  
fabricante.

Los fragmentos de polinucleótido ligados se digieren con BbvI. Tras la digestión, se lleva a cabo una selección  
según el tamaño (por ejemplo separación en gel o fraccionamiento en columna) para eliminar el exceso de adaptadores,  
25 fragmentos terminales cortados y fragmentos de polinucleótido inferior o superior a determinado valor umbral (por  
ejemplo, una distribución duplicada de tamaños).

### Ejemplo 2

#### *Preparación inicial utilizando restricción con MmeI y de tamaños*

El polinucleótido diana se rompió aleatoriamente en trozos de 500 a 1.000 pares de bases mediante rotura me-  
cánica utilizando un dispositivo HydroShear (GeneMachines). En combinación con la sonicación pueden generar-  
se fragmentos más cortos. Para incrementar la eficiencia de la etapa posterior de ligación de extremos romos del  
35 adaptador, los polinucleótidos fragmentados se tratan opcionalmente con ADN polimerasa para rellenar o eliminar  
las regiones protuberantes, tal como se comenta en el Ejemplo 1. Se liga un exceso molar de 10X de un adapta-  
dor de extremos romos de 20 a 40 pb a los polinucleótidos diana fragmentados. La ligación se lleva a cabo bajo  
condiciones optimizadas para la ligación de extremos romos siguiendo las recomendaciones del fabricante. Únicamente  
se fosforiló la cadena antisentido del adaptador para garantizar la ligación unidireccional (además, la cadena  
40 antisentido incorpora un grupo 3'-dideoxi o un grupo bloqueante para evitar la concatamerización). El adaptador  
porta en su mismo extremo 3' un sitio de restricción para MmeI y un sitio de restricción para *EaeI* (o *EamI104I*)  
cadena arriba. La exclusión por tamaños o las columnas de afinidad de espín (por ejemplo la PCR de lavado Mi-  
nElute, de Qiagen) se utilizan convenientemente para eliminar el exceso de adaptadores, así como para cambiar el  
tampón.

45 La digestión con MmeI se llevó a cabo bajo condiciones óptimas para cortar 20 pares de bases de la molécula  
diana y asociarlas con el polinucleótido adaptador. Se utilizó una columna de centrifugación para cambiar el tampón  
(por ejemplo BioSpin-6, de BioRad). A continuación, se ligó un segundo polinucleótido adaptador, que contenía  
una secuencia de reconocimiento para la enzima de restricción *SfaNI*, al polinucleótido diana asociado al adaptador  
50 mediante la región protuberante generada con MmeI (debido a que MmeI genera un extremo 3'-protuberante de  
dos bases, el segundo adaptador es un grupo de adaptadores presentando conjuntamente las 16 permutaciones de dos  
bases en el extremo 3'-protuberante). El sitio *SfaNI* se posiciona para cortar dentro de las cuatro bases terminales de  
la secuencia diana. El adaptador se mantiene no fosforilado para eliminar la amplificación posterior de los polinucleótidos  
diana que contienen dos regiones protuberantes generadas con MmeI (un producto secundario de la digestión del  
55 polinucleótido diana). Se llevó a cabo una amplificación por PCR para amplificar los fragmentos que contenían el  
polinucleótido diana de 20 pb flanqueado por dos polinucleótidos adaptadores. Para proteger los sitios internos de la  
digestión posterior con *EaeI* y *SfaNI*, se llevó a cabo una PCR con dCTP metilada sustituyendo al dCTP normal en la  
mezcla de dNTP.

60 Se utilizó una columna de centrifugación de PCR de lavado para eliminar el exceso de cebadores y la ADN  
polimerasa, y para cambiar el tampón.

El material amplificado por PCR se digiere con *EaeI* (o *EamI104I*) y con *SfaNI* para generar las regiones protube-  
rantes necesarias para la ligación de secuencias de señal y para la conversión bioquímica. Mediante la utilización de  
65 cebadores biotinilados en la PCR, pueden eliminarse los fragmentos terminales digeridos a partir de la afinidad (por  
ejemplo utilizando perlas recubiertas de estreptavidina).

## ES 2 353 059 T3

### Ejemplo 3

#### *Principio de conversión cíclica en la que se convierten cuatro nucleótidos por ciclo*

5 El principio de conversión cíclica de bases de ácidos nucleicos en secuencias de señal legibles se demostró mediante la conversión de una secuencia de 12 bases en un ADN diana en su cadena de señal correspondiente. La conversión tuvo lugar en etapas mediante un procedimiento cíclico, en el que se convirtieron 4 bases en cada ciclo durante un total de 3 ciclos. El producto final fue una cadena de señal consistente de 12 componentes de señal que representan la secuencia de bases relevante en el ADN diana. El experimento se describe esquemáticamente en la figura 2.

10 Un fragmento de ADN diana de 240 pb, que contiene sitios internos para *Sfa*NI y *Eam*1104I se amplificó por PCR a partir del genoma del bacteriófago lambda (1507-1703) utilizando cebadores que contenían sitios de restricción para *Eam*1104I y *Sfa*NI (el fragmento superior en la figura 2), respectivamente. Con el fin de evitar el corte posterior de sitios internos, se llevó a cabo la reacción de PCR con nucleótidos desoxicitosina metilada (dCTP-m5) en lugar de la desoxicitosina (dCTP) habitual. Condiciones de la PCR (50  $\mu$ l): KCl 10 mM, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 10 mM, Tris-HCl 20 mM, Triton-X-100 al 0,1%, pH 8,8 (tampón Thermopol, New England Biolabs), Mg<sup>2+</sup> 2 mM, dNTP (-dCTP) 200  $\mu$ M y dCTP-m5 200  $\mu$ M (Amersham Pharmacia Biotech), 20 pmoles del cebador n° 328 de Lambda (5'-agactggc gatcctggcatcaccctccagcgtgtttat-3'; SEC ID n° 1) y 20 pmoles de cebador n° 329 de Lambda (5'-gcactgataggcgt cactcttcgctgtacgctgccagatg-3'; SEC ID n° 2) (MWG Biotech), 10 ng de genoma de Lambda, 1U de polimerasa Vent (New England Biolabs). El ciclado de PCR se llevó a cabo con un termociclador PTC-200 (MJResearch). Inicio en caliente: 95°C, 5 minutos, 35 ciclos consistentes de: 95°C, 15 segundos; 58°C, 20 segundos; 72°C, 30 segundos. Etapa de extensión completa: 72°C, 5 minutos.

#### Ciclo 1

25 Se introdujo el fragmento en el ciclo de secuenciación (figura 2, A) y se cortó con *Eam*1104I y *Sfa*NI (debido a que los sitios internos se encuentran metilados y protegidos frente al corte, éste tuvo lugar únicamente en las áreas de cebador). Condiciones de la digestión: NaCl 100 mM, Tris-HCl 50 mM, MgCl<sub>2</sub> 10 mM, DTT 1 mM, pH 7,9 (NEB3, New England Biolabs), 1  $\mu$ g de fragmento de PCR metilado, 20 U de *Eam*1104I, 3U de *Sfa*NI e incubación durante una hora a 37°C. Las enzimas se inactivaron por calor a 65°C durante 20 minutos. El fragmento cortado se lavó utilizando el sistema de PCR de purificación de GibcoBRL (Gibco) y después se eluyó con Tris-HCl 10 mM. El fragmento cortado corresponde a la naturaleza de los fragmentos que se obtienen tras una preparación inicial o un ciclo de conversión (descrito en el Ejemplo 1) en el aspecto de que presenta una región protuberante de 3 bases para la ligación de cadena de señal y una región protuberante de 4 bases para la ligación a un adaptador específico en el lado derecho.

35 La etapa de conversión, en la que se une el marcador de selección (figura 2, B), se llevó a cabo mediante la ligación del fragmento con el adaptador específico y su secuencia de señal asociada: Tris-HCl 40 mM, MgCl<sub>2</sub> 10 mM, DTT 10 mM, ATP 0,5 mM, pH 7,8 (tampón de ligasa, Fermentas), 1,6 pmoles de adaptador 1 (5'-biotina-gcctagtctgatgaacgcatcg-3'; SEC ID n° 3 hibridado a 5'-tggacgatgcgttcacgacacctagc-3'; SEC ID n° 4 (MWG-Biotech), 1,9 pmoles de secuencia de señal 1 (1), 7,5 Weiss-U de ADN ligasa de T4 (Fermentas). La incubación tuvo lugar a temperatura ambiente (22°C) durante 2,5 horas. La ligasa se inactivó por calor a 65°C durante 10 minutos.

45 Las secuencias de señal utilizadas en el experimento se prepararon de la manera siguiente: los componentes de señal (cada uno representando a una base) se prepararon mediante hibridación de oligonucleótidos (de aproximadamente 45 bases) que se ligaron entre sí para crear las secuencias de señal deseadas (los componentes de señal se construyeron con diferentes regiones protuberantes que dictan la secuencia de la ligación). Las secuencias de señal se amplificaron mediante PCR y se cortaron con *Sap*I con el fin de generar la región protuberante complementaria necesaria para la ligación de la secuencia de señal al ADN diana. Debido a que la totalidad de los fragmentos de ADN diana presentaban la misma región protuberante para la ligación de la secuencia de señal, todos los fragmentos en una reacción dada se ligaron a la misma secuencia de señal, con independencia de la secuencia de la región protuberante generada con *Sfa*NI. Por lo tanto, la especificidad y la selección se localizan en el adaptador. Debido a que éste se liga al extremo del ADN diana que se cortó con *Sfa*NI, únicamente se ligan los fragmentos que presentan una región protuberante complementaria al adaptador. Por lo tanto, una conversión se basa en el supuesto de que el ADN diana, además de ligarse a una secuencia de señal, también presenta una región protuberante complementaria al adaptador para que la ligación a ésta tenga lugar. En otras palabras, el adaptador garantiza una selección específica de fragmentos que presentan una secuencia de región protuberante correspondiente a la secuencia de señal ligada.

60 La etapa de selección de la conversión de bases consiste de la selección y amplificación de fragmentos a los que se ha ligado tanto una secuencia de señal como un adaptador específico (figura 2, C). Lo anterior se llevó a cabo utilizando una PCR bajo las condiciones siguientes (50  $\mu$ l): KCl 10 mM, (NH<sub>4</sub>)<sub>2</sub>SO<sub>4</sub> 10 mM, Tris-HCl 20 mM, Triton-X-100 al 0,1%, pH 8,8 (tampón Thermopol, New England Biolabs), Mg<sup>2+</sup> 2 mM, dNTP (-dCTP) 200  $\mu$ M y dCTP-m5 200  $\mu$ M (Amersham Pharmacia Biotech), 20 pmoles de adaptador1-cebador n° 332 (5'-biotina-ggctagtctgatgaacgcatcg-3'; SEC ID n° 5) y 20 pmoles de secuencia de señal 1-cebador n° 340 (5'-taatacgaactactatagcatgactcagcctcttcgca-3'; SEC ID n° 5) (MWG-Biotech), aproximadamente 3,5 pmoles de ADN diana ligado y 1U de polimerasa Vent (New England Biolabs). El ciclado de la PCR se llevó a cabo utilizando un termociclador PTC-200 (MJResearch). Inicio en caliente: 95°C, 2 minutos; 20 ciclos consistentes de: 95°C, 15 segundos; 66°C, 20 segundos; 72°C, 30 segundos.

## ES 2 353 059 T3

Etapas de extensión completa: 72°C, 5 minutos. El pocillo 2 en la figura 2 muestra los resultados del ciclo inicial del procedimiento de secuenciación. Se generó un fragmento correcto de 380 pb.

### Ciclo 2

5 Con el fin de convertir las siguientes 4 bases, se limpió el producto de PCR con el sistema de purificación Gibco-BRL (Gibco) y después se cortó con *Sfa*NI y *Eam*1104I. Lo anterior resulta posible debido a que la secuencia de señal y el adaptador del ciclo inicial contienen un sitio para *Eam*1104I y *Sfa*NI, respectivamente. Debido a que los sitios se encontraban localizados en la región del cebador, no resultaron bloqueados por la metilación durante la PCR. Para incrementar la eficiencia de la reacción de corte, los cortes se llevaron a cabo en serie bajo condiciones de corte óptimas. Corte con *Sfa*NI: NaCl 100 mM, Tris-HCl 50 mM, MgCl<sub>2</sub> 10 mM, DTT 1 mM, pH 7,9 (NEB3, New England Biolabs), 1 µg de fragmento de PCR metilado, 4U de *Sfa*NI. Incubación a 37°C durante 1 hora. Inactivación por calor de la enzima a 65°C durante 10 minutos. Se utilizó una columna Micro Bio-Spin 6 (BioRad) para limpiar el fragmento. Corte con *Eam*1104I: Tris-acetato 33 mM, acetato de magnesio 10 mM, acetato de potasio 66 mM, BSA 0,1 mg/ml, pH 7,9 (Tango Y<sup>+</sup>, Fermentas), fragmento de PCR digerido con *Sfa*NI, 10U de *Eam*1104I. Incubación a 37°C durante 1 hora, seguido de la inactivación por calor a 65°C durante 20 minutos. El producto digerido se limpió mediante la utilización del sistema de purificación GibcoBRL (Gibco). La etapa inicial de la conversión se llevó a cabo correspondiente al ciclo inicial, excepto en que se añadió una nueva secuencia de señal y su adaptador específico asociado (correspondiente a las siguientes 4 bases de la secuencia). Para reducir problemas potenciales de contaminación cruzada, la secuencia de señal 2 se diseñó con una región protuberante diferente de la de la secuencia de señal 1. Condiciones de la ligación: Tris-HCl 40 mM, MgCl<sub>2</sub> 10 mM, DTT 10 mM, ATP 0,5 mM, pH 7,8 (tampón ligasa, Fermentas), 0,5 pmoles de adaptador 2 (5'-biotina-cgacagtacgacggaccagcatcc-3'; SEC ID n° 6 hibridada a 5'-acgcgatgctgctcctgctactgtcg-3'; SEC ID n° 7) (MWG-Biotech), 1 pmol de secuencia de señal 2 (2), 1 Weiss-U de ADN ligasa de T4 (Fermentas). La incubación tuvo lugar a temperatura ambiente (22°C) durante 1,5 horas. La ligasa se inactivó por calor a 65°C durante 10 minutos. La amplificación por PCR tuvo lugar bajo las mismas condiciones que en el ciclo inicial, excepto en que se utilizó un conjunto nuevo de cebadores (adaptador2-cebador n° 347: 5'-biotina-cgacagtacgacggaccagcatcc-3'; SEC ID n° 6' y secuencia de señal 2-cebador n° 343: 5'-taatacactactatagcatcgaatgaccgcctctccact-3'; SEC ID n° 8). Son preferibles conjuntos de cebadores diferentes para cada ciclo de secuenciación con el fin de minimizar el peligro de amplificar cualquier residuo del ciclo de secuenciación anterior. El resultado de la amplificación por PCR en el segundo ciclo se muestra en el pocillo 3 en la figura 4. Se generó un fragmento correcto de 523 pb.

### Ciclo 3

35 La conversión de las siguientes 4 bases y la amplificación por PCR siguieron el mismo patrón descrito para el ciclo 2. Tras el corte en serie con *Sfa*NI y *Eam*1104I, el fragmento se ligó con la secuencia de señal 3 (3) y con el adaptador 3 (5'-biotina-atcgagcctggcatagcagcatca-3'; SEC ID n° 9 hibridado con 5'-aaactgatgctatgccaggtcgtgat-3'; SEC ID n° 10) (MWG-Biotech). La amplificación por PCR se llevó a cabo utilizando el conjunto de cebadores siguiente: adaptador3-cebador n° 353 (5'-atcgagcctggcatagcagcatca-3'; SEC ID n° 9) y la secuencia de señal 3-cebador n° 345 (5'-taatacactactatagcaccgggagcatagactcttcaggt-3'; SEC ID n° 11). El resultado de la amplificación en el ciclo 3 se muestra en el pocillo 4, en la figura 4. El pocillo 4 muestra que se formaron dos bandas débiles y una relativamente fuerte. La banda fuerte se aproximada a la longitud esperada, de 666 pb. Las bandas débiles, del tamaño incorrecto, que pueden observarse en los pocillos 4 (ciclo 3) y 3 (ciclo 2), muy probablemente son el resultado de problemas de contaminación cruzada y del cebado incorrecto. La banda débil en el pocillo 3 y las bandas débiles más abajo, en el pocillo 4, corresponden al tamaño de un fragmento de PCR del ciclo anterior. Una posible explicación es que fragmentos no cortados por completo en un ciclo pueden funcionar como moldes para el siguiente ciclo de la PCR. Aunque la utilización de conjuntos nuevos de cebadores en cada ciclo reduce este problema en cierto grado, todavía se encuentra presente el peligro de un cebado incorrecto debido a los sitios de enzimas que presentan en común los cebadores. Sin embargo, este tipo de cebado incorrecto puede eliminarse mediante la utilización de condiciones de hibridación más astringentes durante la PCR (por ejemplo Mg<sup>2+</sup>, temperatura), la utilización de polimerasas más discriminantes, el alejamiento del sitio *Eam*1104I respecto del extremo 3' en el cebador de la cadena de secuencia, la elección de nuevas secuencias o la inmovilización de los fragmentos (por ejemplo el sistema biotina-estreptavidina sobre microperlas) antes del corte, de manera que se garantice que han sido cortados los fragmentos que entran en el ciclo siguiente. La mejor solución para otros tipos de cebado incorrecto sería seleccionar nuevas secuencias de cebador y/o optimizar las condiciones de la PCR. Los sucesos de cebado incorrecto serían causados por la presencia de material de ADN superfluo (por ejemplo un exceso de adaptadores o de ADN no ligado) puede eliminarse mediante la utilización de una estrategia de tio-protección de la cadena retrasada del adaptador. La digestión con una 5'-3' exonucleasa (por ejemplo la exonucleasa de T7 o la exonucleasa de Lambda) previamente a la PCR sólo dejará intacta una cadena de ADN ligada al adaptador. En formas de realización en las que todas las secuencias diana son de igual longitud, puede utilizarse la selección por tamaño para eliminar los fragmentos de longitud incorrecta durante el procedimiento de conversión bioquímica.

### Ejemplo 4

#### *Principio de conversión cíclica en la que se convierten tres nucleótidos por ciclo*

65 El principio de conversión cíclica de las bases de ácidos nucleicos en una secuencia de señal legible fue demostrado mediante la conversión de una secuencia de 12 bases presente en un ADN diana en su secuencia de señal correspondiente. La conversión tuvo lugar en etapas, mediante un procedimiento cíclico, en el que se convirtieron 3 bases en

## ES 2 353 059 T3

cada ciclo durante un total de 4 ciclos. El producto final fue una secuencia de señal que consistía de 12 componentes de señal que representaban la secuencia de bases relevante en el ADN diana. El experimento se describe esquemáticamente en la figura 5.

5 Se generó un fragmento de ADN diana sintético de 66 pb mediante la hibridación de los oligonucleótidos siguientes: n° 003 (5'-PHO-GATCTTGGCTATTCGTCTCTTGGCTTTTCGTCTGATTGTAGACGCCAACGGGACATGAT  
GATGAT-3'; SEC ID n° 12) y n° 004 (5'-[PHO-CATCATCATCATCATGTCCCCTTGGCGTCTACAATCAGAC  
10 GAAAAGCCAAGAGACGAATAGCCAAG-3'; SEC ID n° 13). Tal como resulta evidente a partir de la secuencia, la molécula diana contiene en el extremo 3' cuatro tripletes ATP consecutivos. La molécula hibridada corresponde a la naturaleza de aquellos fragmentos que se obtienen tras una preparación inicial (descrita en el Ejemplo 2) en el aspecto de que presenta una región protuberante de 3 bases para la ligación de secuencia de señal y una región protuberante de 4 bases para la ligación a un adaptador específico.

15 La etapa de conversión en la que se unía el marcador de selección (figura 5, B) se llevó a cabo mediante la ligación de 1 pmol del fragmento con el adaptador específico y su secuencia de señal asociada: Tris-HCl 50 mM, MgCl<sub>2</sub> 10 mM, DTT 10 mM, ATP 1 mM, 25 µg/ml de BSA, pH 7,5 (tampón ligasa, New England Biolabs), 1 pmol de adaptador n° 005 (5'-GATGTAGATGCACTCCCGGACCTC-3'; SEC ID n° 14 hibridada a n° 006: 5'-GAGGTCCGGGAGTGCATC  
TA-3'; SEC ID n° 15) (MWG-Biotech), 1 pmol de secuencia de señal n° 001 (5'-TGTGTCCGCGTGGCTCTTCT  
20 GATCTTGGCTTTTCGTCTCTTGGCTATTCGTCT-3'; SEC ID n° 16, hibridada a n° 002: 5'-PHO-ATCAGACGAA  
TAGCCAAGAGACGAAAAGCCAAGATCAGAAGAGCCACGCGGACACA-3'; SEC ID n° 17) (MWG-Biotech),  
100 unidades de ADN ligasa de T4 (New England Biolabs). La incubación tuvo lugar a 22°C en un termociclador  
PTC-200 (MJResearch) durante 1 hora. La selección y amplificación de aquellas moléculas que se habían ligado con  
éxito tanto al adaptador como al adaptador de la secuencia de señal se llevaron a cabo utilizando una PCR bajo las  
condiciones siguientes (50 µl): KCl 10 mM, (NH<sub>4</sub>)SO<sub>4</sub> 10 mM, Tris-HCl 20 mM, Triton-X-100 al 0,1%, pH 8,8  
25 (tampón Thermopol, New England Biolabs), Mg<sup>2+</sup> 2 mM, 200 µM de dNTP (Amersham Bioscience), 10 pmoles de  
cebador de adaptador n° 006 (5'-GAGGTCCGGGAGTGCATCTA-3'; SEC ID n° 18) (MWG-Biotech) y 10 pmoles  
del cebador del adaptador de señal, n° 007 (5'-TGTGTCCGCGTGGCTCTTCT-3'; SEC ID n° 19) (MWG-Biotech),  
aproximadamente 1 pmol de ADN diana ligado y 0,2U de polimerasa Vent (New England Biolabs). El ciclado de  
PCR se llevó a cabo con un termociclador PTC-200 (MJResearch). Inicio en caliente: 95°C, 2 minutos, 25 ciclos que  
30 consisten en: 95°C, 15 segundos; 59,3°C, 20 segundos; 72°C, 20 segundos. Etapa de extensión completa: 72°C, 60  
segundos. El pocillo 2, en la figura 6, muestra los resultados del ciclo inicial del procedimiento de secuenciación. Se  
generó el fragmento correcto de 142 pb.

### Ciclo 2

35 Con el fin de convertir las siguientes 3 bases, se limpió el producto de PCR utilizando un kit de PCR de limpieza  
MinElute (Qiagen) y después se cortó con *Eam*1104I y *Sfa*NI. Para incrementar la eficiencia de la reacción de corte,  
los cortes se llevaron a cabo en serie bajo condiciones óptimas de corte.

40 Corte de *Eam*1104I: Tris-acetato 33 mM, acetato de magnesio 10 mM, acetato de potasio 66 mM, BSA 0,1 mg/ml,  
pH 7,9 (Tango Y+, Fermentas), 0,5 µg de fragmento de PCR, 20 unidades de *Eam*1104I en un volumen de reacción de  
25 µl. Incubación a 37°C durante 1 hora. Se utilizó una columna Micro Bio-Spin 6 (BioRad) para limpiar el fragmento.  
Corte con *Sfa*NI: NaCl 100 mM, Tris-HCl 50 mM, MgCl<sub>2</sub> 10 mM, DTT 1 mM, pH 7,9 (NEB3, New England Biolabs),  
0,5 µg de fragmento de PCR, 2,5 unidades de *Sfa*NI. Incubación a 37°C durante 15 minutos en un volumen de reacción  
45 de 50 µl.

El producto digerido se limpió mediante PAGE (electroforesis en gel de poliacrilamida) y se purificó a partir del  
gel utilizando un procedimiento de trituración e inmersión. Las partículas de gel se eliminaron utilizando una columna  
Micro Bio-Spin 6 y el ADN se purificó utilizando una columna de limpieza de reacción MinElute (Qiagen).

50 Se llevó a cabo la etapa de conversión correspondiente al ciclo inicial utilizando la misma secuencia de señal y  
su adaptador específico asociado. El resto del procedimiento siguió el mismo patrón descrito para el ciclo 1, excepto  
en que se utilizó un gel de agarosa y un kit de purificación en gel (QiaexII, Qiagen) para purificar la muestra tras la  
digestión. El resultado de la amplificación por PCR en el segundo ciclo se muestra en el pocillo 3, en la figura 6. Se  
55 generó el fragmento correcto de 172 pb.

Los ciclos 3 y 4 siguieron el mismo patrón descrito para el ciclo 2. El resultado de la amplificación por PCR se  
muestra en los pocillos 4 y 5, en los que se generaron los fragmentos correctos de 202 pb y 232 pb.

60

65

REIVINDICACIONES

1. Procedimiento para determinar la secuencia de un polinucleótido diana, que comprende las etapas siguientes:

- 5
- (i) tratar una muestra de un polinucleótido diana bicatenario para crear regiones protuberantes en cada extremo, una de las cuales debe secuenciarse, presentando cada región protuberante un número definido de bases,
- 10
- (ii) dividir la muestra en compartimientos de reacción y poner en contacto cada muestra separada con una secuencia polinucleótida de señal bicatenaria diferente y un polinucleótido adaptador bicatenario correspondiente, representando cada secuencia de señal una secuencia polinucleótida específica de la misma longitud que la de la región protuberante que debe secuenciarse, y que comprende una región protuberante que permite la hibridación y la ligación con el extremo del polinucleótido diana situado enfrente del que se está secuenciando, y comprendiendo cada adaptador una región protuberante que presenta una secuencia complementaria a la secuencia representada por la secuencia de señal, hibridándose el adaptador con la región protuberante en la muestra diana que debe secuenciarse únicamente en el caso de que la región protuberante del adaptador sea complementaria a la región protuberante que se está secuenciando,
- 15
- 20
- (iii) llevar a cabo la reacción de polimerasa de la muestra o muestras utilizando cebadores que se hibridan con los extremos de la secuencia de señal y de la secuencia adaptadora, en la que el producto de la reacción de polimerasa comprende un sitio de restricción que permite el corte del adaptador para formar una nueva región protuberante que debe secuenciarse, opcionalmente mediante la repetición de las etapas (i) a (iii) utilizando enzimas de restricción para crear las regiones protuberantes, e
- 25
- (iv) identificar qué secuencias de señal se encuentran presentes en los productos amplificados, y en qué orden, para determinar de este modo la secuencia del polinucleótido diana.

30 2. Procedimiento según la reivindicación 1, en el que la región protuberante que se liga a la secuencia de señal presenta por lo menos 3 bases.

35 3. Procedimiento según la reivindicación 1 ó 2, en el que la región protuberante que debe secuenciarse presenta 4 bases.

40 4. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la combinación de la totalidad de las secuencias representadas por la secuencia de señal en cada compartimiento de reacción corresponde a la totalidad de las permutaciones de una secuencia que comprende el número de bases en la región protuberante que debe secuenciarse.

45 5. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el sitio de restricción es específico para una enzima de clase II.

6. Procedimiento según la reivindicación 5, en el que la enzima es *Sfa*NI o *Ear*I.

7. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que la reacción de polimerasa se lleva a cabo utilizando metil-dCTP como sustituto de dCTP.

8. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que el adaptador se inmoviliza sobre un material de soporte.

50 9. Procedimiento según cualquiera de las reivindicaciones anteriores, en el que secuencias de señal secuenciales, y opcionalmente los adaptadores secuenciales, comprenden sitios de reconocimiento para diferentes oligonucleótidos cebadores.

55

60

65

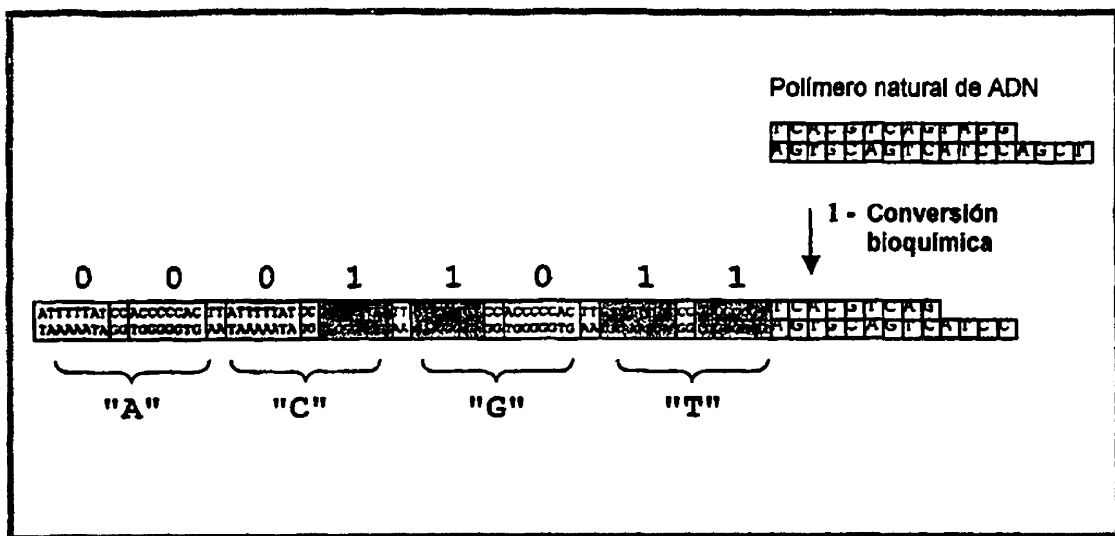


Figura 1

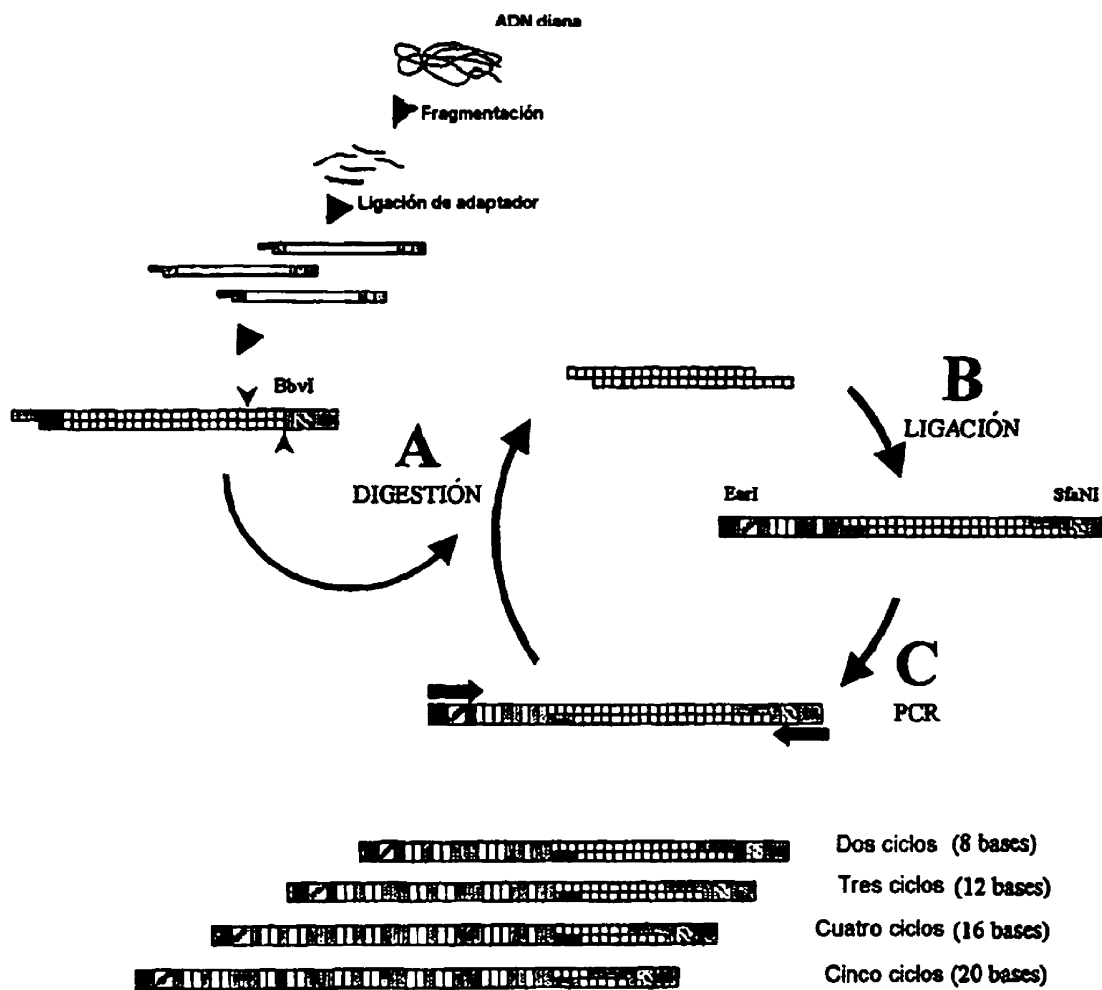


Figura 2

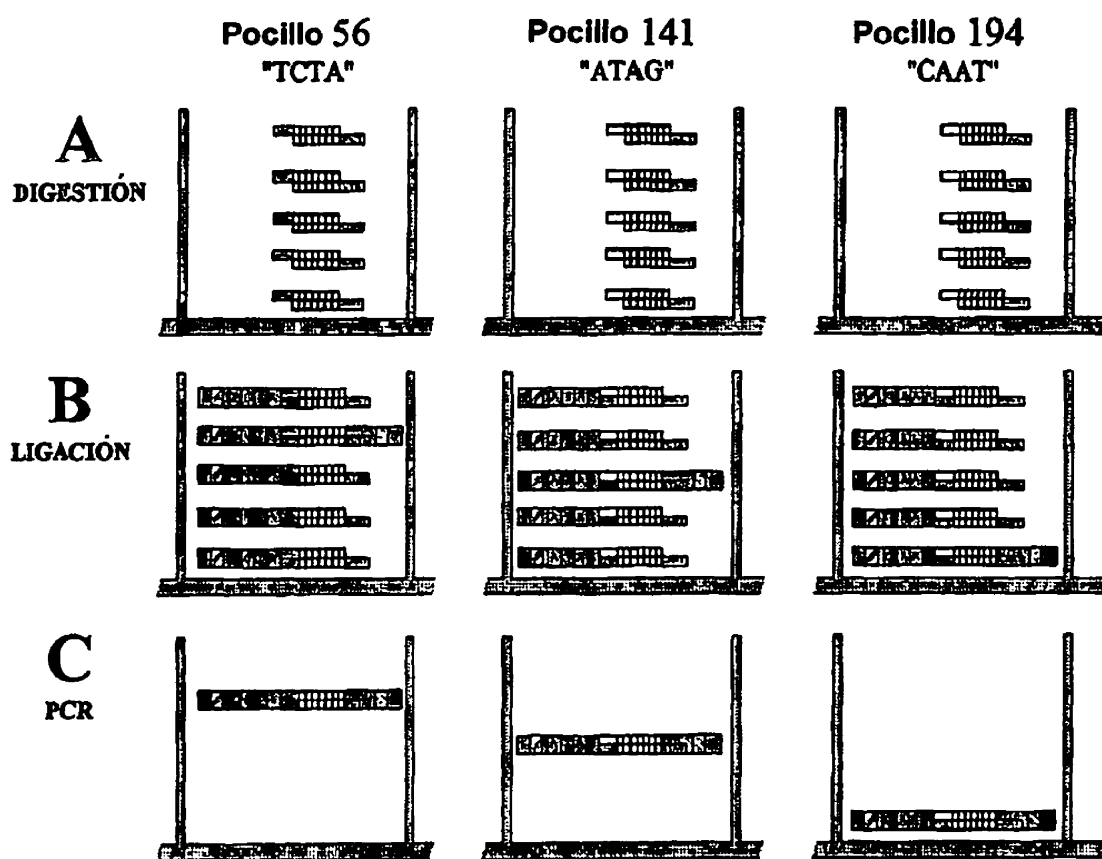


Figura 3

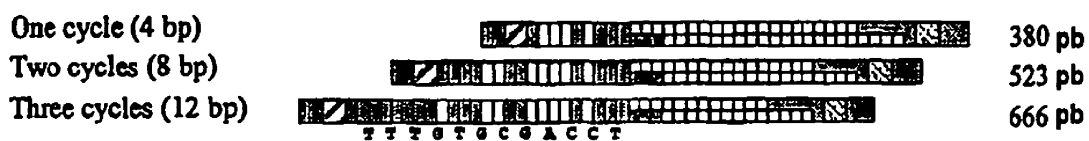
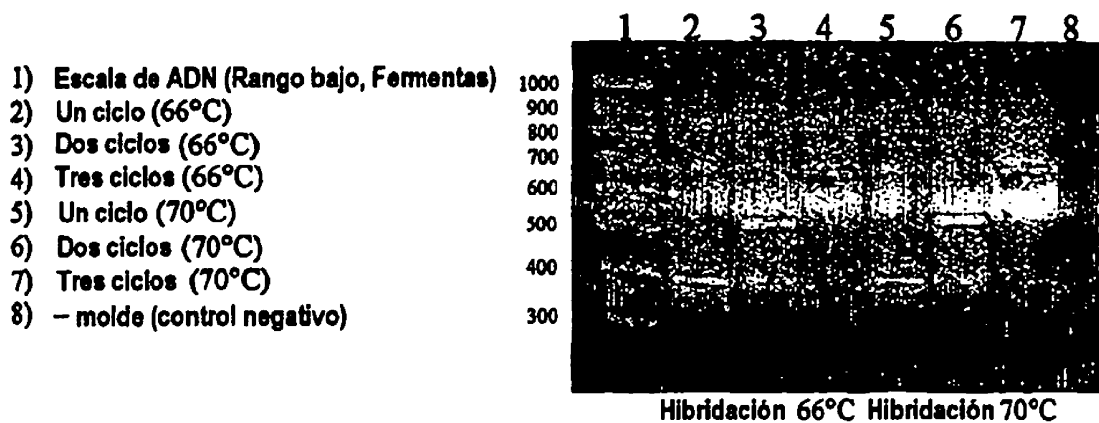


Figura 4

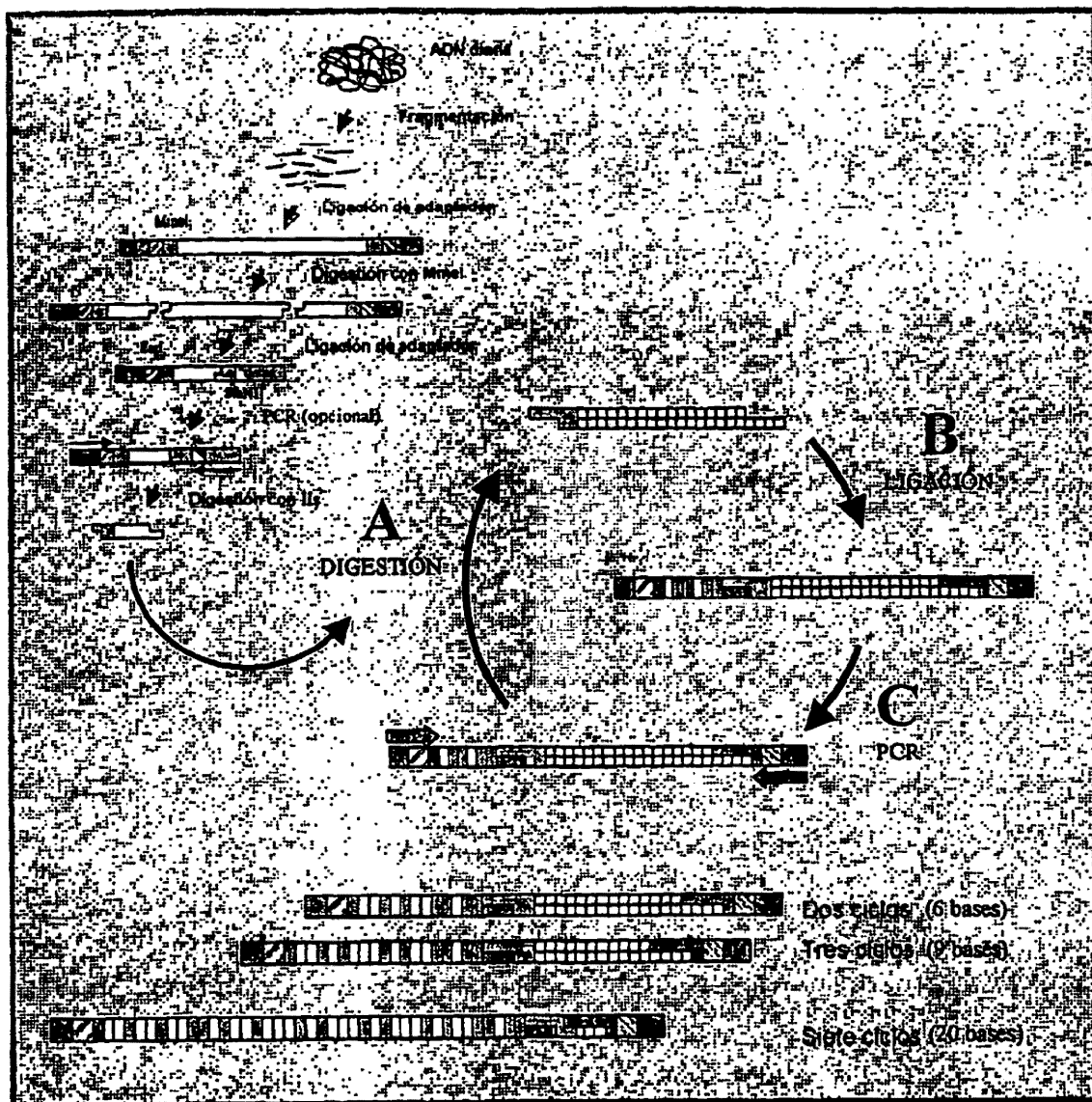


Figura 5

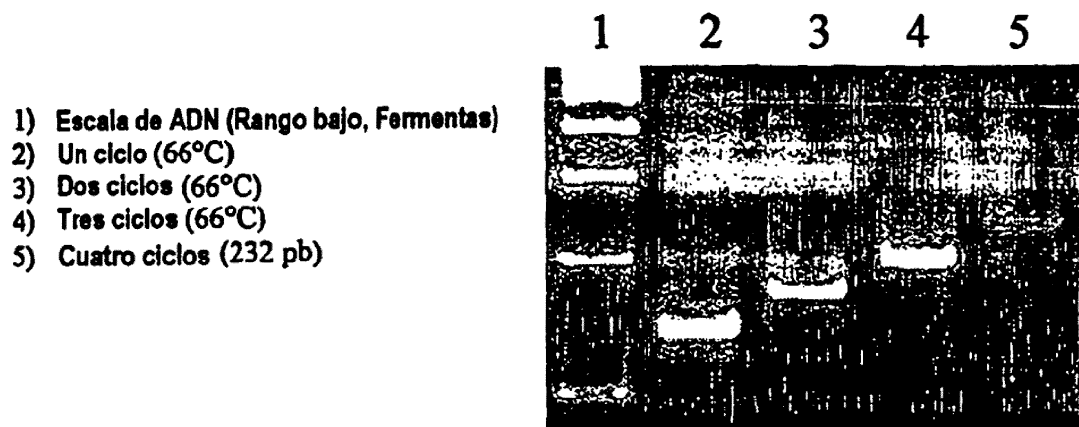


Figura 6