



19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

11 Número de publicación: **2 364 920**

51 Int. Cl.:
C12N 15/09 (2006.01)
C12N 15/10 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Número de solicitud europea: **08017649 .8**
96 Fecha de presentación : **08.10.2008**
97 Número de publicación de la solicitud: **2175018**
97 Fecha de publicación de la solicitud: **14.04.2010**

54 Título: **Proceso de clonación limpia.**

45 Fecha de publicación de la mención BOPI:
16.09.2011

45 Fecha de la publicación del folleto de la patente:
16.09.2011

73 Titular/es: **Icon Genetics GmbH**
Briennerstrasse 12A
80333 München, DE

72 Inventor/es: **Marillonnet, Sylvestre;**
Engler, Carola;
Kandzia, Romy;
Thieme, Frank y
Weber, Ernst

74 Agente: **Ungria López, Javier**

ES 2 364 920 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

Proceso de clonación limpia

5 **Campo de la invención**

La presente invención se refiere a un proceso de clonación de una secuencia de ácido nucleico que flanquea a una secuencia de ácido nucleico de secuencia de nucleótidos conocida en ADN genómico o ARN. La invención también se refiere a un proceso de inserción de una secuencia de ácido nucleico de interés, que puede ser de secuencia de nucleótidos desconocida, en un ácido nucleico aceptor o vector aceptor. La secuencia de ácido nucleico de secuencia de nucleótidos desconocida puede secuenciarse después de la etapa de inserción o clonación, por ejemplo, para determinar la localización de la secuencia de ácido nucleico de secuencia de nucleótidos conocida en el genoma.

15 **Antecedentes de la invención**

Desde su desarrollo, la reacción en cadena de la polimerasa (PCR) ha revolucionado el campo de la biología molecular. La PCR es un proceso por el que cualquier secuencia de ADN de interés (de un tamaño razonable) localizada entre dos secuencias conocidas puede amplificarse usando dos cebadores homólogos a estas dos secuencias conocidas, siendo un cebador un cebador directo y siendo el otro cebador (que se une a la cadena complementaria) un cebador inverso. La PCR permite clonar fácilmente cualquier secuencia de interés, con tal de que estén disponibles dos secuencias flanqueantes conocidas. No es necesario ningún conocimiento de la secuencia localizada entre los dos sitios de unión a cebador.

Un problema relacionado aunque diferente en la biología molecular es la identificación de secuencias desconocidas que flanquean a una región de secuencia de nucleótidos conocida. La PCR no puede usarse directamente para amplificar un fragmento que contiene la secuencia conocida y desconocida, puesto que se conoce la secuencia en sólo un extremo del fragmento a amplificar. Los ejemplos de problemas de este tipo incluyen la determinación de la secuencia de nucleótidos que flanquea a un transgén integrado de forma estable (por ejemplo, en un ADN-T), la secuencia de nucleótidos que flanquea a una inserción de transposón, o la secuencia de nucleótidos de la región variable de un anticuerpo para el que sólo se conoce el isotipo. Los dos primeros ejemplos se refieren a moléculas de ADN, mientras que el tercer ejemplo se refiere a moléculas de ARN. Esta diferencia no es importante ya que las moléculas de ARN pueden convertirse en ADNc por transcripción inversa usando un cebador que se una a la región conocida.

Con los años, se han desarrollado muchos protocolos para proporcionar soluciones para identificación de secuencias desconocidas que flanquean a secuencias conocidas. Muchos de estos protocolos usan la unión de un adaptador al extremo de la secuencia desconocida o usan una PCR que usa uno o varios cebadores inespecíficos (cebadores que contienen de hecho tanto una secuencia constante conocida, una secuencia adaptadora, seguida de una secuencia variable inespecífica o aleatoria) que se unen aleatoriamente al ADN, incluyendo a secuencias en las proximidades de la secuencia conocida. Después se realizan dos o tres PCR usando combinaciones de cebadores adaptadores y cebadores específicos de región conocida (o cebadores específicos de gen, abreviados en la presente memoria mediante "gsp"). Después de la primera PCR, se obtienen típicamente productos tanto específicos como inespecíficos. La proporción de productos específicos aumenta en la segunda PCR realizada usando un cebador adaptador y un cebador específico de secuencia conocida anidado, pero todavía están presentes muchos productos inespecíficos. La identificación de la secuencia desconocida puede realizarse secuenciando el producto amplificado. Sin embargo, si se espera que se amplifiquen en la misma amplificación varios productos específicos (por ejemplo, un genoma podría contener varios transgenes o varios transposones, o una población de ARN podría contener un gran número de anticuerpos diferentes), la secuenciación directa no será útil. Más bien, el producto amplificado tendrá que clonarse y los plásmidos recombinantes secuenciarse individualmente.

Hay muchas estrategias para la clonación de productos de PCR. La clonación se realiza típicamente ligando entre sí fragmentos de ADN que se han preparado por digestión con enzimas de restricción de tipo II. Este proceso requiere habitualmente varias etapas: (1) los plásmidos (o productos de PCR) que contienen el fragmento a subclonar o el vector receptor se digieren con una o dos enzimas de restricción, (2) los fragmentos digeridos se separan usando electroforesis en gel y, después, los fragmentos deseados se extraen del gel, (3) los fragmentos de vector purificado e inserto se ligan entre sí usando una ADN ligasa tal como la ADN ligasa T4, (4) la ligación se usa para transformar células de *E. coli* competentes.

Una limitación de técnicas de clonación convencionales es que los sitios de restricción seleccionados para la clonación no deben estar presentes dentro del fragmento a clonar. Puesto que para el problema analizado en la presente memoria, no se conoce una parte de la secuencia a clonar, puede esperarse que el uso de enzimas de restricción para la clonación dé como resultado la pérdida de parte de la secuencia de algunos de los fragmentos amplificados, o incluso impida completamente la clonación de algunos de los productos.

Se ha desarrollado una estrategia de clonación totalmente diferente que no requiere enzimas de restricción: esta estrategia depende de la generación de extremos de ADN que son monocatenarios tanto en el inserto de ADN como en el vector. Los salientes monocatenarios complementarios en el inserto de ADN y en el vector de clonación hibridarán. Si la región hibridada es de más de 12-15 nucleótidos y los dos extremos de un inserto lineal hibridan con los dos extremos de un vector plasmídico linealizado, puede conseguirse una clonación sin ligación. Las células hospedadoras tales como células de *E. coli* transformadas con el producto hibridado repararán las mellas, conduciendo a la formación de un plásmido circular capaz de replicarse (Li y Evans, 1997, *Nucleic Acids Res.*, 25, 4165-4166).

Uno de los primeros métodos de clonación que se desarrollaron basándose en este principio fue la clonación con UDG. Se realizó la amplificación por PCR de un fragmento de ADN usando cebadores que contenían una extensión de 12 nucleótidos arbitrarios que contenían al menos 4 uracilos. El vector también se amplificó por PCR usando cebadores que contenían una extensión de 12 nucleótidos complementaria a la extensión en los cebadores para el inserto. Después de la amplificación por PCR, el inserto y el vector se trataron con uracilo ADN glicosilasa, que abre las extensiones de ADN en el vector y en el inserto (la UDG cataliza la hidrólisis del enlace N-glicosílico entre el uracilo y el azúcar), creando extremos de ADN monocatenarios. Después de hibridar el vector y el inserto, la mezcla se usó para transformar *E. coli* usando células químicamente competentes. Después, las *E. coli* recortan y reparan los extremos en el sitio de unión (Nisson P. E., Rashtchian A., y Watkins P. C., 1991, *PCR Methods Appl.*, 1: 120-123).

La ventaja de esta estrategia es que es eficaz y da como resultado muy pocas construcciones de vectores vacíos. Los inconvenientes son que (1) los cebadores que contienen uracilo son caros, (2) el vector completo tiene que amplificarse por PCR y (3) la extensión de cebadores tiene que contener 4 uracilos y 4 adeninas (complementarias a los uracilos en la extensión del fragmento complementario) y, por lo tanto, no puede consistir en cualquier secuencia de 12 nucleótidos de elección.

Otra estrategia desarrollada todavía anteriormente, la clonación independiente de ligación (LIC), se desarrolló basándose en la actividad exonucleasa de 3' a 5' de la polimerasa T4 (Aslanidis C. y de Jong, P. J., 1990, *Nucleic Acids Res.*, 18: 6069-6074). Como con la clonación con UDG, se amplifica un fragmento de PCR con dos cebadores que contienen extensiones de 12 nucleótidos, careciendo estas extensiones de uno de los 4 nucleótidos, por ejemplo G. Este fragmento amplificado se trata después con polimerasa T4 en un tampón que contiene dGTP pero ninguno de los otros nucleótidos. La actividad exonucleasa de 3' a 5' de la polimerasa T4 elimina todos los nucleótidos hasta que se encuentra la primera G, punto en el cual se detiene puesto que se obtiene un equilibrio entre eliminación e incorporación de este nucleótido. Por lo tanto, este tratamiento crea una extensión monocatenaria de 12 nt (nt representa nucleótido en la presente memoria) que puede usarse para la hibridación de vector e inserto. El vector contiene extremos que son compatibles con el inserto y se generan por digestión con polimerasa T4 en presencia de dCTP.

En un trabajo similar, se trató el ADN y el inserto con polimerasa T4, pero sin adición de ningún nucleótido (Yang et al., 1993, *Nucleic Acids Res.*, 21: 1889-1893). Puesto que la extensión monocatenaria podría ser más larga de 12-15 nucleótidos, la mezcla hibridada se trató por adición de los 4 desoxinucleótidos y de la polimerasa T4 para rellenar los huecos monocatenarios.

Otro procedimiento similar a ambas técnicas descritas anteriormente (clonación independiente de ligación y secuencia: SLIC) se describe en el documento US 2007/0292954. El procedimiento es similar al trabajo de Yang et al, pero los huecos en el heterodúplex hibridado no se rellenan con ADN polimerasa. Este trabajo también describe el ensamblaje de hasta 10 fragmentos en un vector.

Estas tres estrategias relacionadas son eficaces y permiten la clonación sin usar sitios de restricción convencionales en los sitios de unión entre vector e inserto. Sin embargo, pueden clonarse productos de amplificación inespecíficos obtenidos cuando uno de los cebadores hibrida con una secuencia aleatoria inespecíficamente o dímeros de cebadores además de los productos específicos. Por lo tanto, existe la necesidad de una estrategia de clonación más limpia.

55 Descripción general de la invención

Partiendo de la técnica anterior, un problema de la invención es proporcionar un proceso de clonación de una secuencia de ácido nucleico de interés que comprende un segmento de secuencia U de secuencia de nucleótidos desconocida y un segmento de secuencia de secuencia conocida en un ácido nucleico aceptor tal como un vector de clonación, por lo que se suprime la clonación de productos de PCR inespecíficos.

Este problema se resuelve mediante los procesos descritos en la presente memoria. Por lo tanto, la invención proporciona:

- 65 (1) Un proceso de inserción de una secuencia de ácido nucleico de interés en un ácido nucleico aceptor, que comprende las etapas siguientes:

- amplificar por PCR un ADN que comprende, en el siguiente orden, un segmento de secuencia U, un segmento de secuencia de ácido nucleico K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3 usando un cebador directo que define un primer extremo del ADN amplificado y un cebador inverso que define un segundo extremo del ADN amplificado,
- 5 tratar las moléculas de ADN bicatenario lineales contenidas en el producto de PCR obtenido en la etapa anterior con una exonucleasa para obtener un saliente monocatenario en el primer extremo del ADN y un saliente monocatenario que comprende los segmentos de ácido nucleico K2 y K3 en el segundo extremo del ADN;
- 10 hibridar el producto de la etapa anterior con un ácido nucleico aceptor bicatenario linealizado que tiene en un primer extremo del mismo un saliente monocatenario complementario al saliente monocatenario del primer extremo del ADN y, en un segundo extremo del mismo, un saliente monocatenario complementario al segmento de secuencia monocatenaria K2 del segundo extremo del ADN; y usar el producto de reacción obtenido en la etapa anterior para transformar una célula hospedadora.
- 15 (2) El proceso de acuerdo con (1), que comprende además generar un molde para dicha PCR, que comprende unir una secuencia de unión a cebador a una secuencia de ácido nucleico que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3, hibridando dicho cebador directo con la secuencia de unión a cebador.
- 20 (3) El proceso de acuerdo con (2), en el que la secuencia de unión a cebador es un segmento de secuencia de nucleótidos homooligomérica unido usando una desoxirribonucleótido transferasa terminal.
- 25 (4) El proceso de acuerdo con (2) o (3), en el que dicho cebador directo comprende un segmento de secuencia adaptadora y un segmento de secuencia complementario a la secuencia de unión a cebador.
- 30 (5) El proceso de acuerdo con (1), que comprende además generar un molde para dicha PCR, que comprende unir una secuencia adaptadora a una secuencia de ácido nucleico que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U en el primer extremo de la secuencia de ácido nucleico, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3.
- 35 (6) El proceso de acuerdo con uno cualquiera de (1) a (5), en el que dicho ácido nucleico aceptor no contiene una porción de secuencia complementaria al segmento de secuencia monocatenaria K3 en el saliente monocatenario en el segundo extremo del ácido nucleico aceptor.
- 40 (7) El proceso de acuerdo con uno cualquiera de (1) a (6), en el que la exonucleasa es una exonucleasa 3'-5'. Son ejemplos para la exonucleasa la ADN polimerasa T4 de *E. coli*, el fragmento grande de la polimerasa I fragmento grande de *E. coli* (polimerasa Klenow), nucleasa lambda, nucleasa T7 o exonucleasa III, preferentemente es la ADN polimerasa T4 de *E. coli*.
- 45 (8) El proceso de acuerdo con uno cualquiera de (1) a (7), en el que dicha célula hospedadora es una célula bacteriana tal como *E. coli* o *Agrobacterium tumefaciens*.
- 50 (9) El proceso de acuerdo con uno cualquiera de (1) a (8), en el que se realiza una etapa de preamplificación por PCR antes de dicha reacción de PCR definida en la reivindicación 1, usando un cebador directo que define el primer extremo del producto de PCR de la etapa de preamplificación, y un cebador inverso que termina en su extremo 3' en una secuencia de nucleótidos de segmento de secuencia de ácido nucleico K3.
- 55 (10) El proceso de acuerdo con (2) ó (5), en el que la etapa de unión se realiza en una mezcla que comprende ADN genómico aislado a partir de células procariotas o eucariotas.
- (11) El proceso de acuerdo con (2) ó (5), en la que dicha etapa de unión se realiza en el extremo 3' de un ADNc retrotranscrito a partir de un ARNm.
- 60 (12) El proceso de acuerdo con cualquiera de (1) a (11), en el que el ácido nucleico aceptor no contiene un segmento de secuencia de más de 10 nt que tenga homología con el segmento K3 dentro de un intervalo hecho monocatenario en la etapa de tratamiento con exonucleasa del segundo extremo del ácido nucleico aceptor, tal como dentro de una región de 100 nt desde el segundo extremo del ácido nucleico aceptor.
- 65 (13) Un proceso de inserción de una secuencia de ácido nucleico de interés, que comprende un segmento de secuencia de ácido nucleico U de secuencia de nucleótidos desconocida en un ácido nucleico aceptor, que comprende las etapas siguientes:
 unir una secuencia de unión a cebador a una secuencia de ácido nucleico que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U en el primer extremo de la secuencia de ácido nucleico, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de

secuencia de ácido nucleico de secuencia conocida K3;
 amplificar por PCR un ADN que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U, segmento de secuencia de ácido nucleico K2 y un segmento de secuencia de ácido nucleico K3 usando un cebador directo que hibrida con la secuencia de unión a cebador y que define un primer extremo del ADN amplificado, y un cebador inverso que define un segundo extremo del ADN amplificado, terminando dicho cebador inverso en su extremo 3' en una secuencia de nucleótidos de segmento de secuencia K3;
 tratar las moléculas de ADN bicatenario lineales contenidas en el producto de PCR obtenido en la etapa anterior con una exonucleasa para obtener un saliente monocatenario en el primer extremo del ADN y un saliente monocatenario que comprende los segmentos de ácido nucleico K2 y K3 en el segundo extremo del ADN;
 hibridar el producto de la etapa previa con un ácido nucleico aceptor bicatenario linealizado que tiene en un primer extremo del mismo un saliente monocatenario complementario al saliente monocatenario del primer extremo del ADN y en un segundo extremo del mismo un saliente monocatenario complementario al segmento de secuencia monocatenaria K2 del segundo extremo del ADN; y
 usar el producto de reacción obtenido en la etapa anterior para transformar una célula hospedadora.

(14) Un proceso de inserción de una secuencia de ácido nucleico de interés, que comprende un segmento de secuencia de ácido nucleico U de secuencia de nucleótidos desconocida en un ácido nucleico aceptor, que comprende las etapas siguientes:

unir una secuencia adaptadora a una secuencia de ácido nucleico que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U en el primer extremo de la secuencia de ácido nucleico, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3;
 amplificar por PCR un ADN que comprende, en el siguiente orden, una secuencia adaptadora, el segmento de secuencia de ácido nucleico U, segmento de secuencia de ácido nucleico K2 y segmento de secuencia de ácido nucleico K3 usando un cebador directo que define un primer extremo del ADN amplificado y un cebador inverso que define un segundo extremo del ADN amplificado, terminando dicho cebador inverso en su extremo 3' en una secuencia de nucleótidos de segmento de secuencia K3;
 tratar las moléculas de ADN bicatenario lineales contenidas en el producto de PCR obtenido en la etapa anterior con una exonucleasa para obtener un saliente monocatenario que comprende la secuencia adaptadora en el primer extremo del ADN y un saliente monocatenario que comprende los segmentos de ácido nucleico K2 y K3 en el segundo extremo del ADN;
 hibridar el producto de la etapa anterior con un ácido nucleico aceptor bicatenario linealizado que tiene en un primer extremo del mismo un saliente monocatenario complementario al saliente monocatenario del primer extremo del ADN y en un segundo extremo del mismo un saliente monocatenario complementario al segmento de secuencia monocatenaria K2 del segundo extremo del ADN; y
 usar el producto de reacción obtenido en la etapa anterior para transformar una célula hospedadora.

(15) Un proceso de inserción de un ADNc en una secuencia de ARN de interés en un ácido nucleico aceptor, que comprende las etapas siguientes:

aislar ARN a partir de una célula;
 retrotranscribir ARN para formar una cadena no codificante de ADNc;
 amplificar por PCR dicho ADNc que comprende, en el siguiente orden, un segmento de secuencia U, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3 usando un cebador directo que define un primer extremo del ADN amplificado y un cebador inverso que define un segundo extremo del ADN amplificado, terminando dicho cebador inverso en su extremo 3' en una secuencia de nucleótidos de segmento de secuencia K3;
 tratar las moléculas de ADN bicatenario lineales contenidas en el producto de PCR obtenido en la etapa anterior con una exonucleasa para obtener un saliente monocatenario en el primer extremo del ADN y un saliente monocatenario que comprende segmentos de ácido nucleico K2 y K3 en el segundo extremo del ADN;
 hibridar el producto de la etapa anterior con un ácido nucleico aceptor bicatenario linealizado que tiene en un primer extremo del mismo un saliente monocatenario complementario al saliente monocatenario del primer extremo del ADN y en un segundo extremo del mismo un saliente monocatenario complementario al segmento de secuencia monocatenaria K2 del segundo extremo del ADN; y
 usar el producto de reacción obtenido en la etapa anterior para transformar una célula hospedadora.

(16) El proceso de acuerdo con uno cualquiera de (1) a (14), en el que la secuencia de interés comprende en dirección 5' a 3' un segmento de secuencia U, un segmento de secuencia K2 de secuencia de nucleótidos conocida y un segmento de secuencia K3 de secuencia de nucleótidos conocida.

(17) El proceso de acuerdo con (14), que comprende proporcionar un molde para dicha PCR uniendo una secuencia de unión a cebador al extremo 3' de dicho ADNc, en el que el cebador directo usado en dicha PCR

hibrida con dicha secuencia de unión a cebador.

(18) El proceso de acuerdo con (14), en el que el ARN es ARNm.

5 (19) El proceso de acuerdo con (14), en el que el ARN codifica una cadena de inmunoglobulina, la secuencia U es parte de una región variable de la inmunoglobulina y las secuencias conocidas K2 y K3 son parte de una región constante de la inmunoglobulina.

10 En la presente invención, se clonaron productos de PCR basándose en la homología entre las secuencias presentes tanto en el ácido nucleico aceptor (que puede ser un vector aceptor o de clonación) como en el ADN amplificado (inserto). Los salientes monocatenarios presentes en el ADN amplificado y en el ácido nucleico aceptor son complementarios y pueden hibridar. Los procesos de la invención permiten la clonación de un producto de PCR que contiene una región conocida (región K) y una región desconocida (U) muy eficazmente, religándose una cantidad muy pequeña de vector vacío; lo que es más importante, productos no deseados tales como dímeros de cebadores no se clonan en ninguna medida significativa. En procesos de inserción de ácidos nucleicos de la técnica anterior que comprenden un segmento de secuencia de nucleótidos desconocida y conocida, la especificidad para la clonación del producto deseado se determina exclusivamente mediante el uso de un cebador de PCR en el lado del ácido nucleico que tiene la secuencia de nucleótidos conocida. Cualquier unión inespecífica del cebador puede, en la técnica anterior, amplificar secuencias de ADN no deseadas que se insertarán en el ácido nucleico aceptor (o vector de clonación), ya que la hibridación con el ácido nucleico aceptor está determinada por la secuencia de nucleótidos introducida en el producto de PCR por el cebador. Por el contrario, se toman dos medidas para conseguir especificidad en la clonación de un ácido nucleico deseado en la presente invención. En primer lugar, la unión de cebador a un segmento de secuencia conocida del molde usado en la etapa de PCR se usa en la etapa de amplificación antes de la hibridación. En segundo lugar, un segmento de secuencia de la secuencia conocida (K2) diferente del segmento de secuencia usado para la unión de cebador se usa para hibridar el producto de PCR con el vector aceptor. Puesto que los ácidos nucleicos producto de PCR inespecíficos obtenidos por unión de cebadores inespecíficos generalmente no contendrán el segmento de secuencia K2 usado para la hibridación con el ácido nucleico aceptor (o vector aceptor), dichos ácidos nucleicos producto de PCR inespecíficos no se insertan en el vector aceptor en una medida significativa. Además, no se insertan dímeros de cebadores en el vector aceptor (Figura 1b). Los procesos de la invención tienen una eficacia extraordinariamente elevada para clonar ácidos nucleicos de interés. Estas dos medidas se consiguen en la invención mediante una elección adecuada de la región de unión a cebador usada en la etapa de PCR y una elección adecuada de la región de secuencia del producto de PCR usado para la clonación por homología con el vector aceptor. Curiosamente, la especificidad de esta estrategia de clonación permite la "limpieza" del producto de PCR, incluso cuando se usa el producto de la primera etapa de amplificación por PCR usando ADN genómico como molde o ARN total aislado de células. Aunque el producto de la primera amplificación por PCR puede ser una mezcla de productos de amplificación específicos e inespecíficos, se insertan esencialmente exclusivamente productos específicos en el vector aceptor. El hecho de que se clonen productos específicos esencialmente exclusivamente permite secuenciar plásmidos de colonias crecidas a partir de células hospedadoras transformadas sin una etapa de preexploración y, por lo tanto, se consigue una determinación conveniente de la secuencia de nucleótidos de segmento U.

Breve descripción de las figuras

45 Figura 1: Principio de la estrategia de clonación limpia. Un inserto que se va a clonar se amplifica por PCR con un cebador con homología con la diana (región K3) y un cebador homólogo a una secuencia adaptadora (A) añadida en el extremo de la región flanqueante de la secuencia desconocida (U).

(Figura 1a). El fragmento se clona por homología en un vector que contiene secuencias (A) y (K2). Puesto que la secuencia de K2 no está presente en los cebadores usados para la amplificación, no pueden clonarse productos inespecíficos y dímeros de cebadores (Figura 1b). A' y K3' son secuencias derivadas de A y K3, respectivamente.

50 Figura 2: Clonación de secuencias flanqueantes a partir de un molde de ARN: preparación de inserto para clonación limpia. gsp: cebador específico de gen; BAPpC: cebador usado para PCR que contiene una secuencia adaptadora denominada "BAP" y una cola 3' de nucleótidos de citosina. La secuencia oligo-G mostrada es la SEC ID N°: 1. BAP-CCCCCCCC es la SEC ID N°: 9. La región denominada "variable" es un segmento de secuencia cuya secuencia de nucleótidos puede ser desconocida (también denominado en la presente memoria segmento de secuencia U), tal como una secuencia variable de un anticuerpo.

55 La región denominada "constante" es un segmento de secuencia K de secuencia de nucleótidos conocida, que comprende un segmento de secuencia K2. Sin embargo, pueden aparecer polimorfismos en la región constante en muestras de pacientes diferentes. La región de secuencia constante puede ser la secuencia constante de un anticuerpo. El segmento de secuencia K2 no se usa como región de unión a cebador para PCR, pero puede usarse para clonación por homología en un ácido nucleico aceptor. En el lado a la derecha del segmento de secuencia K2, se muestra el segmento de secuencia K3 que se usa para la unión de cebador y la amplificación por PCR. Si se realiza una segunda PCR anidada en la etapa 6, la primera PCR se considera una preamplificación. En general, sin embargo, no es necesaria una preamplificación para conseguir cantidades suficientes de producto específico en la estrategia mostrada en la Figura 2.

65 Figura 3: Preparación de vector para clonación limpia y clonación limpia de un inserto. El vector de clonación

mostrado aquí contiene un gen LacZ para selección por azul-blanco. También puede usarse un vector que no contiene LacZ. El vector se linealiza usando una enzima de restricción tal como PstI. Después de la hibridación, el vector se transforma en una célula hospedadora tal como *E. coli*. Después de la clonación, el inserto puede secuenciarse usando un cebador específico de vector seqpr.

Figura 4: Clonación de secuencias flanqueantes a partir de un molde de ADN: preparación de inserto para clonación limpia, estrategia 1. Se aísla ADN genómico que generalmente contiene una pluralidad de diferentes fragmentos de ADN genómico (se muestran dos fragmentos de ADN bicatenario diferentes en la parte superior). En la etapa siguiente, a los fragmentos de ADN genómico se les añade una cola de G usando una transferasa terminal en sitios en los que el ADN se rompe durante el procedimiento de extracción. Se realiza una (o dos) PCR (anidadas). Después, el producto de PCR amplificado se somete a tratamiento con exonucleasa para obtener salientes monocatenarios 5'. Generalmente se realizan en esta estrategia tanto la PCR1 como la PCR2, ya que la proporción de productos de PCR deseados respecto a productos de PCR inespecíficos puede no ser satisfactoria si sólo se realiza una PCR.

Figura 5: Clonación de secuencias flanqueantes a partir de un molde de ADN: preparación de inserto para clonación limpia, estrategia 2. En lugar de depender de extremos de ADN rotos en el ADN genómico como en la estrategia 1 (Figura 4), se realiza una amplificación lineal por PCR usando solamente un cebador específico de gen *gsp*. Después, a este producto se le añade una cola de G usando una transferasa terminal. Después, se realiza una PCR usando cebadores BAPpC y un segundo cebador específico de gen *Gsp2*, seguido de tratamiento con exonucleasa.

Figura 6: Esquema que muestra la construcción del vector de clonación limpia pICH36166. Se realiza una PCR usando los cebadores igclon1 y nospclon1 usando pICH30086. Los segmentos C y D de los cebadores también están presentes en el vector pICH36833. *bap2* es una secuencia adaptadora que se insertará en el vector aceptor pICH36166 que se producirá. CS representa la secuencia de captura, que es un ejemplo para un segmento de secuencia K2. La secuencia de nucleótidos de CS se muestra y se denomina en la presente memoria SEC ID N°: 2. El producto de PCR puede insertarse en el vector pICH36833 mediante clonación por homología, que implica el tratamiento tanto del producto de PCR como del pICH36833 con la actividad exonucleasa de la polimerasa T4.

Figura 7: Estructura del ARN de anticuerpo IGM con secuencias constantes y variables, productos de PCR con cola de G y vectores de clonación limpia usados en los ejemplos.

Figura 8: Posición de secuencias de captura (segmentos de secuencia K2) y cebadores de amplificación Mu2N y Mu3N usados en secuencias constantes de IgM (cadena Mu de inmunoglobulina) humana. La secuencia de nucleótidos mostrada se denomina en la presente memoria SEC ID N°: 3.

Figura 9: Posición de secuencias de captura y cebadores de amplificación usados en secuencias constantes de IgK (cadena Kappa de inmunoglobulina) humana. La secuencia de nucleótidos mostrada se denomina en la presente memoria SEC ID N°: 4.

Figura 10: Ensayo de actividad exonucleasa T4 a diferentes temperaturas y concentraciones de nucleótidos. L: Escalera de ADN Plus de 1 kb GeneRuler (Fermentas). Condiciones de reacción:

ADN: pICH10990 digerido con SacI/NdeI (10 µl/50 µl), (Tamaños de fragmento: 3595, 1563, 1158, 125 pb) 2,5 µl/10 µl usados para tratamiento con polimerasa T4 (Pol. T4).

Polimerasa T4: 0,5 µl/10 µl = 1.5 U/10 µl en 2 NEB + BSA

Incubación: 10 min a 25°C, 20°C, 15°C o 10°C, todo inactivado térmicamente: 5 min a 80°C; C = control sin T4; H = control de inactivación térmica (después de la adición de polimerasa T4 5 min a 80°C). Tratamiento con nucleasa de judía Mung: adición de 1,5 µl de tampón de nucleasa de judía Mung, 1 µl de nucleasa de judía Mung (10 U/µl), a 15 µl de H₂O, 20 min 25°C.

Figura 11: Clonación limpia con T4: ensayo de diferentes tiempos de digestión con polimerasa T4 a temperatura ambiente.

Figura 12: Clonación de secuencia flanqueante en ADN de una planta de *Nicotiana benthamiana* transgénica para la construcción pICH30649. El ADN-T contiene en el borde derecho el terminador Nos y una secuencia de vector viral (3' NTR). Se define una "secuencia de captura" (CS, equivalente a la región K2) en la 3'NTR y se clona en un vector de clonación pICH36166. *bap2*: secuencia adaptadora. "Desconocida" indica un segmento de secuencia de secuencia de nucleótidos desconocida U. La secuencia de captura CS se denomina en la presente memoria SEC ID N°: 5. 3'NTR: región no traducida 3'. Se indican las regiones K1, K2 y K3. K1 es un segmento de secuencia de secuencia de nucleótidos conocida entre el segmento de secuencia desconocida U y el segmento de secuencia K2.

Figura 13: PCR de secuencias flanqueantes (F-PCR) en plantas de *Nicotiana benthamiana* transgénicas para la construcción pICH30649.

Figura 14: PCR de colonias sobre productos de F-PCR clonados (transgén pICH30649).

Figura 15: F-PCR con y sin preamplificación. La secuencia de captura CS se denomina en la presente memoria SEC ID N°: 2.

Figura 16: F-PCR y clonación de secuencias flanqueantes de pICH18693 a partir de la planta transgénica nbi160-7.

Figura 17: Ensayo de actividad exonucleasa de polimerasa Klenow con diferentes temperaturas y tiempos de incubación, y ensayo de clonación. c = inserto, PCR en pICH3491x (biblioteca de Arabidopsis) ecolib3+4.

Tratamiento con Klenow:

- 2,5 μl /10 μl de pICH10990 digerido con SacII/NdeI (10 μl /50 μl), (Tamaños de fragmento: 3595, 1563, 1158, 125 pb)
- 1 μl /10 μl = 10 U/10 μl de Klenow en 2 NEB + BSA
- 30, 60, 90, 120 min a 37°C o temperatura ambiente (25°C), todo inactivado térmicamente: 5 min a 80°C; 0 min = control sin T4;
- **Tratamiento con nucleasa de judía Mung:**
- + 1,5 μl de tampón de nucleasa de judía Mung, 1 μl de nucleasa de judía Mung (10 U/ μl), a 15 μl de H₂O, 20 min a 25°C

Clonación con Klenow:

- 1 μl /10 μl de vector = pICH31401 PstI, 0,5 μl de inserto = PCR sobre pICH3491x (biblioteca de Arabidopsis) ecolib3 + 4,
- 0,5 μl = 2,5 U/10 μl de Klenow,
- 30, 60, 120 min a 37°C, 5 min a 80°C
- Los 10 μl se usan para transformar células DH10B químicamente competentes, se añaden 500 μl de LB, se siembran 25 μl en placas de LB con Carb y X-Gal después de una agitación de 15 min a 37°C.

Descripción detallada de la invención

La secuencia de ácido nucleico de interés que se va a clonar o insertar en un ácido nucleico aceptor puede ser ADN o ARN. Si es ARN, se clona después de su transcripción inversa en ADNc. El ácido nucleico aceptor es generalmente ADN bicatenario. Por lo tanto, la clonación de una secuencia de ARN de interés significa clonar un ADNc de la secuencia de ARN de interés. La clonación de una secuencia de ácido nucleico de interés en la presente memoria significa insertar la secuencia de ácido nucleico de interés en un ácido nucleico aceptor. La secuencia de ácido nucleico de interés puede insertarse en el ácido nucleico aceptor como parte de una molécula de ácido nucleico de mayor tamaño que contiene, además de la secuencia de ácido nucleico de interés, una secuencia adaptadora y un segmento K2 que se usan para la clonación por homología. La clonación por homología en la presente memoria significa insertar una secuencia de ácido nucleico de interés en un ácido nucleico aceptor, por lo que tanto la secuencia de ácido nucleico de interés como el ácido nucleico aceptor se proporcionan con salientes monocatenarios, y por lo que los salientes monocatenarios en ambos lados de la secuencia de ácido nucleico de interés hibridan con salientes monocatenarios del ácido nucleico aceptor linealizado por emparejamiento de bases. La clonación puede comprender además usar el ácido nucleico aceptor que tiene integrado el ácido nucleico de interés para transformar una célula hospedadora adecuada. La clonación puede comprender además cultivar clones separados a partir de las células hospedadoras transformadas.

El ácido nucleico aceptor es generalmente un vector de clonación. Las expresiones “vector de clonación” y “ácido nucleico aceptor” significan una molécula de ADN bicatenario que puede mantenerse en células hospedadoras adecuadas. El ácido nucleico aceptor puede tener un origen de replicación para permitir la replicación en las células hospedadoras empleadas. Además, o adicionalmente, el ácido nucleico aceptor (o vector de clonación) puede tener un marcador de selección que permita la selección de las células hospedadoras adecuadas que contienen el ácido nucleico aceptor.

Los procesos de la invención tienen su aplicación principal en la clonación, determinación e identificación de secuencias de ácido nucleico desconocidas que flanquean a una secuencia de ácido nucleico de secuencia de nucleótidos conocida. Sin embargo, el concepto de la presente invención no requiere que una secuencia de ácido nucleico de interés que se va a clonar tenga un segmento de secuencia de ácido nucleico de secuencia de nucleótidos desconocida en un extremo. Si se conoce la secuencia de nucleótidos en ambos extremos del ácido nucleico de interés que se va a clonar, las etapas de unir una secuencia de unión a cebador o una secuencia adaptadora pueden omitirse, y tanto el cebador directo como el inverso usados en la etapa de PCR pueden diseñarse para unirse específicamente con fines de PCR a ambos extremos del ácido nucleico de interés presente en el molde de PCR. En cualquier caso, después de haber unido una secuencia de unión a cebador o una secuencia adaptadora a un segmento de secuencia de ácido nucleico U de secuencia desconocida presente en una secuencia de ácido nucleico de interés, la secuencia de ácido nucleico obtenida contendrá en ambos extremos segmentos de secuencia de secuencia de nucleótidos conocida que pueden usarse para la unión de cebadores en la etapa de PCR de la invención.

Generalmente, la secuencia de ácido nucleico de interés comprende un segmento de secuencia U que puede ser de secuencia de nucleótidos desconocida. La secuencia de ácido nucleico de interés puede comprender además un segmento de secuencia de ácido nucleico (en resumen “segmento”) de secuencia de nucleótidos conocida tal como un segmento K2 o parte del mismo. El segmento K2 se usa para insertar un ácido nucleico de interés en el ácido nucleico aceptor por emparejamiento de bases entre el segmento monocatenario K2 y un segmento monocatenario complementario del vector aceptor denominado K2'. El segmento K2 también se denomina en la presente memoria

“secuencia de captura”, ya que se captura por el segmento K2' del ácido nucleico aceptor. El segmento K3 tiene la secuencia de nucleótidos del cebador inverso en la etapa de amplificación por PCR de la invención y define el segmento K3 en el ADN amplificado. Un segmento de secuencia en el molde de PCR usado en la etapa de PCR que es homólogo al cebador inverso de modo que se consigue la hibridación de cadenas complementarias en dicha PCR también puede denominarse segmento K3. El segmento K3 no se inserta en el ácido nucleico aceptor y por lo tanto no es parte de la secuencia de ácido nucleico de interés. En la presente memoria, el extremo de una secuencia de ácido nucleico que contiene un segmento U más próximo al segmento de secuencia U se denomina en la presente memoria “primer extremo” del ácido nucleico respectivo. El otro extremo de un ácido nucleico que está más próximo al segmento de secuencia de ácido nucleico K2 se denomina “segundo extremo” del ácido nucleico respectivo.

El proceso de la invención comprende una etapa de amplificar por PCR un ADN que comprende, en el siguiente orden, un segmento de secuencia U, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3 usando un cebador directo que define un primer extremo del ADN y un cebador inverso que define un segundo extremo del ADN. El ADN mencionado en esta etapa de proceso es el ADN amplificado en esta etapa. Por lo tanto, este ADN también se denomina en la presente memoria “ADN amplificado”. El cebador inverso termina en su extremo 3' en una secuencia de nucleótidos de segmento de secuencia de ácido nucleico K3. El ADN termina en su primer extremo mediante la secuencia de nucleótidos determinada por el cebador directo, y termina en su segundo extremo mediante la secuencia de nucleótidos determinada por el cebador inverso. Aparte de los segmentos U, K2 y K3, el ADN puede tener además otras porciones de secuencia presentes en el molde usado para PCR, tales como un segmento de secuencia de ácido nucleico adicional K1 entre los segmentos U y K2. El segmento K1 puede ser de secuencia de nucleótidos conocida. Los segmentos K2 y K3 no tienen que ser contiguos. En su lugar, los segmentos K2 y K3 pueden estar separados por otro segmento de secuencia. En general, sin embargo, los segmentos K2 y K3 no están separados por más de 100 nt (“nt” representa nucleótidos), preferentemente son contiguos. Si la secuencia de nucleótidos de segmento U no se conoce, el ADN amplificado puede contener además en su primer extremo una secuencia adaptadora usada para hibridar el primer extremo del ADN con el primer extremo del ácido nucleico aceptor en la etapa de hibridación.

El segmento K2 se define como un segmento de ácido nucleico homólogo al segmento de secuencia K2' presente en el segundo extremo del ácido nucleico aceptor linealizado para permitir la hibridación del segundo extremo monocatenario del ADN amplificado con el segundo extremo monocatenario del ácido nucleico aceptor por emparejamiento de bases entre cadenas complementarias. La longitud del segmento de K2 y el grado de homología con el segmento K2' deberían ser tales que pueda conseguirse una hibridación estable en la etapa de hibridación. Con este fin, la longitud del segmento K2 es de al menos 10 y debería ser de al menos 12 nucleótidos, mejor de al menos 16 nucleótidos, aún mejor de al menos 20 nucleótidos y, preferentemente, de al menos 25 nucleótidos. No existe un límite superior estricto para la longitud del segmento K2. Por comodidad, será generalmente de casi 100 nt. Cuando se compara el grado de homología entre los segmentos K2 y K2', se comparan segmentos de secuencia (o “bloques”) de una longitud idéntica en términos del número de nucleótidos. El grado de homología entre el segmento K2 y el segmento K2' debería ser tan alto como fuera posible para asegurar una hibridación específica. Por ejemplo, los segmentos K2 y K2' deberían contener un bloque de al menos 10 nucleótidos contiguos de secuencia de nucleótidos idéntica. Sin embargo, el bloque de secuencias de nucleótidos idénticas en los segmentos K2 y K2' puede ser más largo de 10 nt, tal como de al menos 12, al menos 15 o al menos 20 nucleótidos contiguos. Cuanto mayor sea la homología entre los segmentos K2 y K2', mejor pueden tolerarse polimorfismos desconocidos que puedan aparecer en los segmentos K2 derivados de muestras de diferentes individuos o pacientes sin impedir una clonación con éxito. La relación entre la tendencia a la hibridación de cadenas de ADN complementarias y la similitud de las cadenas pertenece al conocimiento general del experto en la materia, de modo que siempre es fácil para el experto en la materia diseñar un segmento K2' adecuado dependiente de un segmento K2 dado. Si, en la presente memoria, se hace referencia a la secuencia de nucleótidos del segmento K2 como una secuencia de nucleótidos conocida, esto no significa que la secuencia de nucleótidos deba conocerse exactamente. En su lugar, es suficiente que el segmento de secuencia K2 hibride en la etapa de hibridación con el segmento de secuencia K2' que se obtiene por ingeniería genética de modo que las cadenas complementarias de los segmentos K2 y K2' puedan hibridar en la etapa de hibridación de la invención.

El segmento K3 se define como el segmento de secuencia de ácido nucleico introducido en el ADN amplificado en la etapa de PCR mediante el cebador inverso (y la secuencia complementaria del cebador inverso producida en dicha PCR). Respecto a la longitud y a la secuencia de nucleótidos que se empleará para el cebador inverso, son aplicables consideraciones como las que se aplican generalmente en la PCR por el experto en la materia. Generalmente, el cebador inverso tiene una longitud de entre 10 y 50 nt o de entre 14 y 30 nt. La secuencia de nucleótidos del cebador inverso y, por lo tanto, del segmento K3, se seleccionan de modo que se consiga una unión adecuada al molde usado en la PCR.

En la presente memoria, los segmentos K2 y K3 son segmentos separados que no solapan en un grado significativo. Preferentemente, no solapan en más de 3 nt o más de 1 nt. Esto significa que se usan segmentos separados para la amplificación por PCR por un lado, y para la inserción en el ácido nucleico aceptor por el otro lado en el segundo extremo de tanto el ADN como el ácido nucleico aceptor. La ausencia de solapamiento entre segmentos K2 y K3 puede conseguirse diseñando el cebador inverso de modo que termine en su extremo 3' en una secuencia de

nucleótidos de segmento de secuencia de ácido nucleico K3. En otras palabras, el cebador inverso no debería prolongarse con su extremo 3' hacia el segmento K2, que tiene homología con un segmento de secuencia de nucleótidos K2' presente en el segundo extremo del ácido nucleico aceptor.

5 Para evitar que el ADN amplificado hibride con el segmento monocatenario K3 con el segundo extremo monocatenario del ácido nucleico aceptor en la etapa de hibridación, los segmentos K2 y K3 deben ser suficientemente diferentes, es decir, carecer de una homología significativa. El experto en la materia puede juzgar a partir de conocimientos generales cuándo dos secuencias de nucleótidos son suficientemente diferentes para evitar la hibridación de cadenas complementarias. Una ausencia de homología significativa puede significar que K2 y K3
10 no contienen un bloque de 9 o más nucleótidos contiguos de secuencias de nucleótidos idénticas y, preferentemente, no contienen un bloque de 5 o más nucleótidos contiguos de secuencias de nucleótidos idénticas. Son aplicables consideraciones similares a la secuencia de nucleótidos del adaptador A y del segmento K2.

15 Después de dicha etapa de amplificación por PCR, las moléculas de ADN contenidas en el producto de PCR se tratan con una exonucleasa para obtener un saliente monocatenario en el primer extremo del ADN y un saliente monocatenario que comprende los segmentos de ácido nucleico K2 y K3 en el segundo extremo del ADN. Si se desea, las moléculas de ADN del producto de PCR pueden separarse de otros componentes de la PCR por métodos conocidos en general antes del tratamiento con exonucleasa. La exonucleasa puede ser una exonucleasa 5'-3' o una exonucleasa 3'-5'. Se conocen en la técnica enzimas que tienen actividad exonucleasa. Son ejemplos las usadas en el documento US 2007/0292954. Son ejemplos la ADN polimerasa T4 de *E. coli*, el fragmento grande de la polimerasa de *E. coli* conocido como polimerasa Klenow, la nucleasa lambda, la nucleasa T7 o la exonucleasa III, preferentemente es la polimerasa Klenow de *E. coli*. No es necesario añadir dNTP a la mezcla de reacción, lo que proporciona más flexibilidad al proceso de la invención. Puesto que la actividad exonucleasa de diferentes enzimas que tienen actividad exonucleasa difiere, se recomienda establecer condiciones adecuadas para el tratamiento con exonucleasa tales como temperatura, concentración de exonucleasa y duración del tratamiento, de modo que se elimine un número adecuado de nucleótidos del ADN y se obtengan salientes de una longitud adecuada en ambos extremos. Una longitud adecuada en el primer extremo y el segundo extremo del ADN amplificado es una longitud que expone completamente las porciones de secuencia necesarias para la hibridación con el ácido nucleico aceptor. En el segundo extremo, todo el segmento K3 y esencialmente todo el segmento K2 debería hacerse monocatenario.
20 Las condiciones adecuadas para la ADN polimerasa T4 se proporcionan en los ejemplos. También se describe en los ejemplos un método para identificar condiciones adecuadas para el tratamiento con exonucleasa (Ejemplo 3). Pueden aplicarse métodos análogos para identificar condiciones adecuadas para otras enzimas que tengan actividad exonucleasa. Aunque es posible realizar la etapa de tratamiento con exonucleasa con una mezcla de dos o más exonucleasas, no debería combinarse una exonucleasa 5'-3' y una exonucleasa 3'-5' puesto que los salientes deseados pueden no obtenerse en este caso. Después del tiempo de reacción predeterminado, la actividad exonucleasa puede destruirse por métodos conocidos tales como por calentamiento.

El ácido nucleico aceptor usado en la presente invención es un ADN bicatenario linealizado tal como un vector de clonación. El ácido nucleico aceptor puede crearse linealizando un ácido nucleico aceptor bicatenario que comprende, en el siguiente orden, un segmento de secuencia homólogo al primer extremo del ADN amplificado, un segmento de secuencia de cadena principal de vector que comprende un marcador de selección y un segmento de secuencia K2', que es homólogo al segmento de secuencia K2. En una realización, el ácido nucleico aceptor puede crearse linealizando una molécula circular que comprende, en el siguiente orden,

45 un primer sitio de escisión por endonucleasa de restricción, el segmento de secuencia homólogo al segmento de secuencia adaptadora en el primer extremo del ADN amplificado, el segmento de secuencia de cadena principal de vector que comprende un marcador de selección, el segmento de secuencia K2' homólogo al segmento de secuencia K2 y, opcionalmente,
50 un segundo sitio de escisión por endonucleasa de restricción y un segmento de secuencia espaciadora que puede eliminarse del ácido nucleico aceptor por restricción en dicho primer y en dicho segundo sitio de escisión por endonucleasa de restricción. El segmento de secuencia espaciador puede contener lacZ para la selección por azul/blanco de clones obtenidos después de transformar el producto de reacción de la etapa de hibridación en células hospedadoras, consúltese la Figura
55 3.

La linealización puede conseguirse por tratamiento con una endonucleasa de restricción que reconozca el primer y/o el segundo sitio de escisión por endonucleasa de restricción. Es posible y se prefiere que la linealización se realice en la misma mezcla de reacción en la que se realiza el tratamiento con exonucleasa del ADN amplificado y del ácido nucleico aceptor linealizado.
60

El ácido nucleico aceptor contiene por lo tanto en su primer extremo un segmento de secuencia homólogo a un segmento de secuencia en el primer extremo del ADN amplificado. Si el segmento U es de secuencia de nucleótidos desconocida, el primer extremo del ADN amplificado estará formado por una secuencia adaptadora A añadida al primer extremo del ADN amplificado en dicha PCR o en una etapa anterior (véase adicionalmente a continuación). En su segundo extremo, el ácido nucleico aceptor contiene un segmento K2', que es homólogo al segmento K2 del
65

ADN amplificado, de modo que pueden hibridar las cadenas complementarias del segmento K2' y del segmento K2. El término "complementaria", por lo tanto, no significa que se formen exclusivamente pares de bases G-C y A-T tras la hibridación. En su lugar, se tolera algún emparejamiento erróneo con tal de que pueda conseguirse una hibridación estable (véase anteriormente respecto al grado de homología entre los segmentos K2 y K2'). Los segmentos K2 y K2' son preferentemente de una longitud idéntica. Es posible que un segmento de secuencia de cualquier secuencia de nucleótidos esté presente entre el segundo extremo del ácido nucleico aceptor y el segmento K2'. Si está presente, dicho segmento de secuencia es más corto que 50 nt, preferentemente más corto que 20 nt. En una realización, el segundo extremo coincide con el segmento K2'. El ácido nucleico aceptor no contiene un bloque de secuencia de más de 9, preferentemente de más de 5 nucleótidos contiguos, que también está presente con una secuencia de nucleótidos idéntica en el segmento K3 dentro de un intervalo que se ha vuelto monocatenario en la etapa de tratamiento con exonucleasa del ácido nucleico aceptor, tal como dentro de una región de 100 nt del segundo extremo del ácido nucleico aceptor.

Antes de la etapa de hibridación, el ácido nucleico aceptor bicatenario linealizado se proporciona en ambos extremos con los salientes monocatenarios necesarios para la etapa de hibridación. Los salientes monocatenarios pueden obtenerse por tratamiento con exonucleasa de forma similar a la descrita anteriormente para el ADN amplificado. Si se usa una exonucleasa 5'-3' para el ADN amplificado, también se usa una exonucleasa 5'-3' para el ácido nucleico aceptor. Si se usa una exonucleasa 3'-5' para el ADN amplificado, también se usa una exonucleasa 3'-5' para el ácido nucleico aceptor. En una realización, el tratamiento con exonucleasa del ADN amplificado y del ácido nucleico aceptor se realizan en el mismo recipiente de reacción, es decir, como una reacción de un recipiente.

La etapa de hibridación de la invención se realiza incubando el ADN amplificado tratado con exonucleasa con el ácido nucleico aceptor bicatenario linealizado que tiene en un primer extremo del mismo un saliente monocatenario complementario al saliente monocatenario del primer extremo del ADN, y en un segundo extremo del mismo un saliente monocatenario complementario al segmento de secuencia monocatenario K2 del segundo extremo del ADN. La incubación se lleva a cabo en condiciones que permitan dicha hibridación.

Después de la hibridación, la mezcla de reacción de la etapa de hibridación se usa para transformar células hospedadoras. Las células hospedadoras deberían adaptarse al ácido nucleico aceptor de modo que la replicación del ácido nucleico aceptor sea posible en la célula hospedadora. Por lo tanto, el ácido nucleico aceptor debería tener un origen de replicación que pueda reconocerse por la ADN polimerasa presente en la célula hospedadora. Es ventajoso que el ácido nucleico aceptor contenga en su cadena principal de vector un marcador de selección que permita la selección de células hospedadoras que contienen el ácido nucleico aceptor. La célula hospedadora rellenará los huecos en el ácido nucleico aceptor hibridado y escindirá los salientes del ADN o del ácido nucleico aceptor que no estén implicados en la hibridación. Pueden obtenerse clones de células hospedadoras, y el ADN de vector reparado puede aislarse de clones de células hospedadoras, por ejemplo, por secuenciación del ácido nucleico de interés insertado en el ácido nucleico aceptor.

La secuencia adaptadora así como el segmento K2' determinan la idoneidad de un ácido nucleico aceptor dado para insertar un ADN amplificado dado. Puede que tengan que generarse diferentes ácidos nucleicos aceptores para ADN amplificados que difieran significativamente en la secuencia de nucleótidos en el primer extremo o en el segmento K2. Sin embargo, la construcción del ácido nucleico aceptor es sencilla. La Figura 6 muestra una estrategia para ello, en la que la secuencia adaptadora (bap2 en este ejemplo) y el segmento K2' (CS en este ejemplo) se proporcionan al ácido nucleico aceptor mediante cebadores (igclon1 y nosplon2 en la Figura 6). La clonación por homología puede usarse para insertar el producto de PCR en una cadena principal de vector (pICH36833 en la Figura 6). El ácido nucleico aceptor (pICH36166 en la Figura 6) puede linealizarse con la endonucleasa de restricción PstI. Es conveniente usar la misma secuencia adaptadora en el primer extremo del ADN amplificado y el ácido nucleico aceptor para clonar proyectos de secuencias de interés que tengan diferentes segmentos K2.

En lo siguiente, se describe el molde usado en la etapa de PCR así como las formas de obtenerlo. El ácido nucleico de molde comprende el ácido nucleico de interés y, por lo tanto, comprende, en este orden, al menos segmento U, segmento K2 y un segmento homólogo al segmento K3. Como anteriormente, la secuencia de nucleótidos del segmento K2 es esencialmente conocida. La del segmento homólogo a K3 también es esencialmente conocida, de modo que el cebador inverso puede diseñarse para hibridar con el segmento homólogo al segmento K3 del molde. Si se conoce la secuencia de nucleótidos en el otro extremo del ácido nucleico de interés en el molde, el cebador directo puede diseñarse para hibridar con fines de PCR con el primer extremo del molde.

En una realización importante, la secuencia de interés a insertar en un ácido nucleico aceptor tiene en su primer extremo un segmento de secuencia U cuya secuencia de nucleótidos se desconoce. En este caso, no es posible diseñar un cebador directo específico y preparar un ácido nucleico aceptor que tenga en su primer extremo un segmento de secuencia homólogo al primer extremo del ácido nucleico de interés. Se han descrito diversas soluciones a este problema en la técnica anterior que pueden usarse para la presente invención. El material de partida para esta realización puede ser ADN genómico aislado a partir de células eucariotas o procariotas. En una realización alternativa, el material de partida puede ser ARN aislado de células eucariotas o procariotas. El material de partida también puede ser, o contener en el ARN aislado total, ARNm de células eucariotas o procariotas.

En una realización en la que la secuencia de ácido nucleico es ADN (por ejemplo, en una preparación de ADN genómico), una secuencia de unión a cebador se une a una secuencia de ácido nucleico que tiene en un extremo un segmento de secuencia de ácido nucleico U de secuencia de nucleótidos desconocida (véase la Figura 4 para una ilustración). La secuencia de unión a cebador puede ser un segmento de secuencia de nucleótidos homooligomérica que se une al segmento U usando una desoxirribonucleótido transferasa terminal y NTP del tipo de la cola homooligomérica deseada que se va a unir. En una realización preferida, el segmento de secuencia de nucleótidos homooligomérica (o "cola") consiste en bases de guanina, y la unión que usa una desoxirribonucleótido transferasa terminal se conoce como "cola de G". En lugar de GTP, pueden usarse otros NTP tales como ATP, CTP o TTP. Una vez que se une una secuencia de unión a cebador, la secuencia de nucleótidos del primer extremo del ácido nucleico de interés obtenido se conoce, y un cebador directo adecuado de secuencia complementaria puede diseñarse para la PCR. La secuencia de unión a cebador debería tener una longitud de al menos 10 nt. Sin embargo, se prefiere que la secuencia usada para la clonación por homología en el primer extremo del ADN en el ácido nucleico aceptor sea más larga que esta longitud. Por lo tanto, el cebador directo usado puede comprender en su extremo 5' un segmento de secuencia adaptadora además de un segmento de secuencia complementario a la secuencia de unión a cebador. El ADN amplificado contendrá entonces la secuencia adaptadora en su primer extremo, que se volverá monocatenario en la etapa de tratamiento con exonucleasa. El ácido nucleico aceptor se proporcionará con un segmento de secuencia homólogo a la secuencia adaptadora, de modo que el tratamiento con exonucleasa del ácido nucleico aceptor genera un saliente monocatenario en el primer extremo del ácido nucleico aceptor que es complementario al saliente monocatenario en el primer extremo del ADN amplificado.

En la realización del párrafo anterior, es ventajoso realizar una etapa de preamplificación por PCR (denominada "PCR1", etapa 3 en la Figura 4). La etapa de amplificación por PCR de la invención es entonces la segunda PCR (denominada "PCR2" en la Figura 4). En esta realización, el cebador inverso usado en la segunda PCR corresponde al segmento K3 del ADN amplificado. El cebador inverso usado en la primera PCR también se une a un segmento de secuencia de secuencia de nucleótidos conocida del molde de PCR de la primera PCR.

En otra realización en la que la secuencia de ácido nucleico es ADN (por ejemplo, en una preparación de ADN genómico), se realiza una síntesis de primera cadena sobre ADN, tal como ADN genómico, usando un cebador que se une a una secuencia conocida (véase la Figura 5 para una ilustración). Después de la eliminación de cebadores y dNTP, una secuencia de unión a cebador puede unirse al extremo 3' de la primera cadena sintetizada. Las etapas siguientes pueden ser las mismas que en la realización anterior ilustrada en la Figura 4. Sin embargo, generalmente no se requiere una preamplificación.

En una realización alternativa en la que la secuencia de ácido nucleico es ADN (por ejemplo, en una preparación de ADN genómico), una secuencia adaptadora de una longitud suficiente se une por ligación a una secuencia de ácido nucleico que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U en el primer extremo de la secuencia de ácido nucleico, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3 antes de la etapa de amplificación por PCR. No es necesario introducir una secuencia adaptadora en el primer extremo mediante el cebador directo. El cebador directo usado en la PCR está diseñado para unirse a la secuencia adaptadora unida. Preferentemente, en esta realización, el cebador directo termina en su extremo 5' en la secuencia adaptadora unida al ácido nucleico.

En una realización adicional, se emplea TAIL-PCR como se describe por Liu et al en *Biotechniques* 43 (2007) 649-656 en la etapa de amplificación por PCR.

En una realización importante, un ADNc de una secuencia de ARN de interés tal como una secuencia de ARNm de interés se inserta en un ácido nucleico aceptor, consúltese la Figura 2. En este caso, el material de partida para el proceso de la invención es ARN. Después de haber aislado ARN de células eucariotas o procariontas, se genera una cadena de ADNc del ARN por transcripción inversa usando un cebador específico de gen ("gsp1" en la Figura 2). Si el ARN es ARNm, puede usarse un cebador oligo-dT como alternativa para la transcripción inversa del ARNm. El experto en la materia entiende que se obtienen muchos más productos si se usa un oligo-dT como cebador para la síntesis de cadenas de ADNc en comparación con un caso en el que se usa un cebador específico de gen en una muestra de ARNm celulares. El uso de un cebador oligo-dT en lugar de un gsp1 para la síntesis de la primera cadena tiene la ventaja de que cualquier polimorfismo desconocido que pueda estar presente en el segmento K2 en el ARNm de diferentes individuos no conduce al fracaso de la síntesis de la primera cadena de dicho polimorfo. Como etapa siguiente, el molde de ARN puede digerirse después y el cebador usado para la retrotranscripción puede eliminarse. Las etapas restantes del proceso de la invención pueden ser las mismas que se han descrito anteriormente. La Figura 2 ilustra en la etapa 6 una primera PCR y una segunda PCR anidada opcional. Si se realizan dos reacciones de PCR en el molde de ADN, la segunda PCR se considera la etapa de amplificación de la invención y el cebador inverso de la segunda PCR define el segmento K3 del producto de PCR. Generalmente, la última amplificación por PCR se considera la etapa de amplificación de la invención. Si sólo se realiza una PCR sobre el molde de ADN, se considera que esta PCR es la etapa de amplificación de la invención.

Una estrategia alternativa para unir un adaptador a un segmento de secuencia U es ligar un cebador de ARN en el extremo 5' del ARNm usando la ARN ligasa T4 como se describe en Trout, A. B., McHeyzer-Williams, M. G.,

Pulendran, B. y Nossal G. J. V., 1992, PNAS, Vol 89: 9823-9825. Este ARN modificado se somete después a transcripción inversa y se usa para la amplificación por PCR como en otros protocolos descritos anteriormente.

Una estrategia alternativa usa la actividad de intercambio de molde de la transcriptasa inversa del virus de la leucemia murina de Moloney para añadir 2-4 citosinas extra al extremo 3' de la cadena de ADNc recién sintetizada al alcanzar la estructura de caperuza del ARNm, como se describe en el documento US 5.962.272. El ADNc resultante puede generarse para que contenga una secuencia adaptadora en el lado del segmento U. Esta estrategia no requiere la adición de cola de G con transferasa terminal y el producto puede usarse directamente para PCR como en otros protocolos.

Si el material de partida del proceso de la invención es ADN, tal como ADN genómico aislado de células eucariotas o procariontas, pueden usarse diversas estrategias para proporcionar el molde para la reacción de PCR de la invención. Se muestran esquemáticamente dos estrategias posibles en la Fig. 4 y la Fig. 5. En la estrategia representada en la Fig. 4 se usan fragmentos de ADN genómico directamente para unir un adaptador o una cola homopolimérica. Se ha descubierto sorprendentemente por los presentes inventores que no es necesario aplicar ninguna medida para promover la fragmentación del ADN, tal como tratamiento con ADNasa o aplicación de esfuerzo cortante. En su lugar, puede usarse el ADN genómico según se aísla en condiciones suaves mediante un kit comercial para el aislamiento de ADN genómico. Los fragmentos de ADN genómico serán sustancialmente más grandes en comparación con un caso en el que se induce la fragmentación del ADN. Puesto que los fragmentos más pequeños se amplifican preferente por PCR, el tener fragmentos más grandes de ADN genómico significa que se produce menos fondo de productos de PCR inespecíficos durante la amplificación por PCR que en los métodos de la técnica anterior. Aun así, es posible amplificar esencialmente todos los ácidos nucleicos de interés, puesto que para la mayoría de las secuencias de interés habrá un fragmento de ADN genómico en una preparación típica que se localice lo bastante próximo a un extremo de un fragmento para permitir una amplificación por PCR eficaz.

Para la identificación de secuencias desconocidas en el ADN, por ejemplo, inserciones de transgenes o transposones, o para desplazamiento sobre el cromosoma, se han descrito diversos protocolos en la bibliografía. Estas estrategias se incluyen en varias clases.

En la primera clase, el ADN se digiere con una enzima y se liga un adaptador en los extremos. Después, las secuencias flanqueantes se amplifican usando cebadores específicos de secuencia (gsp) conocidos y cebadores adaptadores. Habitualmente se realizan dos PCR anidadas usando dos cebadores específicos de secuencia conocidos anidados. Un inconveniente importante de los protocolos que dependen de la digestión de ADN con enzimas es que no se sabe si una enzima corta en las proximidades de la secuencia conocida y, por lo tanto, es necesario usar varias enzimas en paralelo para asegurarse de que una proporcione resultados (produzca un fragmento que no sea demasiado corto y que tenga suficiente información, o no demasiado largo puesto que no se amplificaría por PCR). Después, el producto de PCR puede secuenciarse directamente. Sin embargo, esto no producirá información si está presente más de una inserción y se han amplificado ambas secuencias flanqueantes. Esta mezcla de productos de PCR puede clonarse y varios plásmidos secuenciarse para obtener las dos secuencias flanqueantes.

Sin embargo, un problema principal cuando se analizan secuencias flanqueantes usando un adaptador ligado a extremos de ADN producidos por digestión con enzimas de restricción es que no se puede asegurar que se hayan determinado todas las secuencias flanqueantes de todas las inserciones en este genoma. Esto es porque los productos de PCR obtenidos para diversas secuencias flanqueantes tendrán todos tamaños diferentes, y la competición durante la PCR significará que se amplificarán preferentemente los más pequeños. Una solución será realizar una identificación de secuencia con varias enzimas diferentes en paralelo, lo que aumentará la cantidad de trabajo necesario para cada línea a analizar.

Una solución a este problema es no digerir el ADN usando enzimas de restricción, sino usar fragmentación aleatoria (digestión parcial de ADN usando un agente de corte frecuente, uso de ADNasa o sonicación). Estas estrategias serán menos propensas a amplificación no aleatoria pero son más pesadas de realizar.

Una segunda clase de protocolos se basa en PCR inversa (Ochman H., Gerber A. S., y Hartl D. L., 1988, Genetics, 120: 621-623). El ADN se digiere con una enzima de restricción y después se vuelve a ligar usando ADN ligasa. A baja concentración de ADN, el acontecimiento de ligación más frecuente es la recirculación de fragmentos de ADN. Esto permite la amplificación usando dos cebadores específicos de gen. Una ventaja de la iPCR es que las secuencias flanqueantes pueden identificarse después de una sola PCR (dependiendo del tamaño del genoma del hospedador) o de dos PCR anidadas. Los inconvenientes son como se han descrito para los métodos previos: será necesario el uso de varias enzimas diferentes para encontrar todas las inserciones.

Una tercera clase de métodos se basa en la amplificación por PCR usando un cebador específico de región (gsp) conocido y un cebador aleatorio inespecífico. Un ejemplo de dicha estrategia es la TAIL-PCR. La TAIL-PCR requiere tres rondas de PCR anidada y varias reacciones en paralelo con diferentes conjuntos de cebadores aleatorios (Yao-Guang Liu y Robert F. Whittier, 1995, Genomics, 25: 674-681). Una ventaja de este protocolo es que es un método de sólo PCR y, por lo tanto, puede automatizarse. Un inconveniente es que no siempre funcionará y no identificará

todas las secuencias flanqueantes cuando estén presentes múltiples inserciones. Además, es un protocolo bastante pesado puesto que cada muestra requiere muchas amplificaciones por PCR.

5 Se ha desarrollado un protocolo que permite la identificación de todas las inserciones presentes en un genoma en un solo experimento. Se prepara ADN genómico usando un kit convencional con una columna. El ADN producido usando kits comerciales convencionales está generalmente compuesto por fragmentos que varían de promedio de 10 kb a 25 kb. Esto significa que para cualquier secuencia conocida, aparecerá un punto de rotura del ADN en de 0 nt a 25 kb en cualquier secuencia dada, de promedio. Por lo tanto se aprovecha esta observación y se añade una cola directamente a los extremos rotos de ADN genómico con una cola de G usando transferasa terminal (Fig. 4), 10 siendo extremadamente sencillo de realizar y extremadamente útil, como muestran los resultados. Como se realiza para la amplificación de extremos de ADNc, se realiza una primera PCR usando un cebador BAPpC y un cebador específico de secuencia (gsp1) conocido. Esta primera PCR amplificará los productos específicos, pero también fragmentos inespecíficos pequeños con cebador BAPpC solamente. Esto es porque a todos los fragmentos de ADN genómico se les añaden colas en ambos lados mediante transferasa terminal. Se realiza una segunda PCR con un 15 cebador BAP y un cebador específico de secuencia (gsp2) conocido anidado para enriquecer la proporción de secuencia específica de gen respecto a secuencias inespecíficas. Incluso después de esta segunda PCR, se espera que se amplifiquen productos tanto específicos como inespecíficos. Por lo tanto, para clonar sólo los productos específicos, se usa el proceso de clonación limpia de la presente invención. Se prepara un vector de clonación lineal que contiene en un extremo homología con el adaptador y en el otro extremo homología con la secuencia localizada 20 justo cadena arriba del gsp2. Este vector y el inserto se tratan mediante la polimerasa T4. Sólo se obtienen productos específicos de gen.

Este protocolo es extremadamente fiable: en un solo experimento de clonación pueden identificarse todas las 25 secuencias flanqueantes de todas las inserciones. La adición de colas totalmente aleatoria permite la amplificación cuantitativa de secuencias flanqueantes; por ejemplo, un genoma que contiene una inserción homocigota y la otra heterocigota identificará la primera inserción en dos tercios de los plásmidos secuenciados y la segunda inserción y las restantes en un tercio de las secuencias. Además, la clonación limpia permite secuenciar directamente todos los plásmidos obtenidos sin explorar previamente para plásmidos recombinantes que contengan una inserción 30 específica.

Un protocolo alternativo para la clonación de secuencias flanqueantes consiste en realizar como primera etapa unos pocos ciclos de amplificación con un solo cebador, el cebador específico de gen (gsp1, Fig. 5). Al producto de PCR se le añade después una cola usando transferasa terminal y después se realizan una o dos PCR. Después, el 35 fragmento de PCR se somete a clonación limpia como se ha descrito anteriormente.

En el Ejemplo 4 se identificaron de forma fiable las secuencias que flanquean inserciones de ADN-T en plantas transgénicas. En el caso de la identificación de secuencias de ARN (Ejemplo 2) se identificó con éxito las secuencias de regiones variables de inmunoglobulinas específicas de tumores que se presentan en la superficie de células B malignas. Las inmunoglobulinas, especialmente sus regiones variables, pueden usarse como antígenos vacunales 40 individualizados para el tratamiento del Linfoma No Hodgkin (Inoges et al., 2006, J Natl Cancer Inst., 98: 1292-1301; McCormick et al., 2008, Proc. Natl. Acad. Sci. USA, 105: 10131-10136).

Ejemplos

45 EJEMPLO 1

Construcción de vectores de clonación limpia

a) Construcción de pICH36166

50 El fragmento alfa de LacZ se amplificó por PCR a partir del plásmido pICH36166 (idéntico al pUC19 excepto por que el gen bla que proporciona resistencia a ampicilina se sustituyó por un gen de resistencia a kanamicina). Los cebadores para la amplificación fueron Igclon1 (SEC ID N°: 6) (5'-GGA GGG TTG AAG ACT **TGT CCA GAG CCG TCC AGC AAC TGC AGG CAG CTG GCA CGA CAG GTT TC**-3') y nospclon1 (SEC ID N°: 7) (5'-GAT CCT AGA TGT GGA AGA CTT TAC **CGG GTC ATA ACG TGA CTC CCT TAA TTC TCC GCT GCA GCG CGC GTT TCG GTG ATG ACG**-3') (Fig. 6). La parte 3' de ambos cebadores (subrayada) es específica para el fragmento alfa de LacZ en pICH30086. Esta secuencia está precedida por un sitio de reconocimiento de PstI (en cursiva). La parte del medio (línea de puntos en negrita) contiene las secuencias homólogas para la clonación limpia (bap2 en Igclon1 y la "secuencia de captura CS", en este caso específica de promotor Nos en nospclon1). La parte 5' de cada cebador es 60 homóloga a los extremos de un vector linealizado usado para la clonación, en este caso pICH36833 (regiones de secuencia C y D en la Fig. 6). Debería mencionarse que el vector de clonación limpia puede de hecho prepararse a partir de cualquier vector de clonación convencional distinto de pICH36833, siendo el único requisito seleccionar secuencias C y D para el diseño del cebador que coincidan con el vector seleccionado en la región que flanquea el sitio usado para linealizarlo. 65

El fragmento LacZ amplificado a partir de pICH30086 se amplificó por PCR en un volumen de 50 µl. El producto de

5 PCR se purificó a través de una columna (NucleoSpin Extract II, Macherey-Nagel, Düren, Alemania; volumen de elución 50 µl). El pICH36833 se cortó con Bpil y se purificó en columna también. Para realizar la clonación con T4, 2 µl del producto de PCR, 1 µl de vector digerido con Bpil, 2 µl de tampón de polimerasa T4 10x (New England Biolabs, Ipswich MA, Estados Unidos), 0,5 µl de ADN polimerasa T4 (NEB, 20 unidades/µl) y 14,5 µl de agua se mezclaron y se incubaron durante 5 minutos a temperatura ambiente. La mezcla se usó para transformar *E. coli* DH10b químicamente competentes mediante transformación por choque térmico, y se sembraron en placas en medios que contenían X-Gal y Carbenicilina (el pICH36833 tiene un gen bla que proporciona resistencia a carbenicilina). Los clones positivos pueden detectarse usando selección por azul/blanco, conteniendo las colonias azules la construcción de vector deseada (el pICH36833 no tiene un gen LacZ funcional, pero el vector de clonación limpia resultante si lo tiene). La región que contiene Bap2 y la "secuencia de captura" se secuencian antes del uso para comprobar que la secuencia de clonación Bap2 y la "secuencia de captura" son correctas.

10 Como se ha mencionado anteriormente, los vectores de clonación limpia resultantes contienen un casete LacZ α para la selección por azul/blanco flanqueado por dos sitios Pst1, y después flanqueado en la parte 5' de LacZ por una secuencia correspondiente al cebador específico de cola de G Bap2 (SEC ID N°: 8): gtccagagccgtccagcaa, y en el lado 3' por una secuencia de captura de 20-52 pb. En el caso del pICH36166 la secuencia de captura corresponde a parte del terminador Nos. Para la clonación limpia, los vectores de clonación limpia se digieren primero con PstI, que escinde en ambos lados del fragmento LacZ, produciendo un vector lineal que contiene en ambos extremos secuencias del Bap2 y de la secuencia de captura. Este fragmento lineal está listo entonces para tratamiento con exonucleasa y clonación (véase a continuación).

b) Construcción de otros vectores de clonación limpia

25 Todos los demás vectores de clonación limpia pueden prepararse con exactamente el mismo protocolo que se ha descrito anteriormente, con la única excepción de que el cebador nospclon1 tiene que reemplazarse por un cebador diferente en el que la secuencia de la secuencia de captura se cambia por cualquier secuencia que el usuario desee "capturar".

Ejemplo 2

Identificación de la región variable de una inmunoglobulina

a) Descripción general de los métodos:

35 Se aisló ARN a partir de 0,2 ml de suspensiones de células tumorales (5-15 millones de células) de biopsias de ganglio linfático a partir de pacientes con Linfoma No Hodgkin usando el kit de RNeasy (Quiagen) de acuerdo con el protocolo del fabricante. El ARN se eluyó usando 40 µl de agua sin ARNasa. La calidad del ARN se ensayó mediante electroforesis en gel de agarosa y la concentración se midió con un NanoDrop ND-1000 (Thermo Scientific). 0,5-1,0 µg de ARN se sometieron a transcripción inversa en ADNc usando el kit de transcriptasa inversa SuperScript III de InVitrogen y usando un Oligo dT20 (InVitrogen). Una mezcla de 10 µl que contenía el ARN, dNTP 1 mM, oligo dT20 5 µM en agua sin ARNasa se desnaturizó durante 5 minutos a 65°C en un termociclador de PCR (BIO-RAD). La mezcla se enfrió durante al menos 1 minuto a 4°C, después se añadieron 2 µl de tampón RT 10x, 4 µl de MgCl₂ 25 mM, 2 µl de DTT 0,1 M, 1 µl de ARNasa OUT y 1 µl de transcriptasa inversa SuperScript III. La mezcla de reacción se incubó en un termociclador de PCR durante 50 minutos a 50°C y después se inactivó térmicamente a 85°C durante 5 minutos. El ADNc resultante no se trató con ARNasaH, pero se purificó en columna (kit MN Extract II, Machery-Nagel) y se eluyó con 25 µl de tampón de elución. Se usaron 10 µl de ADNc purificado en columna para la adición de colas de G usando 5 unidades de transferasa terminal (NEB), dGTP 0,25 mM, CoCl₂ 0,25 mM en tampón NEB 4 1x en una mezcla de reacción de 50 µl. La reacción de adición de colas de G se incubó durante 45 minutos a 37°C en un termociclador de PCR. Se inactivó la transferasa terminal durante 15 minutos a 70°C. La mezcla de reacción de cola de G inactivada se usó directamente para una PCR para amplificar fragmentos de inmunoglobulina. La PCR para la amplificación de fragmentos de anticuerpo a partir de ADNc de 1^a cadena con cola de G se llevó a cabo usando un cebador específico de cola de G (Bap2pc (SEC ID N°: 9): gtccagagccgtccagcaa cccccccccccc) y cebadores específicos de inmunoglobulina (véase a continuación la descripción para amplificaciones específicas) derivados de la región constante. La reacción se preparó usando ADN polimerasa Taq Hot Start (Fermentas) en una mezcla de reacción de 50 µl que contenía 1 µl de ADNc con cola de G, dNTP 0,2 mM, MgCl₂ 2 mM, 2 unidades de Taq Hot Start, cebador específico de cola de G 0,2 µM, cebador específico de región constante de inmunoglobulina 0,2 µM. La amplificación por PCR usaba una etapa de 4 minutos de activación de la Taq Hot Start y desnaturización a 95°C, y 36 ciclos de 20 segundos de desnaturización a 95°C, 30 segundos de hibridación a 58°C y un minuto de elongación a 72°C. Se analizaron 10 µl de la reacción de PCR mediante electroforesis en gel de agarosa. Después, el producto de PCR se purificó en columna (usando el kit Machery-Nagel MN Extract II) para eliminar los cebadores y los dNTP restantes que podían inhibir la actividad exonucleasa 3'-5' de la ADN polimerasa T4 durante la etapa siguiente. El producto de PCR se eluyó de la columna en 50 µl de tampón de elución.

65 Se aislaron vectores de clonación limpia a partir de 1 ml de *E. coli* DH10B en fase estacionaria usando el kit Nucleospin Plasmid QuickPure (Macherey-Nagel) de acuerdo con el protocolo del fabricante. Se eluyó ADN de vector con 50 µl de tampón de elución. Se digirieron 3 µl del ADN de vector purificado en una reacción de 30 µl con

*Pst*I, se inactivó térmicamente durante 20 minutos a 80°C y se analizó mediante electroforesis en gel de agarosa. Para la clonación limpia con T4, se usó ADN polimerasa T4 (NEB) para mascar los extremos del vector y del producto de PCR, dando como resultado regiones monocatenarias de ADN en los extremos de los fragmentos de inserto y vector. Para realizar la clonación limpia con T4, 2 µl del producto de PCR, 1 µl del vector de clonación limpia con T4 digerido con *Pst*I, 2 µl de tampón de polimerasa T4 10x (NEB), 0,5 µl de ADN polimerasa T4 (NEB, 20 unidades/µl) y 14,5 µl de agua se mezclaron y se incubaron durante 5-15 min a temperatura ambiente. Las reacciones se enfriaron en hielo durante 1 minuto y se usaron directamente para transformar 100 µl de células de *E. coli* DH10B quimiocompetentes. Los clones se seleccionaron en placas de agar LB complementado con carbenicilina y X-Gal. Se usó selección por azul-blanco para identificar los clones que contenían insertos (blancos), mientras que los clones azules llevan vectores de clonación limpia con T4 sin digerir. Para analizar la eficacia de la clonación de secuencias de inmunoglobulina, se secuenciaron 48 clones seleccionados aleatoriamente de cada clonación usando un cebador de vector.

b) Clonación de secuencias de inmunoglobulina: resultados

Se analizaron dos muestras de tumor, conteniendo ambas una inmunoglobulina asociada a tumor del isotipo IgM, K (cadena pesada Mu, cadena ligera Kappa). Para cada biopsia y para tanto la cadena ligera como la cadena pesada, se realizaron dos productos de PCR independientes con diferentes cebadores y se clonaron. Esto se realizó para asegurarse de que se determinaba la secuencia asociada a tumor correcta. Puesto que se usaron dos cebadores específicos de región constante diferentes (por separado) para cada identificación de secuencia, también se generaron dos vectores de clonación limpias separados correspondientes. Por ejemplo, se generaron dos vectores de clonación limpia para la cadena pesada: pICH36101 y pICH36113 (Fig. 7 y 8) y se generaron dos vectores de clonación limpia para la cadena ligera, pICH36078 y pICH36083 (Fig. 9). En cada caso, la secuencia de captura no solapa con el cebador usado para la amplificación. Obsérvese que la secuencia de captura no tiene que estar exactamente cadena arriba del cebador usado para la amplificación; de hecho, el cebador usado para la amplificación y la secuencia de captura pueden estar separados por 10, 20 o incluso 30 nucleótidos y todavía funcionará la clonación limpia.

Para la muestra 1 (T096), los productos de PCR obtenidos por amplificación con el cebador **Mu2N** (SEC ID N°: 10) (tc tgc tga tgt cag agt tg) se clonaron en el vector de IgM **pICH36101**, y los productos obtenidos a partir de la amplificación con el cebador **Mu3N** (SEC ID N°: 11) (aag tcc tgt gcg agg cag) se clonaron en el vector de IgM **pICH36113**. Se analizaron por secuenciación 48 y 47 clones de *E. coli* resultantes, respectivamente. Pudieron obtenerse 92 secuencias, tres reacciones de secuencia fallaron. Se encontró que 87 de estas secuencias correspondían a fragmentos de longitud completa que contenían 5'UTR, líder, región variable y región constante parcial de las inmunoglobulinas clonadas. Las otras cinco secuencias también correspondían a inmunoglobulinas pero no eran de longitud completa. La eficacia de clonación para las secuencias de inmunoglobulina era del 100%, y el 95,6% de las secuencias correspondían a fragmentos de longitud completa.

Para la muestra 2 (T104) los productos de PCR obtenidos por amplificación con el cebador **Mu2N** se clonaron en el vector de IgM **pICH36101**, y los productos obtenidos a partir de la amplificación con el cebador **Mu3N** se clonaron en el vector de IgM **pICH36113**. 48 clones de *E. coli* resultantes se analizaron cada uno por secuenciación. Pudieron obtenerse 96 secuencias. 89 de estas secuencias correspondían a fragmentos de longitud completa que contenían 5'UTR, líder, región variable y región constante parcial de las inmunoglobulinas clonadas. Las otras siete secuencias también correspondían a inmunoglobulinas pero no eran de longitud completa. La eficacia de clonación para secuencias de inmunoglobulina era del 100% y el 93,7% de las secuencias correspondían a fragmentos de longitud completa.

Para la muestra 1 (T096), los productos de PCR obtenidos por amplificación con el cebador **KC2N** (SEC ID N°: 12) (GGA GGG CGT TAT CCA CCT TCC) se clonaron en el vector de IgK **pICH36078**, y los productos obtenidos a partir de la amplificación con cebador **KC3N** (SEC ID N°: 13) (tca gca ggc aca caa cag agg) se clonaron en el vector (**pICH36083**). 48 clones de *E. coli* resultantes se analizaron cada uno por secuenciación. Pudieron obtenerse 95 secuencias, una reacción de secuencia falló. 91 de estas secuencias correspondían a fragmentos de longitud completa que contenían 5'UTR, líder, región variable y región constante parcial de las inmunoglobulinas clonadas. Las otras cuatro secuencias también correspondían a inmunoglobulinas pero no eran de longitud completa. La eficacia de clonación para secuencias de inmunoglobulina era del 100% y el 95,8% de las secuencias correspondían a fragmentos de longitud completa.

Para la muestra 2 (T104), los productos de PCR obtenidos por amplificación con el cebador (**KC2N**) se clonaron en el vector de IgK (**pICH36078**) y los productos obtenidos a partir de la amplificación con el cebador (**KC3N**) se clonaron en el vector de IgK **pICH36083**. 48 clones de *E. coli* resultantes se analizaron cada uno por secuenciación. Pudieron obtenerse 96 secuencias. 86 de estas secuencias correspondían a fragmentos de longitud completa que contenían 5'UTR, líder, región variable y región constante parcial de las inmunoglobulinas clonadas. Las otras 10 secuencias también correspondían a inmunoglobulinas pero no eran de longitud completa. La eficacia de clonación para secuencias de inmunoglobulina era del 100% y el 89,6% de las secuencias correspondían a fragmentos de longitud completa.

Conclusión de estos experimentos: usando clonación limpia con T4, podían clonarse diferentes tipos de secuencias de inmunoglobulina que pertenecían a diferentes clases de inmunoglobulina. Los datos podían usarse para obtener con éxito el idiotipo de tumor exacto para las muestras de linfoma folicular no Hodgkin.

5 Ejemplo 3

Determinación de condiciones adecuadas para tratamiento con exonucleasa

a) Ensayo in vitro para cuantificación de actividad exonucleasa de T4

10 Se desarrolló un ensayo para determinar las condiciones adecuadas u óptimas para el tratamiento con T4. Este ensayo consiste en digerir ADN lineal con polimerasa T4 durante un tiempo dado, después incubar 5 minutos a 80°C para inactivar la polimerasa T4, y después incubar 20 minutos a 25°C con nucleasa de judía Mung para digerir los extremos de ADN monocatenario generados mediante la polimerasa T4. El ADN digerido se procesa después en un gel de agarosa para comparar los tamaños del fragmento resultante con ADN linealizado no tratado. El ADN que se usó es una digestión del plásmido pICH10990, que da como resultado cuatro fragmentos de tamaño de 3595, 1563, 1158 y 125 pb (sólo son útiles los 3 fragmentos más grandes puesto que son más visibles en un gel de agarosa teñido con bromuro de etidio convencional).

20 El ADN se digirió 10 minutos en presencia de polimerasa T4 a 25, 20, 15 y 10° C. Todas las incubaciones dan como resultado un cambio de tamaño (a productos más pequeños), siendo el cambio más marcado a mayores temperaturas (Fig. 10). Una incubación de 10 minutos a 25°C da como resultado una digestión de aproximadamente 100 a 300 nt (es visible una mancha). Considerando que la digestión debe tener lugar en ambos extremos de cada fragmento, la digestión en cada extremo es de aproximadamente 50 a 150 nt, indicando regiones monocatenarias de 50 a 150 nt en cada extremo de los fragmentos lineales después del tratamiento con T4.

30 Se realizó un segundo experimento para descubrir si la adición de bajas concentraciones de dNTP puede disminuir la velocidad de digestión de T4, para hacer que la digestión sea más fácil de manejar. Concentraciones inferiores a 1 µM todavía permiten que se formen regiones monocatenarias, mientras que una concentración de 10 µM y superior inhibe completamente la formación de regiones monocatenarias.

b) Ensayo de diferentes tiempos de digestión para clonación limpia con T4 de regiones variables de inmunoglobulina

35 Se clonaron regiones variables de anticuerpo amplificadas a partir de la biopsia T078 usando clonación limpia con T4, de acuerdo con el protocolo descrito en los ejemplos previos. El inserto se preparó con digestión con T4 a temperatura ambiente de 5 a 20 minutos. Se obtuvieron colonias blancas con los 4 puntos temporales, pero el máximo número de colonias se obtuvo con la digestión de 10 minutos (Fig. 11). Se secuenciaron los insertos de plásmidos de 170 colonias (una selección aleatoria de los diferentes puntos temporales) y se descubrió que todos contenían secuencias de inmunoglobulina.

40

Ejemplo 4

Aislamiento de secuencias que flanquean inserciones de ADN-T

45 Las plantas de *Nicotiana benthamiana* nbi157-3, nbi158-3, nbi159-3 y nbi160-7 son plantas transgénicas generadas por transformación mediada por *Agrobacterium* de una construcción (pICH30649) que contiene en el borde derecho secuencias de terminador Nos y secuencias no traducidas 3' NTR de un vector viral basado en cr-VMT (Fig. 12). Se aisló el ADN genómico a partir de 100 mg de material foliar usando el Kit NucleoSpin® Plant II (Macherey-Nagel) de acuerdo con el protocolo del fabricante. Como todos los protocolos de extracción de ADN, este procedimiento de aislamiento da como resultado cierta rotura del ADN genómico.

50

55 En el primer ensayo, se realizó la adición de colas de G sobre el ADN genómico usando diferentes cantidades de ADN (0,1, 0,3 ó 1 µg). Al ADN se le añadieron colas de G usando 5 unidades de transferasa terminal (NEB), dGTP 0,25 mM, CoCl₂ 0,25 mM y tampón NE 4 1x en una mezcla de reacción de 50 µl. La reacción de adición de colas de G se incubó durante 45 minutos a 37°C en un termociclador de PCR. Se inactivó la transferasa terminal durante 15 minutos a 70°C. Se usaron cebadores específicos de 3'-NTR para 2 rondas de PCR anidada. La primera PCR se realizó con *bap2pC* (SEC ID N°: 9) (gtccagagccgtccagcaac ccc ccc ccc ccc c) y 3'-NTRf1 (SEC ID N°: 14) (gcgacgatagcgcatagtg), con las condiciones de PCR siguientes: 4 minutos a 94°C seguido de 35 ciclos de 10 s a 94°C, 30 s a 58°C y 90 s a 72°C, seguido de 4 min a 72°C. La segunda PCR se realizó con los cebadores *bap2* (SEC ID N°: 8) (gtccagagccgtccagcaac) y 3'-NTRf2 (SEC ID N°: 15) (atccgtaggggtggcgtaaac) con las mismas condiciones de PCR. Como se esperaba, es visible una mancha de productos de PCR en ambas PCR (Fig. 13). Los productos de ambas PCR (partiendo de 0,1 µg de ADN) se purificaron en columna (kit NucleoSpin Extract II, Macherey-Nagel, 25 µl de PCR se purificaron y se eluyeron en 25 µl de tampón de elución) antes de la clonación en vector de clonación limpia (pICH36261).

65

El vector de clonación pICH36261 (Fig. 12) se cortó con PstI y se purificó en columna también. La reacción de clonación se preparó con 1,5 µl de vector (aproximadamente 40 ng), 4 µl de producto de PCR, 2 µl de tampón de polimerasa T4 10x y 0,5 µl de ADN polimerasa T4 en un volumen total de 20 µl. La reacción se incubó 5 min a temperatura ambiente y la mezcla completa se usó para transformar 100 µl de células de *E. coli* DH10b químicamente competentes usando transformación por choque térmico. Las células transformadas se seleccionaron en medio complementado con X-gal. Se exploraron las colonias blancas mediante PCR de colonias (usando dos cebadores en el vector de clonación que flanquean la inserción, kanseqf (SEC ID N°: 16): tggaaaaacgccagcaacgc y kanseqr (SEC ID N°: 17): tgtctcatgagcggatacat) para determinar la presencia y el tamaño del inserto (Fig. 14). El vector vacío (religado en el sitio PstI) proporciona un fragmento de 295 nt. Todos los fragmentos de PCR de mayor tamaño indican que el vector contiene un inserto.

Sólo había unas pocas colonias blancas (0-6) después de la clonación de productos de PCR1 (indicando que sólo estaba presente una pequeña cantidad de productos específicos en la primera PCR), mientras que los productos de la PCR2 proporcionaban 20-120 colonias blancas. Se exploraron 24 colonias por cada planta mediante PCR de colonias (Fig. 14). Se encuentra principalmente vector vacío (o sin producto) en los clones de la PCR1 (16 de 17 clones), mientras que en los clones de la PCR2 se encuentran insertos de aproximadamente 150 pb hasta >1000 pb. Se secuenciaron 78 clones. Un clon (de la clonación después de la primera PCR) contenía un producto inespecífico. Las otras 77 secuencias eran específicas de transgén.

Ejemplo 5

Aislamiento de secuencias que flanquean inserciones de ADN-T con o sin preamplificación

Existen muchas formas por las que pueden amplificarse secuencias flanqueantes de secuencias de ADN conocidas. Se ha ejemplificado una forma en el ejemplo anterior. Sin embargo, la clonación limpia es adecuada para productos de PCR que se obtienen con cualquier otra estrategia diseñada para amplificar secuencias flanqueantes. Aquí se proporciona un ejemplo para ilustrar un protocolo de amplificación ligeramente diferente. Sin embargo, la clonación limpia se realiza exactamente como para el ejemplo anterior.

La planta *Nicotiana benthamiana* nbi160-7 es un transgénico para la construcción pICH30649, pero también contiene un segundo transgén de la construcción pICN18693 (Fig. 15). Para esta inserción, el promotor Nos se localiza próximo al borde derecho. Se realizó un experimento de clonación para comparar la clonación limpia después de la adición de colas de G directa de ADN genómico o después de la adición de colas de G del producto de una preamplificación lineal realizada con un solo cebador.

La preamplificación se realizó en una reacción de 50 µl que contenía 100 ng de ADN_g, dNTP (0,2 mM de cada uno) y 0,2 µM de cebador específico de inserto nospromfwd (SEC ID N°: 18) (5'-TTT GCT AGC TGA TAG TGA CCT TAG GCG AC-3'). Después de la desnaturalización inicial (2 min, 94°C), se realizaron 10 ó 25 ciclos del programa siguiente: 10 s a 94°C, 30 s a 58°C, 90 s a 72°C. Las muestras se purificaron a través de una columna (kit NucleoSpin Extract II, Machery-Nagel, 50 µl de producto de preamplificación se purificaron y se eluyeron en 30 µl de tampón de elución). La muestra completa se usó para unir una cola de G a los extremos libres del ADN usando una transferasa terminal (NEB) en un volumen de reacción de 50 µl como se ha descrito en el ejemplo anterior. También se realizó la adición de colas de G sobre 100 ng de ADN genómico sin preamplificación.

Después se usaron dos µl de ADN con colas de G en una reacción de PCR de 50 µl usando cebador específico de pICH18693 nosprev2 (SEC ID N°: 19) (5'-TGC GGT TCT GTC ACT TCC AAA CG-3') y el cebador específico de cola de G bap2pc (SEC ID N°: 9) (5'-GTC CAG AGC CGT CCA GCA ACC CCC CCC CCC-3'). Se realizó una PCR anidada con 1 µl del producto de PCR con nosprev2 en una reacción de 50 µl usando el cebador específico anidado bvipcr3 (SEC ID N°: 20) (5'-CGG CTT GTC CCG CGT CAT C-3') y el cebador bap2 (SEC ID N°: 8) (5'-GTC CAG AGC CGT CCA GCA AC-3'). Los resultados de estas PCR se muestran en la Fig. 16.

Productos de todas las PCR se purificaron usando el Kit NucleoSpin Extract II (Machery-Nagel). El vector de clonación pICH36166 (Fig. 15) se cortó con PstI y se purificó en columna también. La reacción de clonación se preparó con 1,5 µl de vector (aproximadamente 40 ng), 4 µl de producto de PCR, 2 µl de tampón de polimerasa T4 10x y 0,5 µl de ADN polimerasa T4 (NEB) en un volumen total de 20 µl. La reacción se incubó 5 min a temperatura ambiente y después se usó completamente para transformar 100 µl de células de *E. coli* DH10b químicamente competentes, usando transformación por choque térmico. Las células transformadas se seleccionaron en medio complementado con X-gal. Las colonias blancas se exploraron mediante PCR de colonias (usando dos cebadores en el vector de clonación que flanquean la inserción, kanseqf y kanseqr) para determinar la presencia y el tamaño del inserto (Fig. 16). Usando una preamplificación (1b, 1c, 2b, 2c), todas las colonias ensayadas llevaban un inserto de un tamaño variable (aproximadamente 100-800 pb), incluso cuando se usaba el producto de la 1ª PCR (1b, 1c). Por el contrario, la clonación eficaz de muestras sin preamplificación sólo era posible después de la PCR anidada (2a frente a 1a).

Se aisló el ADN plasmídico de 71 clones y se secuenciaron los insertos. Las 71 secuencias identificaron productos específicos. En algunos casos en los geles de la Fig. 16 pueden observarse dos bandas después de la exploración

por PCR de colonias; la secuenciación puso de manifiesto que estos son los resultados de dos colonias que se escogieron juntas, y son de hecho una mezcla de dos colonias, cada una con un inserto específico.

Ejemplo 6

5

Uso de Klenow para la generación de extremos de ADN monocatenarios

a) Ensayo *in vitro* para cuantificación de la actividad exonucleasa de Klenow

10 Klenow es conocida por tener una actividad exonucleasa que es mucho más débil que la de la polimerasa T4 en condiciones de digestión similares. Para cuantificar esto, se ensayó la actividad exonucleasa de klenow usando el ensayo descrito anteriormente en el Ejemplo 3. Se descubrió que a temperatura ambiente (25°C) no puede detectarse una actividad exonucleasa significativa con el ensayo (Fig. 17). Por el contrario, se forman regiones monocatenarias a 37°C con incubaciones que varían de 30 a 90 minutos. Esta observación es muy útil, puesto que en condiciones de pipeteo convencional (20 a 25°C) casi no hay presente actividad exonucleasa. Esto permite pipetear todas las muestras sin preocuparse de que las primeras muestras estén sobredigeridas mientras que las últimas muestras no estén listas todavía. Cuando se termina el pipeteo de todas las muestras, todas las muestras pueden transferirse a un baño de agua o termociclador a 37°C, donde tendrá lugar una reacción más controlada.

20 b) Ensayo de clonación con Klenow

Se realizó una clonación usando una biblioteca de fragmentos de *Arabidopsis thaliana* aleatorios amplificados por PCR (c en la Fig. 17) en un vector de clonación linealizado pICH31401. La homología entre el inserto y el vector es de 20 nt en un extremo (cgccggtcacaaggtcagct (SEC ID N°: 21)) y 22 nt en el otro extremo (gccaggatctgtgtctcaatt (SEC ID N°: 22)). El mayor número de colonias se obtuvo con la muestra para la que se realizó una digestión con klenow durante 30 minutos a 37°C.

ANEXO

30 Secuencia de pICH36166 (SEC ID N°: 23)

La secuencia de Bap2 (gtccagagccgtccagcaa comenzando por el principio) está en negrita, así como la secuencia de captura (cggagaattaagggagtcacgttatgaccc). Los dos sitios PstI están en cursiva y subrayados.

gtccagagccgtccagcaa*ctgcaggcagctggc*acgacaggttcccgactggaaagcgggcagtgagcgcgaacgc
 aatagtgagttagctcactcattaggcaccaggttacactttatgctccggctcgtatgtgtggaattgtgagcggata
 acaatttcacacaggaaacagctatgacatgattacgcaagcttgatgcctgcaggtcgactctagaggatccccgggta
 ccgagctcgaattcactggccgtcgtttacaacgtcgtgactgggaaaaccctggcgttacccaactaatcgcttgcagcac
 atcccccttcgccagctggcgtaatagcgaagaggcccgcaccgatcgccctccaacagttgcgagcctgaatggcga
 atggcgctgatgcggtattttctccttacgcatctgtcgggtatttcacaccgcataatggcactctcagtaaatctgctgatg
 ccgcatagttaagccagccccgacaccgccaacaccgctgacgcgcccctgacgggctgtctgctcccggcatccgctta
 cagacaagctgtgaccgtctccgggagctgcatgtgtcagaggtttaccgctatcaccgaaacgcgcg*ctgcaggcggaga*
attaagggagtcacgttatgacccggtaaagcttccacatctaggatctgccaggaaccgtaaaaagggcgtgctggc
 gttttccataggctccgccccctgacgagcatcaaaaaatcgacgctcaagtcagaggtggcgaaacccgacaggacta
 taaagataccagggcgtttcccctggaagctccctcgtgcgctcctgttccgaccctgccgcttaccggatacctgtccgcctt
 ctcccctcgggaagcgtggcgtttctcatagctcacgctgtaggtatctcagttcgggttaggtcgttcgctccaagctgggctgtg

35

**tgacgaacccccggtcagcccagccgctgcgccctatccggtaactatcgtcttgagtccaacccggtgaagacacgacttat
 cgccactggcagcagccactggtaacaggattagcagagcgaggtatgtaggcggtgctacagagttctggaagtgggtggcct
 aactacggctacactagaaggacaglatttggtatctgcgctctgctgaagccagttacctcggaaaaagagttggtagctctg
 atccggcaaaacaaaccaccgctggtagcgggtggtttttgtttgcaagcagcagattacgcgagaaaaaaaggatctcaag
 aagatcctttgatctttctacggggctgacgctcagtggaacgaaaactacgftaagggattttggtcatgagattatcaaaaa
 ggatcttacctagatccttttaaaatfaaaalgaagtttaaatcaatctaaaagtataatgagtaaaacttggtctgacagttacca
 tgcttaatcagtgaggcacctatctcagcgaictgctatttcgttcaiccatagttgcctgactccccgctcgtgtagataaactacgat
 acgggagggcttacatctggcccagtgctgcaatgataaccgagacccacgctcaccggctccagattatcagcaata
 aaccagccagccggaaggccgagcgcagaagtggtcctgcaactttatccgctccatccagttatattgttgccggga
 agctagagtaagtagtccagtaaatggttgcgaacggttgccattgctacagggatcgtgggtgacgctcgtcgtttggt
 atggcttattcagctccggttcccaacgatcaaggcgagttacatgatccccatggttgcaaaaaagcgggttagctccttcgg
 tcctccgatcgttgcagaagtaagttggccgagtgatcactcatgggtatggcagcactgcataattcttactgtcatgccat
 ccgtaagatgctttctgtgactggtgagtactcaaccaagtcattctgagaatagtgatgcgggcagccgagttgctcttgcccgg
 cgtaatacgggataataaccgcccacatagcagaactttaaagtgtctcatcattggaaaaacggttctcggggcgaaaactct
 caaggatctaccgctgtgagatccagttcagatgtaaccactcgtgcaccaactgatcttcagcatctttactttcaccagcgt
 ttctgggtgagcaaaaacaggaaggcaaaatgccgcaaaaaagggataaagggcgacacggaaatgttgaatactcatac
 tctccttttcaatattatgaagcatttatcaggggtattgtctcatgagcggatacatattgaaatgtattagaaaaataaacaat
 aggggttccgcgacgaattggccagcgtgccattttgggtgagggccggtcgcggccgagggggcgagcccctggggggg
 atgggaggcccgcttagcggggccgggaggggtgaagactt**

LISTADO DE SECUENCIAS

- 5 <110> Icon Genetics GmbH
- <120> Proceso de clonación limpia
- <130> EAD-14942
- 10 <160> 23
- <170> PatentIn versión 3.3
- 15 <210> 1
- <211> 10
- <212> ADN
- <213> Artificial
- 20 <220>
- <223> Cola de G
- <400> 1
- ggggggggggg 10
- 25 <210> 2
- <211> 30
- <212> ADN
- <213> Artificial
- 30 <220>
- <223> Secuencia para hibridación por homología
- <400> 2
- 35 cggagaatta agggagtcac gttatgacct 30
- <210> 3
- <211> 312
- <212> ADN
- 40 <213> Artificial
- <220>
- <223> Porción de secuencia del vector pICH36113

	<400> 3	
	gggagtgcat cegcccaaac ccttttcccc ctegtctcct gtgagaattc cccgtcggat	60
	acgagcagcg tggccgttgg ctgcctcgca caggacttcc ttcccgactc catcacttts	120
	tcttggaat acaagaacaa ctctgacatc agcagyaccg ggggcttccc atcagtcctg	180
	agagggggca agtacgcagc cacctcacag gtgctgctgc cttccaagga cgtcatgcag	240
	ggcacagacg aacacgtggt gtgcaaagtc cagcacccca acggcaacaa agaaaagaac	300
	gtgcctcttc ca	312
5		
	<210> 4	
	<211> 320	
	<212> ADN	
	<213> Artificial	
10		
	<220>	
	<223> Porción de secuencia del vector pICH36083	
	<400> 4	
15		
	gaactgtggc tgcaccatct gtcttcatct tcccgccatc tgatgagcag ttgaaatctg	60
	gaactgcctc tgttgtgtgc ctgctgaata acttctatcc cagagaggcc aaagtacagy	120
	ggaaggtgga taacgccctc caatcgggta actcccagga gagtgtcaca gagcaggasa	180
	gcaaggacag cacctacagc ctcagcarca ccctgacgct gagcaaagca gactacgaga	240
	aacacaaagt ctacgcctgc gaagtcaccc atcagggcct gagctcgcct gtcacaaaga	300
	gcttcaacag gggagagtgt	320
	<210> 5	
	<211> 25	
	<212> ADN	
	<213> Artificial	
20		
	<220>	
	<223> Secuencia para hibridación por homología	
25		
	<400> 5	
	ttaacaggtg atccaggaaa taagg	25
	<210> 6	
	<211> 62	
	<212> ADN	
	<213> Artificial	
30		
	<220>	
	<223> Cebador de PCR IgclonI	
35		
	<400> 6	
	ggaggggttga agacttgtcc agagccgtcc agcaactgca ggcagctggc acgacaggtt	60
	tc	62
40		
	<210> 7	
	<211> 81	

<212> ADN
 <213> Artificial

 <220>
 5 <223> Cebador de PCR nospclonl

 <400> 7

gatcctagat gtggaagact ttaccggggtc ataacgtgac tcccttaatt ctccgctgca 60
gcgcgcgttt cggatgatgac g 81
 10
 <210> 8
 <211> 19
 <212> ADN
 <213> Artificial
 15
 <220>
 <223> Cebador Bap2

 <400> 8
 20 gtccagagcc gtccagcaa 19

 <210> 9
 <211> 33
 <212> ADN
 25 <213> Artificial

 <220>
 <223> Cebador de PCR Bap2pc
 30
 <400> 9
 gtccagagcc gtccagcaac ccccccccc ccc 33

 <210> 10
 <211> 19
 35 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador de PCR Mu2N
 40
 <400> 10
 tctgctgatg tcagagttg 19

 <210> 11
 <211> 18
 45 <212> ADN
 <213> Artificial

 <220>
 50 <223> Cebador de PCR Mu3N

 <400> 11
 aagtcctgtg cgaggcag 18

 <210> 12
 <211> 21
 55 <212> ADN
 <213> Artificial

 <220>
 60 <223> Cebador de PCR KC2N

<400> 12
 ggagggcgt atccacctc c 21

5 <210> 13
 <211> 21
 <212> ADN
 <213> Artificial

10 <220>
 <223> Cebador de PCR KC3N

<400> 13
 tcagcaggca cacaacagag g 21

15 <210> 14
 <211> 20
 <212> ADN
 <213> Artificial

20 <220>
 <223> Cebador de PCR 3'-NTRf1

<400> 14
 ggcgacgata ggcgatagtg 20

25 <210> 15
 <211> 21
 <212> ADN
 <213> Artificial

30 <220>
 <223> Cebador de PCR 3'-NTRf2

35 <400> 15
 atccgtaggg gtggcgtaaa c 21

<210> 16
 <211> 20
 <212> ADN
 <213> Artificial

40 <220>
 <223> Cebador de PCR kanseqf

45 <400> 16
 tggaaaaacg ccagcaacgc 20

50 <210> 17
 <211> 20
 <212> ADN
 <213> Artificial

55 <220>
 <223> Cebador de PCR kanser

<400> 17
 tgtctcatga gcggatacat 20

60 <210> 18
 <211> 29
 <212> ADN
 <213> Artificial

<220>
 <223> Cebador de PCR nospromfwd
 5 <400> 18
 ttgctagct gatagtgacc ttaggcgac 29

 <210> 19
 <211> 23
 10 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador de PCR nosprev2
 15 <400> 19
 tgcggtctg tcagttcaa acg 23

 <210> 20
 <211> 19
 20 <212> ADN
 <213> Artificial

 <220>
 <223> Cebador de PCR bvipcr3
 25 <400> 20
 cggctgtcc cgcgtcatc 19

 <210> 21
 <211> 20
 30 <212> ADN
 <213> Artificial

 <220>
 <223> Secuencia de homología
 35 <400> 21
 cgccggtctc aaggtcagct 20
 40 <210> 22
 <211> 22
 <212> ADN
 <213> Artificial
 45 <220>
 <223> Secuencia de homología

 <400> 22
 50 gccaggatct gtggtctcaa tt 22

 <210> 23
 <211> 2591
 <212> ADN
 55 <213> Artificial

 <220>
 <223> Secuencia del plásmido pICH36166
 60 <400> 23

ES 2 364 920 T3

gtccagagcc gtccagcaac tgcaggcagc tggcagcaca ggtttccgga ctggaaagcg 60
ggcagtgagc gcaacgcaat taatgtgagt tagctcactc attaggcacc ccaggcttta 120
cactttatgc ttccggctcg tatgttgtgt ggaattgtga gcggataaca atttcacaca 180
ggaaacagct atgaccatga ttacgccaag cttgcatgcc tgcaggtcga ctctagagga 240
tccccgggta ccgagctcga attcactggc cgtcgtttta caacgtcgtg actgggaaaa 300
ccctggcggt acccaactta atcgcttgc agcacatccc cctttcgcca gctggcgtaa 360

tagcgaagag gcccgcaccg atcgcccttc ccaacagttg cgcagcctga atggcgaatg	420
gcgcctgatg cggatatttc tccttacgca tctgtgcggt atttcacacc gcatatgggtg	480
cactctcagt acaatctgct ctgatgccc atagttaagc cagccccgac acccgccaac	540
acccgctgac ggcacctgac gggcttgtct gctcccggca tccgcttaca gacaagctgt	600
gaccgtctcc gggagctgca tgtgtcagag gttttcaccg tcatcaccga aacgcgcgct	660
gcagcggaga attaagggag tcacgttatg acccggtaaa gtcttcaca tctaggatct	720
gccaggaacc gtaaaaaggc cgcgttgctg gcgttttcc ataggctccg cccccctgac	780
gagcatcaca aaaatcgacg ctcaagtcag aggtggcgaa acccgacagg actataaaga	840
taccaggcgt tccccctgg aagctccctc gtgcgctctc ctgttccgac cctgccgctt	900
accggatacc tgtccgcctt tctcccttcg ggaagcgtgg cgtttctca tagctcacgc	960
tgtaggatc tcagttcggg gtaggtcgtt cgctccaagc tgggctgtgt gcacgaacct	1020
cccgttcagc ccgaccgctg cgccttatcc ggtaactatc gtcttgagtc caaccggta	1080
agacacgact tategccact ggcagcagcc actggtaaca ggattagcag agcgaggat	1140
gtaggcggtg ctacagagtt cttgaagtgg tggcctaact acggctacac tagaaggaca	1200
gtatttggtg tctgcgctct gctgaagcca gttaccttcg gaaaagagt tggtagctct	1260
tgatccggca aacaaccac cgctggtagc ggtggttttt ttgtttgca gcagcagatt	1320
acgcgcagaa aaaaaggatc tcaagaagat cctttgatct tttctacggg gtctgacgct	1380
cagtggaaag aaaactcag ttaagggatt ttggtcatga gattatcaaa aaggatcttc	1440
acctagatcc ttttaaatta aaaatgaagt tttaaatcaa tctaaagtat atatgagtaa	1500
acttggctc acagttacca atgcttaatc agtgaggcac ctatctcagc gatctgtcta	1560
ttctgtcat ccatagttgc ctgactcccc gtcgtgtaga taactacgat acgggagggc	1620
ttaccatctg gccccagtgc tgcaatgata ccgcgagacc cacgctcacc ggctccagat	1680
ttatcagcaa taaaccagcc agccggaagg gccgagcgca gaagtggctc tgcaacttta	1740
tccgctcca tccagtctat taattgttgc cgggaagcta gagtaagtag ttcgccagtt	1800
aatagtttgc gcaacgttgt tgccattgct acaggcatcg tgggtgcacg ctctgcttt	1860
ggtatggctt cattcagctc cggttcccaa cgatcaaggc gagttacatg atccccatg	1920
ttgtgcaaaa aagcggttag ctcttcgggt cctccgatcg ttgtcagaag taagttggcc	1980
gcagtgttat cactcatggt tatggcagca ctgcataatt ctcttactgt catgccatcc	2040
gtaagatgct tttctgtgac tgggtgagtac tcaaccaagt cattctgaga atagtgtatg	2100
cggcgaccga gttgctcttg cccggcgtca atacgggata ataccgccc acatagcaga	2160
actttaaaag tgctcatcat tggaaaacgt tcttcggggc gaaaactctc aaggatctta	2220
ccgctgttga gatccagttc gatgtaacct actcgtgcac ccaactgatc ttcagcatct	2280

ES 2 364 920 T3

tttactttca	ccagcgtttc	tgggtgagca	aaaacaggaa	ggcaaaatgc	cgcaaaaaag	2340
ggaataaggg	cgacacggaa	atgttgaata	ctcatactct	tcctttttca	atattattga	2400
agcatttata	agggttattg	tctcatgagc	ggatacatat	ttgaatgtat	ttagaaaaat	2460
aaacaaatag	gggttccgcg	cacgaattgg	ccagcgtctc	catttttggg	gtgaggccgt	2520
tcgcggccga	ggggcgcagc	ccctgggggg	atgggaggcc	cgcgttagcg	ggccgggagg	2580
gttgaagact	t					2591

REIVINDICACIONES

1. Un proceso de inserción de una secuencia de ácido nucleico de interés en un ácido nucleico aceptor, que comprende las etapas siguientes:

5
amplificar por PCR un ADN que comprende, en el siguiente orden, un segmento de secuencia U, un segmento de secuencia de ácido nucleico de secuencia de nucleótidos conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3, usando un cebador directo que define un primer extremo del ADN amplificado y un cebador inverso que define un segundo extremo del ADN amplificado, terminando dicho cebador inverso en su extremo 3' en una secuencia de nucleótidos de segmentos de secuencia de ácido nucleico K3;

10
tratar las moléculas de ADN bicatenario lineal contenidas en el producto de PCR obtenido en la etapa anterior con una exonucleasa para obtener un saliente monocatenario en el primer extremo del ADN y un saliente monocatenario que comprende los segmentos de ácido nucleico K2 y K3 en el segundo extremo del ADN;

15
hibridar el producto de la etapa anterior con un ácido nucleico aceptor bicatenario linealizado que tiene en un primer extremo del mismo un saliente monocatenario complementario al saliente monocatenario del primer extremo del ADN y, en un segundo extremo del mismo, un saliente monocatenario complementario al segmento de secuencia monocatenario K2 del segundo extremo del ADN; y

20
usar el producto de reacción obtenido en la etapa anterior para transformar una célula hospedadora.

25
2. El proceso de acuerdo con la reivindicación 1, que comprende además generar un molde para dicha PCR, que comprende unir una secuencia de unión a cebador a una secuencia de ácido nucleico que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3, hibridando dicho cebador directo con la secuencia de unión a cebador.

30
3. El proceso de acuerdo con la reivindicación 2, en el que la secuencia de unión a cebador es un segmento de secuencia de nucleótidos homooligomérica unido usando una desoxirribonucleótido transferasa terminal.

35
4. El proceso de acuerdo con la reivindicación 2 ó 3, en el que dicho cebador directo comprende un segmento de secuencia adaptadora y un segmento de secuencia complementario a la secuencia de unión a cebador.

40
5. El proceso de acuerdo con la reivindicación 1, que comprende además generar un molde para dicha PCR, que comprende unir una secuencia adaptadora a una secuencia de ácido nucleico que comprende, en el siguiente orden, el segmento de secuencia de ácido nucleico U en el primer extremo de la secuencia de ácido nucleico, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3.

45
6. El proceso de acuerdo con una cualquiera de las reivindicaciones 1 a 5, en el que dicho ácido nucleico aceptor no contiene una porción de secuencia complementaria al segmento de secuencia monocatenario K3 en el saliente monocatenario en el segundo extremo del ácido nucleico aceptor.

7. El proceso de acuerdo con una cualquiera de las reivindicaciones 1 a 6, en el que la exonucleasa es ADN polimerasa T4 de *E. coli*, el fragmento grande de la polimerasa I fragmento grande de *E. coli* (como polimerasa Klenow), nucleasa lambda, nucleasa T7 o exonucleasa III, preferentemente es la ADN polimerasa T4 de *E. coli*.

50
8. El proceso de acuerdo con las reivindicaciones 1 a 7, en el que se realiza una etapa de preamplificación por PCR antes de dicha reacción de PCR definida en la reivindicación 1, usando un cebador directo que define el primer extremo del producto de PCR de la etapa de preamplificación, y un cebador inverso que termina en su extremo 3' en una secuencia de nucleótidos de segmento de secuencia de ácido nucleico K3.

55
9. El proceso de acuerdo con la reivindicación 2 ó 5, en el que la etapa de unión se realiza en una mezcla que comprende ADN genómico aislado a partir de células eucariotas o procariontas.

60
10. El proceso de acuerdo con la reivindicación 2 ó 5, en el que dicha etapa de unión se realiza en el extremo 3' de un ADNc retro transcrito a partir de un ARNm.

65
11. Un proceso de inserción de un ADNc de una secuencia de ARN de interés en un ácido nucleico aceptor, que comprende las etapas siguientes:

aislar ARN de una célula;

retrotranscribir el ARN para formar una cadena no codificante de ADNc;

amplificar por PCR dicho ADNc que comprende, en el siguiente orden, un segmento de secuencia U, un segmento de secuencia de ácido nucleico de secuencia conocida K2 y un segmento de secuencia de ácido nucleico de secuencia conocida K3, usando un cebador directo que define un primer extremo del ADN amplificado y un cebador inverso que define un segundo extremo del ADN amplificado, terminando dicho cebador inverso en su extremo 3' en una secuencia de nucleótidos de segmento de secuencia K3;

tratar las moléculas de ADN bicatenario lineal contenidas en el producto de PCR obtenido en la etapa anterior con una exonucleasa para obtener un saliente monocatenario en el primer extremo del ADN y un saliente monocatenario que comprende los segmentos de ácido nucleico K2 y K3 en el segundo extremo del ADN;

hibridar el producto de la etapa anterior con un ácido nucleico aceptor bicatenario linealizado que tiene en un primer extremo del mismo un saliente monocatenario complementario al saliente monocatenario del primer extremo del ADN, y en un segundo extremo del mismo un saliente monocatenario complementario al segmento de secuencia monocatenaria K2 del segundo extremo del ADN; y

transformar el producto de reacción obtenido en la etapa anterior en una célula hospedadora.

12. El proceso de acuerdo con una cualquiera de las reivindicaciones 1 a 11, en el que la secuencia de interés comprende en dirección 5' a 3' un segmento de secuencia U, un segmento de secuencia K2 de secuencia de nucleótidos conocida y un segmento de secuencia K3 de secuencia de nucleótidos conocida.

13. El proceso de acuerdo con la reivindicación 11, que comprende proporcionar un molde para dicha PCR por unión de una secuencia de unión a cebador al extremo 3' de dicho ADNc, en el que el cebador directo usado en dicha PCR hibrida con dicha secuencia de unión a cebador.

14. El proceso de acuerdo con la reivindicación 11, en el que el ARN es ARNm.

15. El proceso de acuerdo con la reivindicación 11, en el que el ARN codifica una cadena de inmunoglobulina, la secuencia U es parte de una región variable de la inmunoglobulina y las secuencias K2 y K3 conocidas son parte de una región constante de la inmunoglobulina.

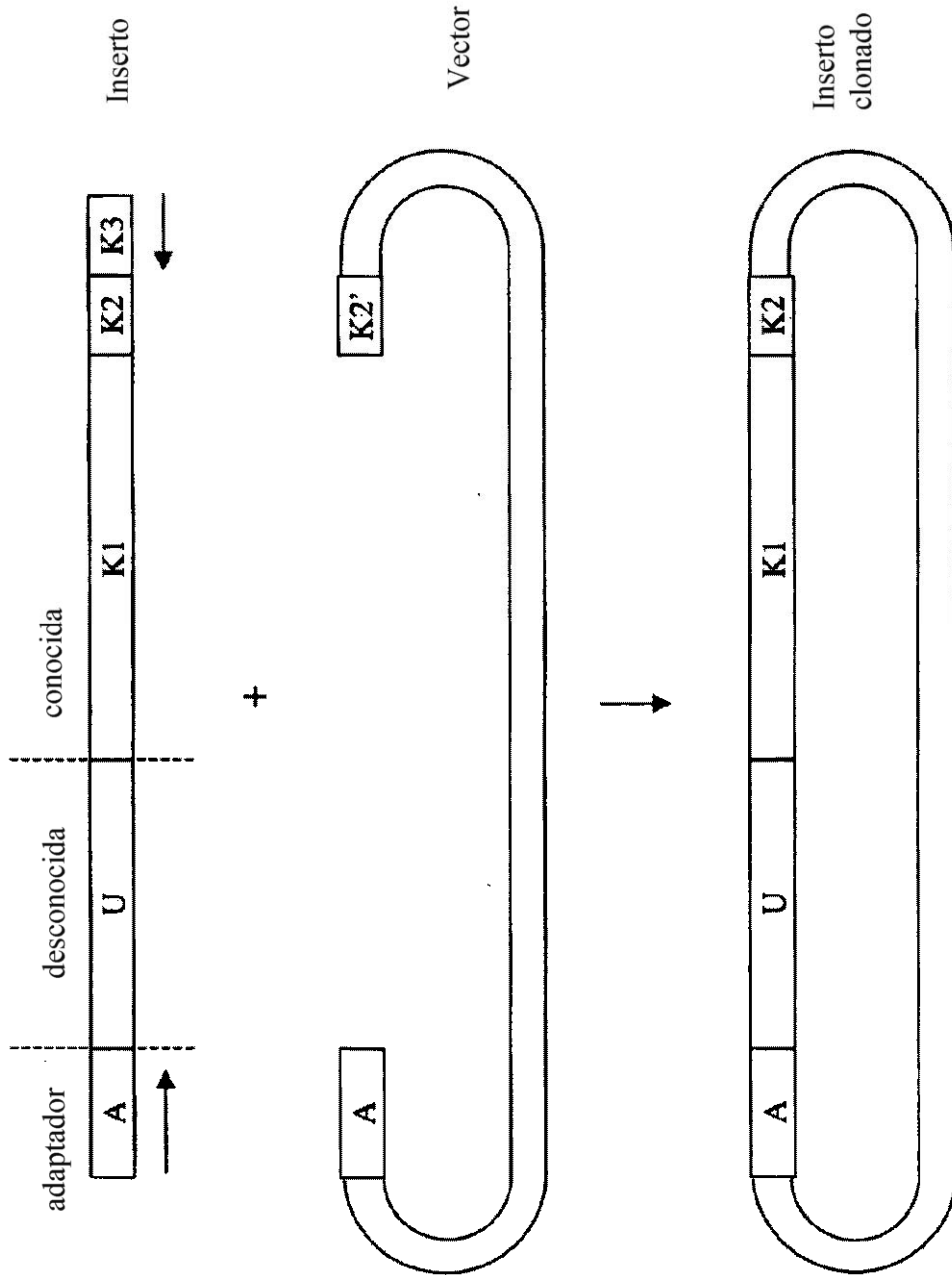
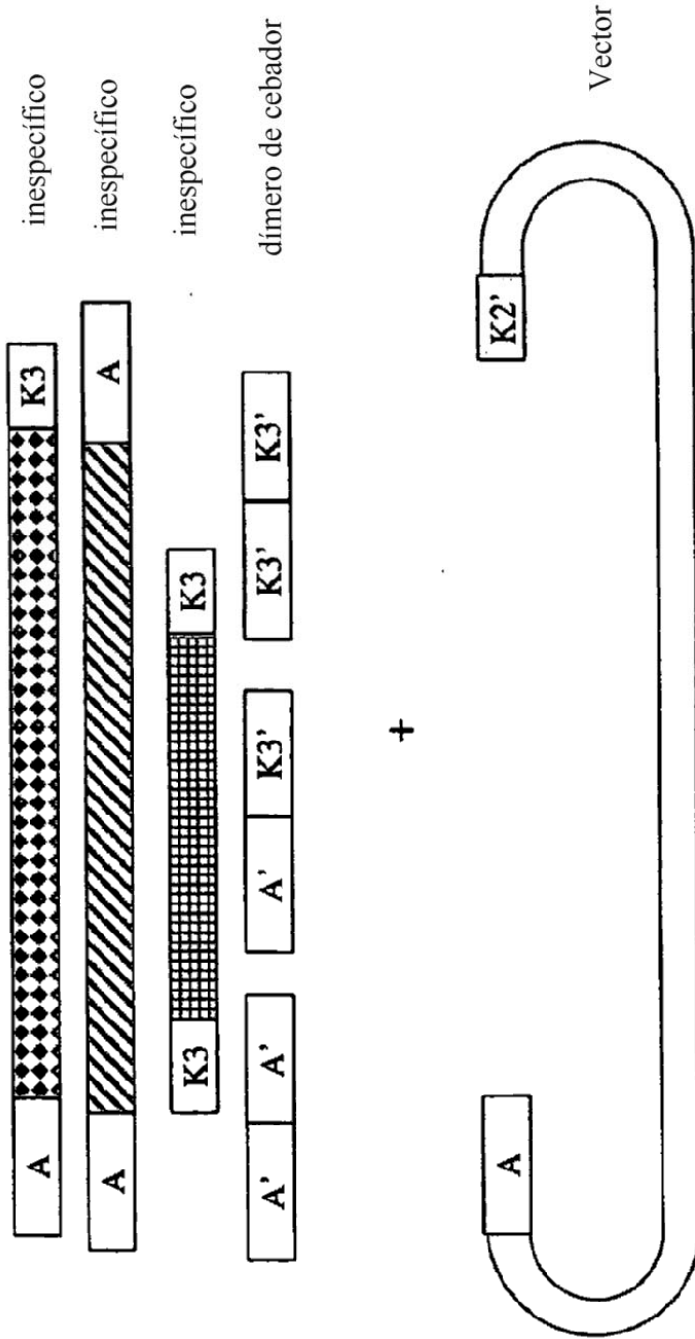


Fig. 1a



No plásmidos recombinantes,
 los productos inespecíficos no pueden clonarse

Fig. 1b

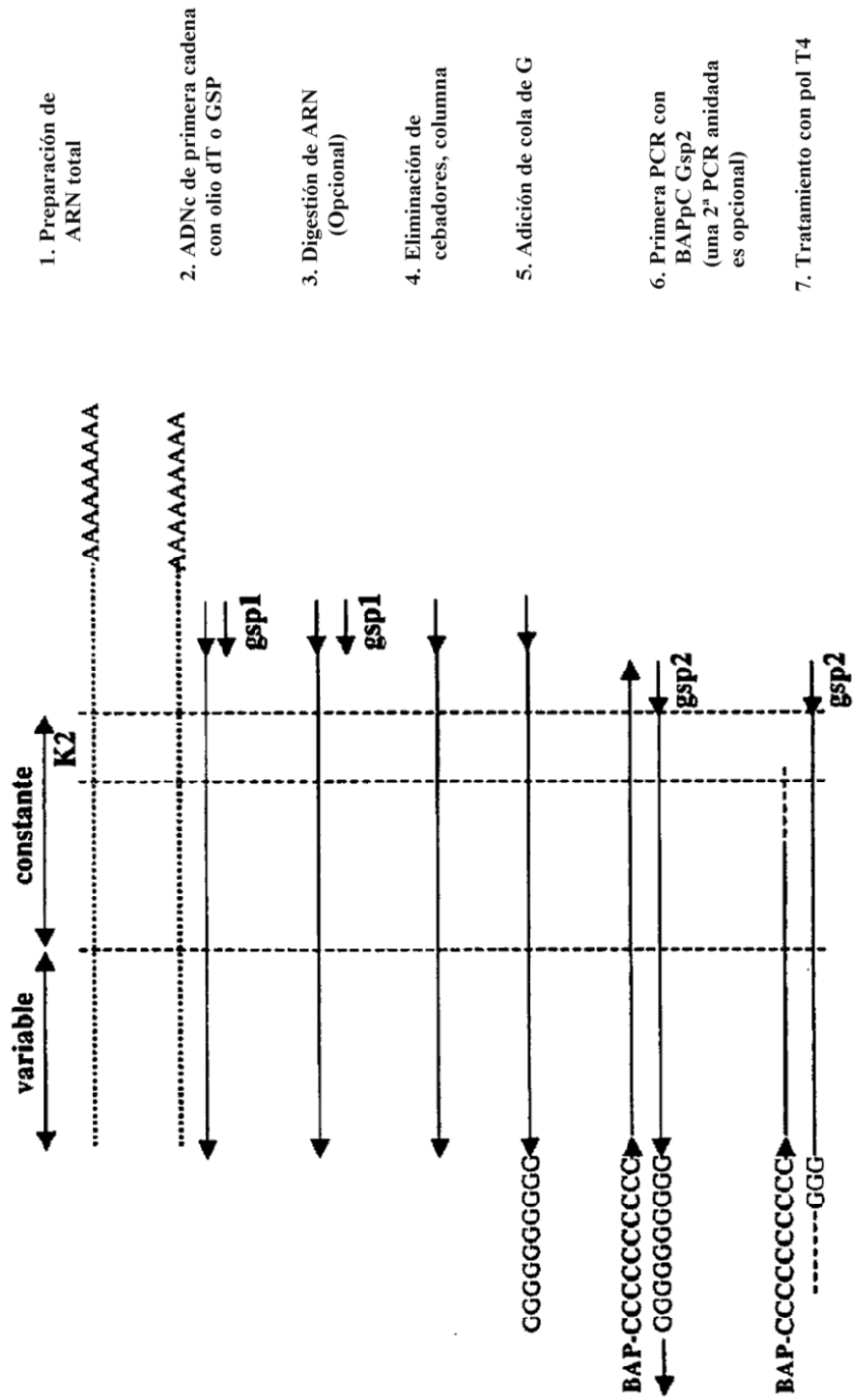


Fig. 2

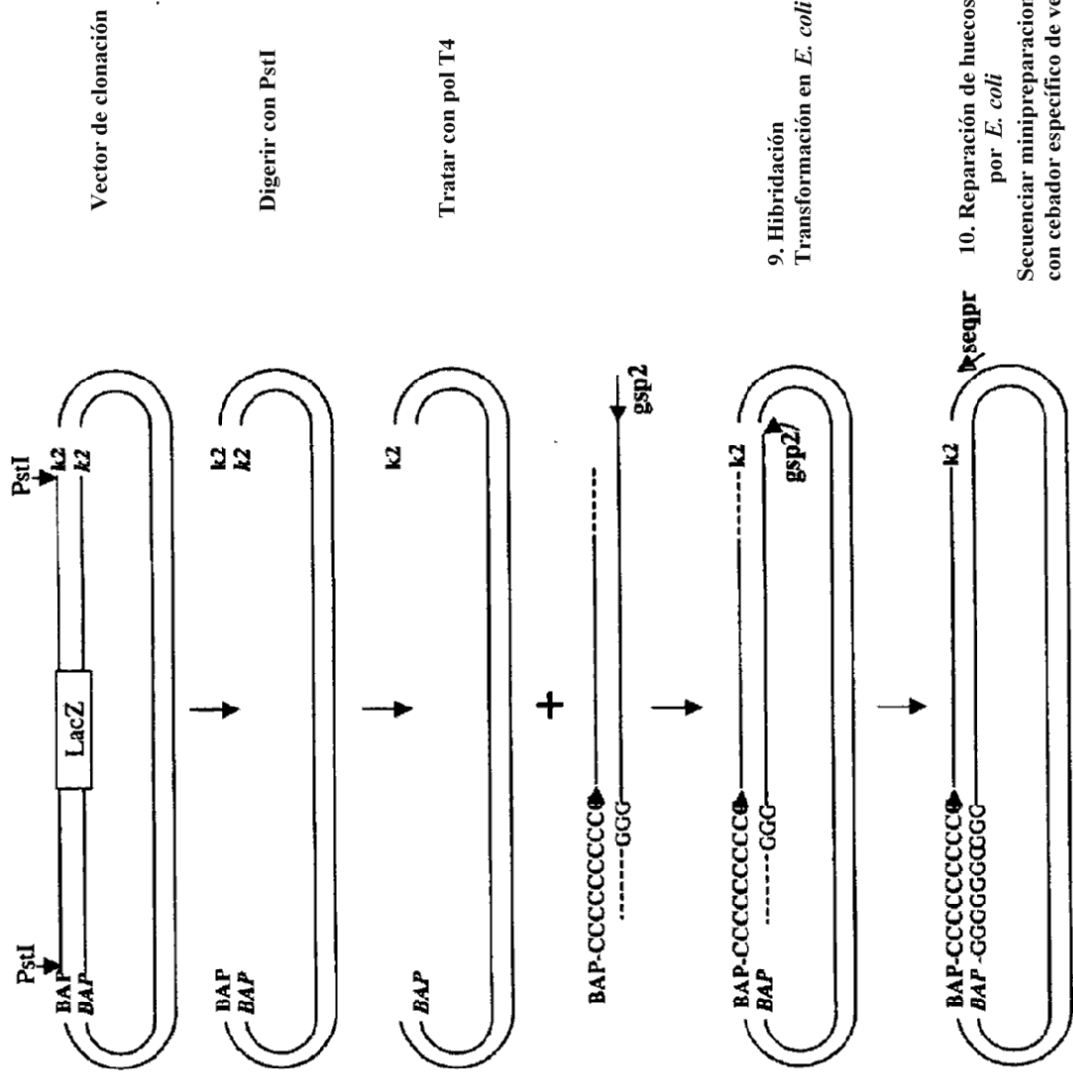


Fig. 3

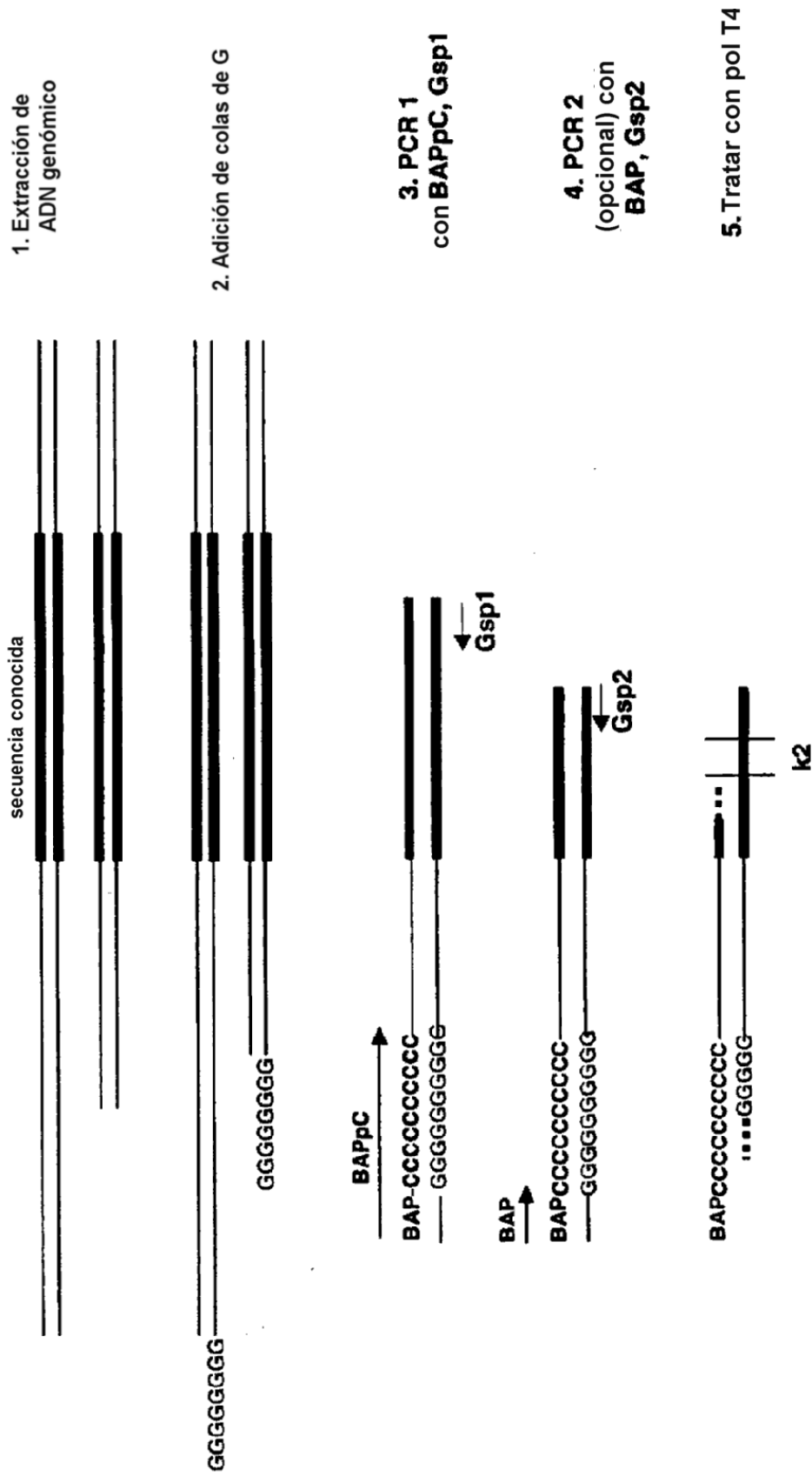


Fig. 4

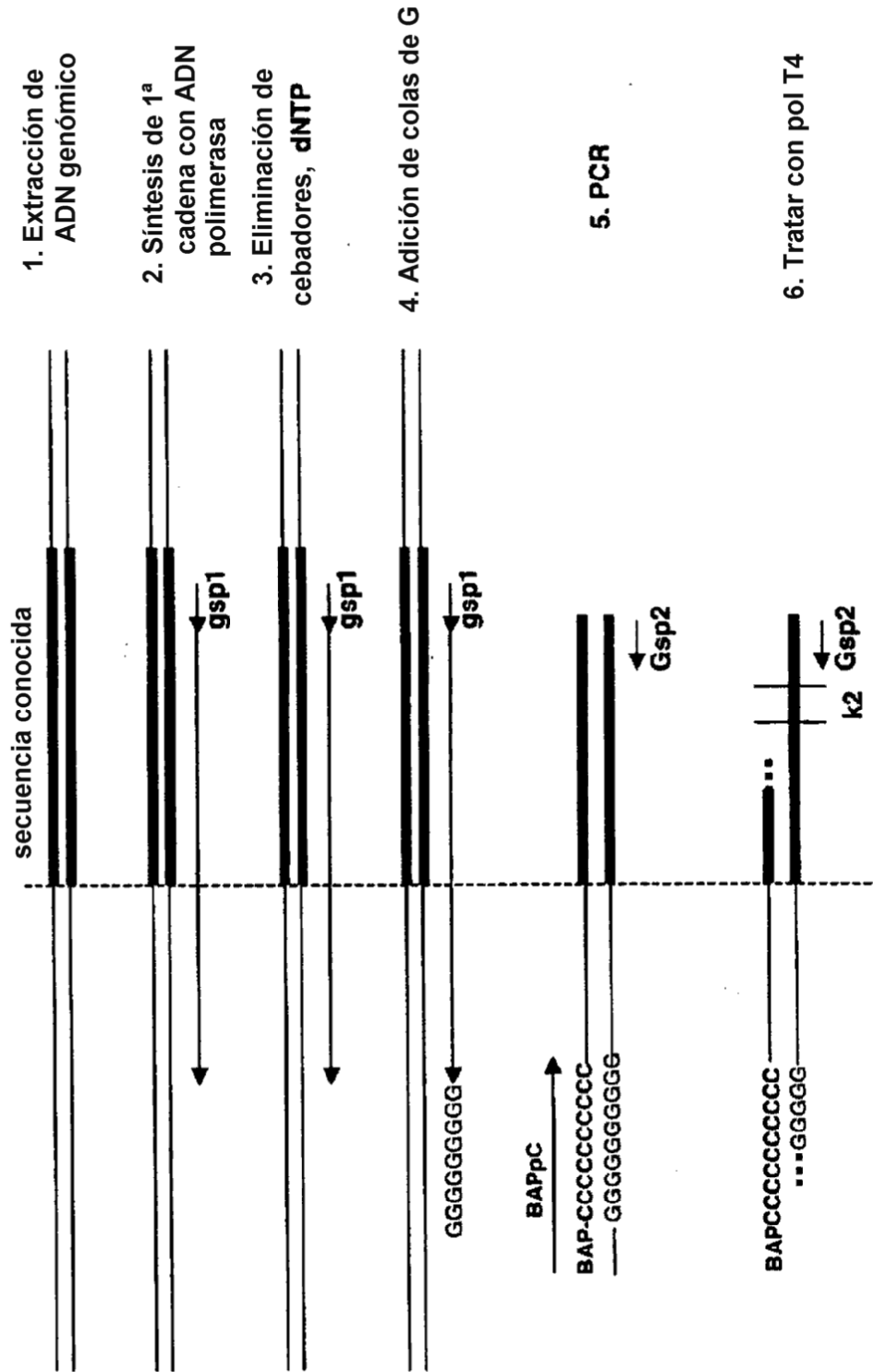


Fig. 5

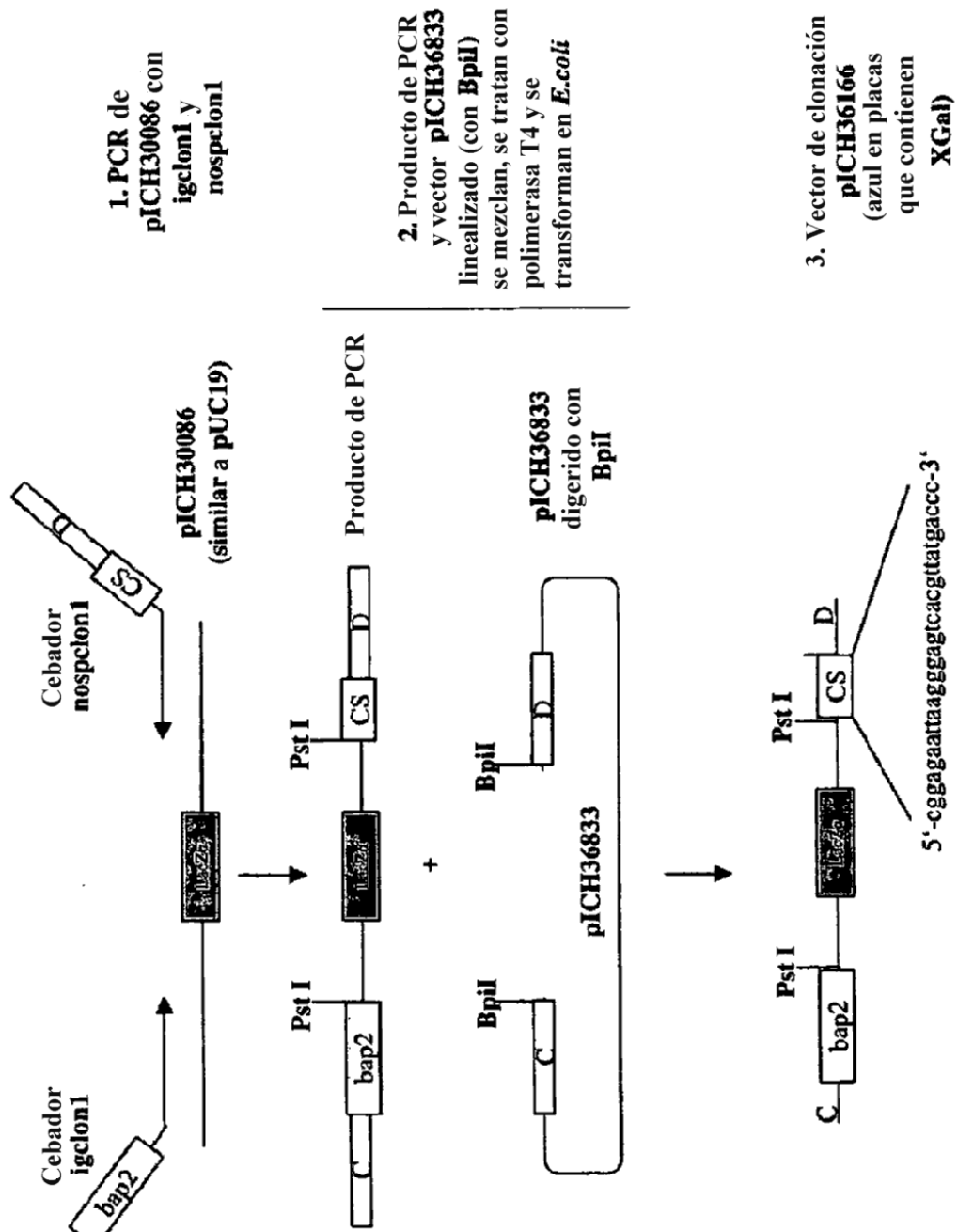


Fig. 6

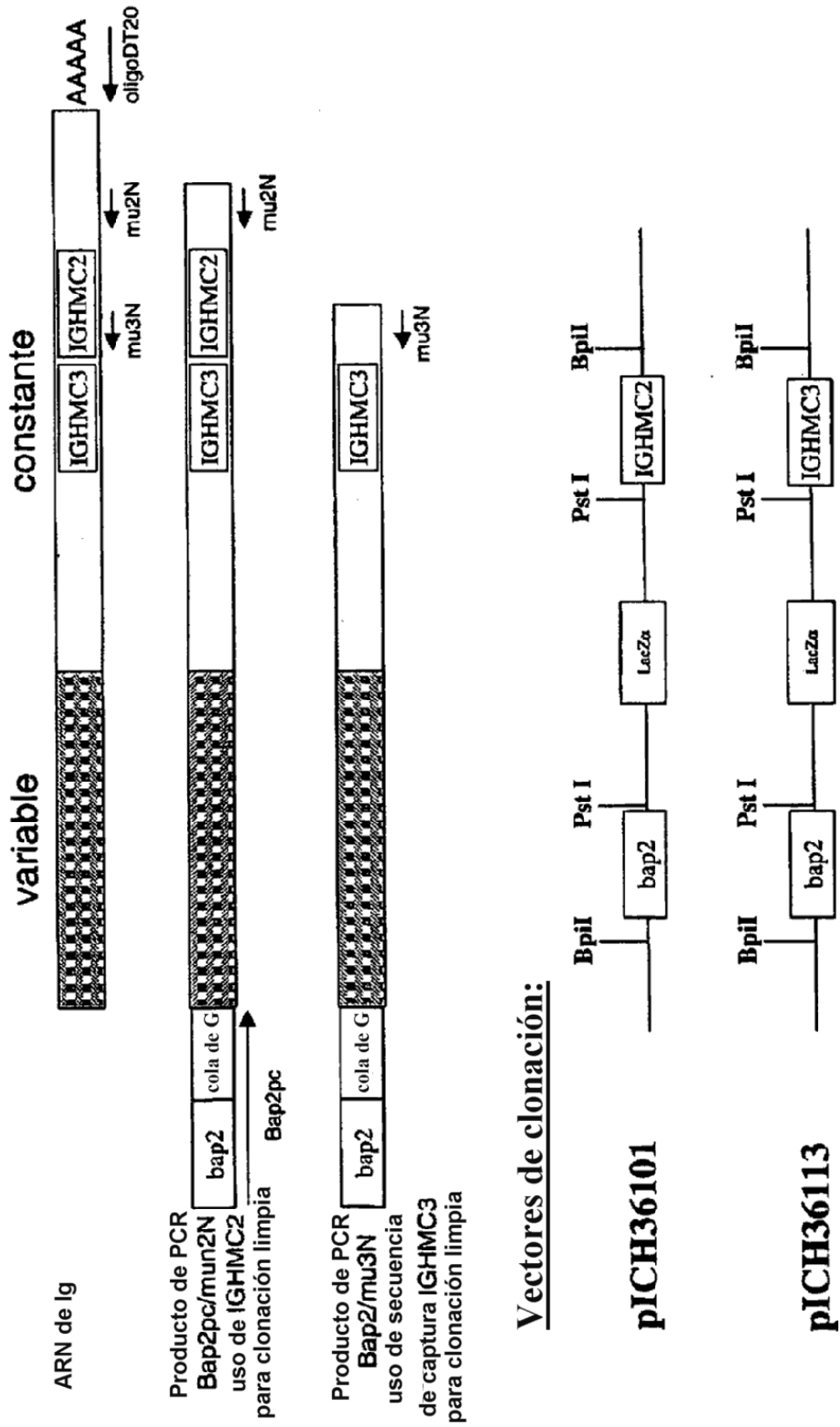


Fig. 7

Secuencia de captura **IGHMC3** en **pICH36113**
 para clonar productos de PCR amplificados con **Bap2PC** y **Mu3N**

(g)ggagtgcac cgcgcccac ccttttccc ctcgtctcct gtgagaattc cccgtcggat acgagcagcg tggccggttgg

Secuencia de captura **IGHMC2** en **pICH36101**
 para clonar productos de PCR amplificados con **Bap2PC** y **Mu2N**

ctgcctcgca caggacttcc ttccgactc catcactttt tcttggaat acaagaacaa ctctgacatc agcag~~y~~accc
 Mu3N Mu2N

ggggcttccc atcagtcccg agagggggca agtacgcag cacctcacag gtgctgctgc ctccaagga cgtcatgcag

ggcacagacg aacacgtggt gtgcaaagtc cagcaccaca acggcaacaa agaaaagaac gtgccctcttc ca

Fig. 8

Secuencia de captura en **pICH36083**
 para clonar productos de PCR amplificados con **Bap2PC** y **KC3N**

GAACTGTGGC **TG**CACCAATCT **GT**CTTCATCT **TCCC**GCCATC **TG**ATGAGCAG **TT**GAAATCTG **GA**ACTGCCCTC **TG**TGTGTGTC
 KC3N

Secuencia de captura en **pICH36078**
 para clonar productos de PCR amplificados con **Bap2PC** y **KC2N**

CTGTGAATA **AC**TTCTATCC **CAG**AGAGGCC **AA**AGTACAGY **GGA**AGGTGGA **TA**ACGCCCTC **CA**ATCGGGTA **AC**TCCCAGGA
 KC3N

GAGTGTCACA **GAG**CAGGASA **GCA**AGGACAG **CAC**CTACAGC **CTC**AGCARCA **CC**CTGACGCT **GAG**CAAAGCA **GAC**TACGAGA

AACACAAAGT **CT**AGGCCCTGC **GA**AGTCACCC **AT**CAGGGCCT **GAG**CTCGCCC **GTC**ACAAAGA **GCT**CAACAG **GGG**AGAGTGT

Fig. 9

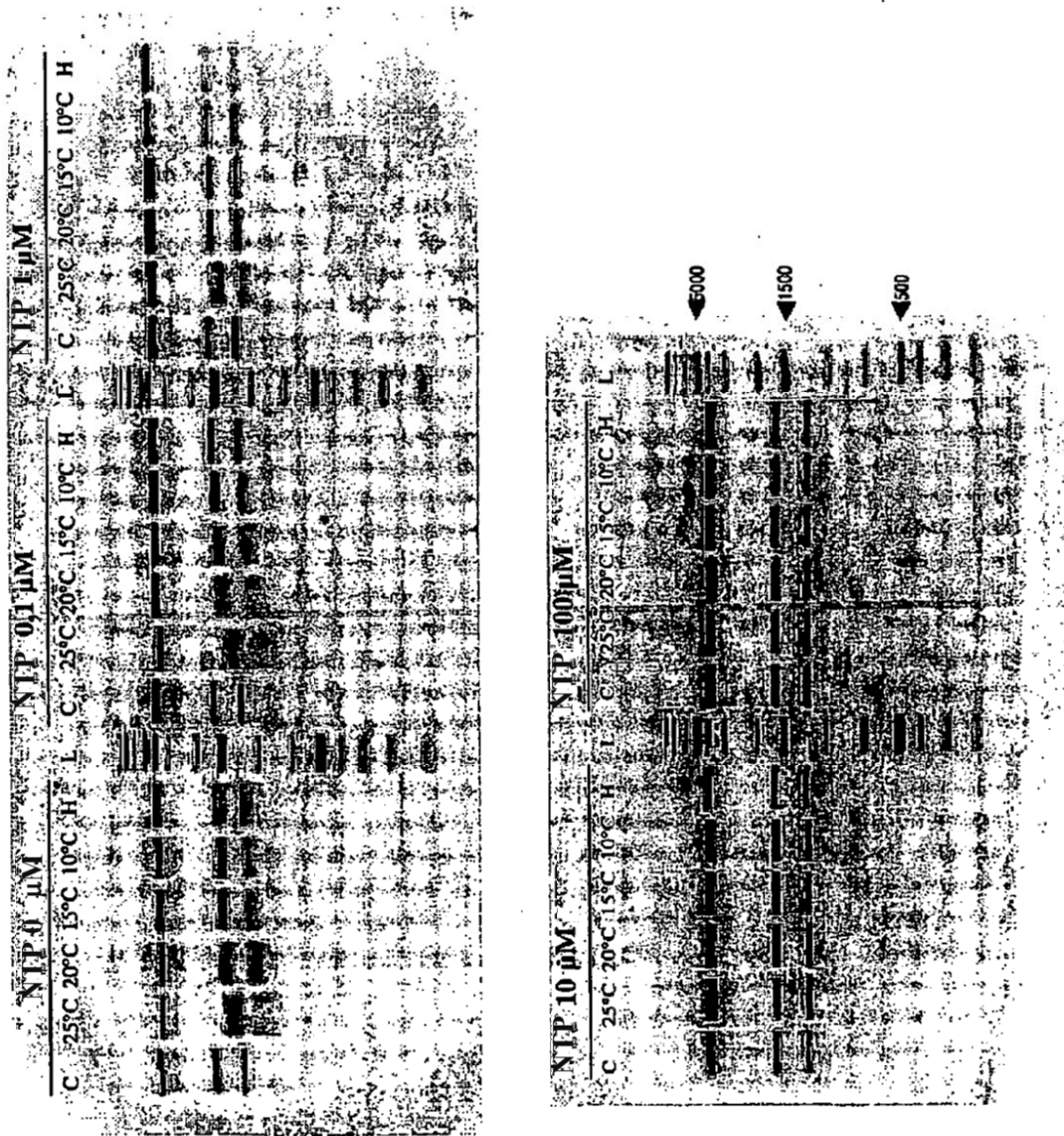
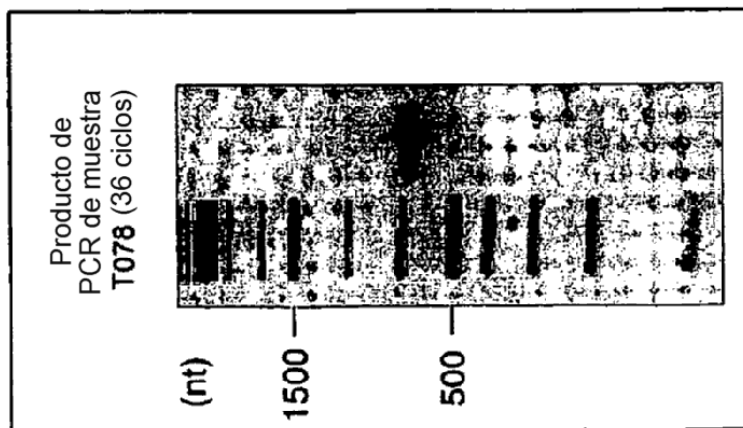


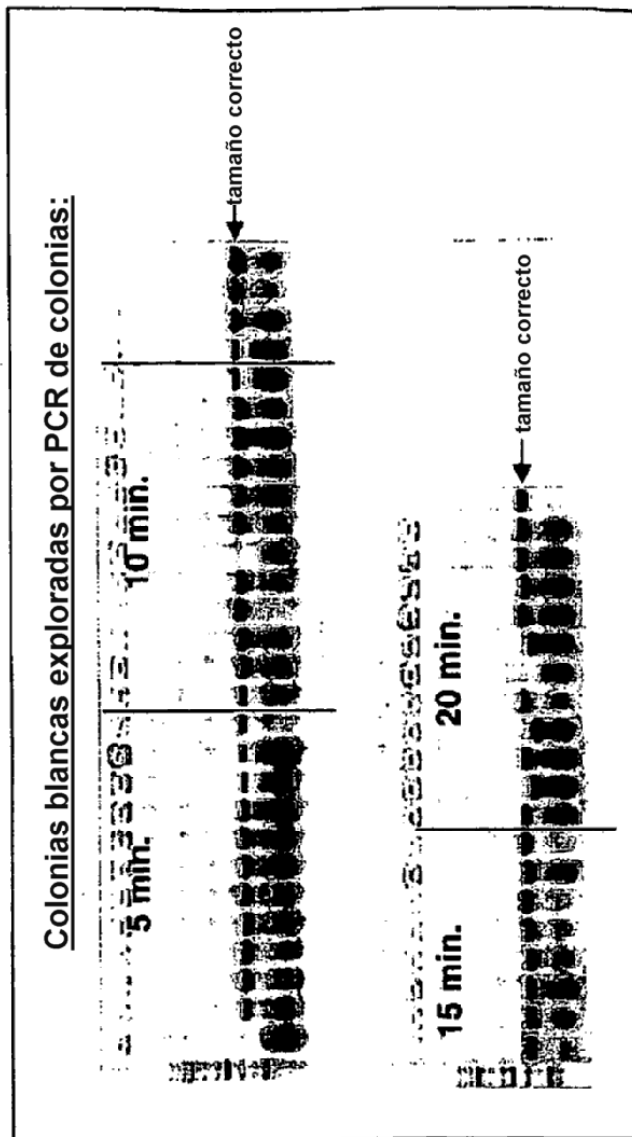
Fig. 10

Fig. 11



COLONIAS BLANCAS:

Mezcla de transformación	50 µl	Resto
5 min.	49	>200
10 min.	123	>200
15 min.	35	>200
20 min.	7	101



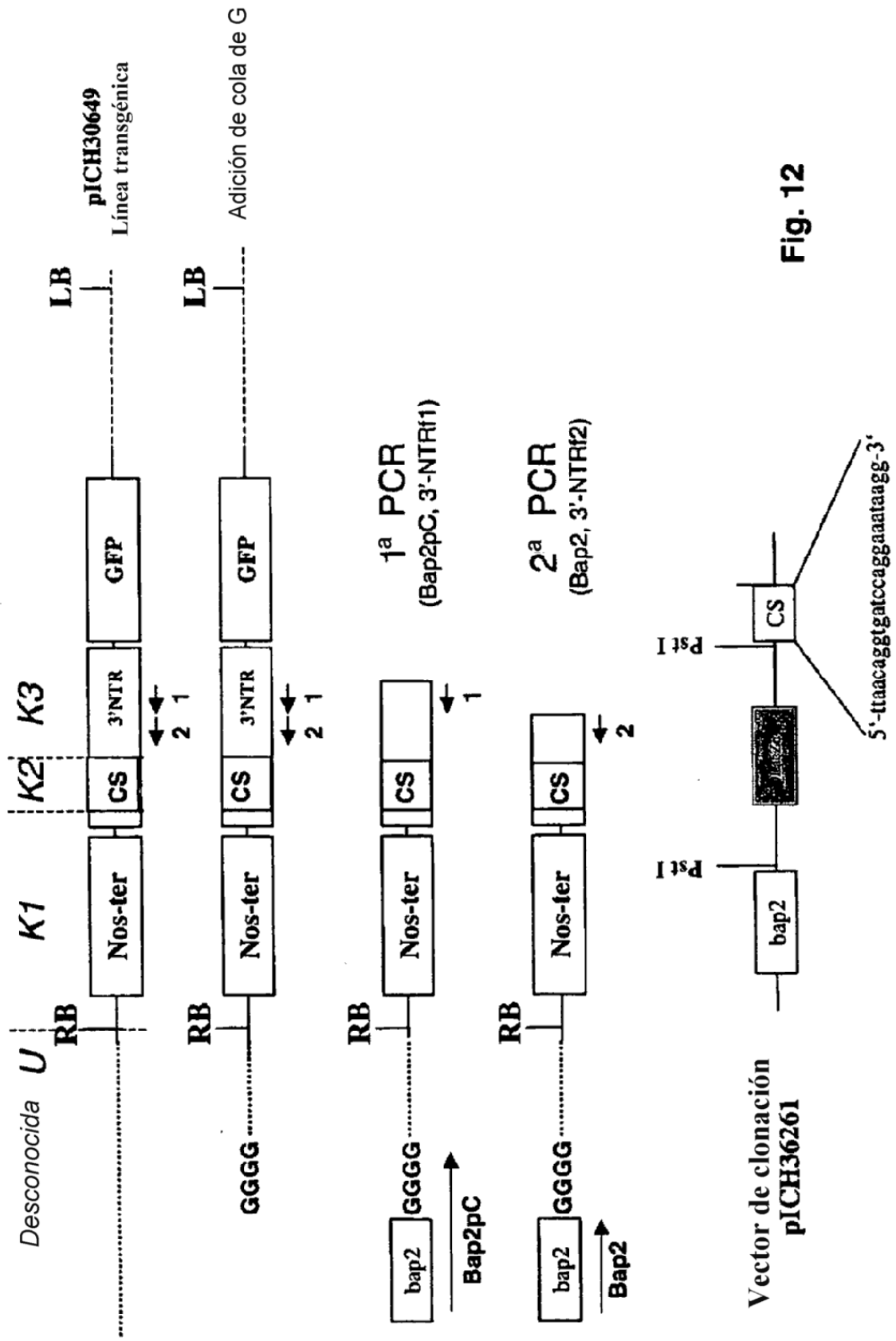


Fig. 12

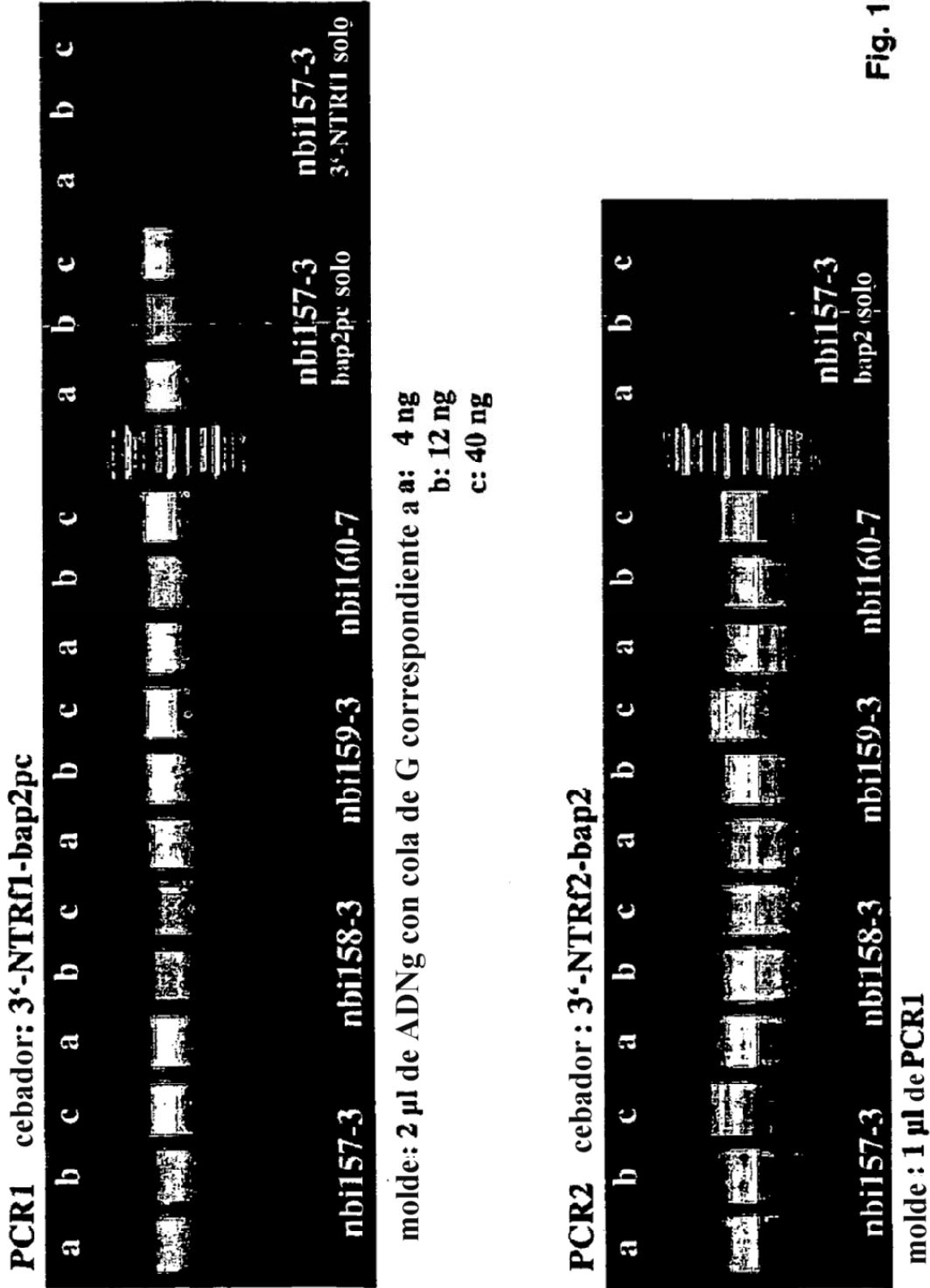


Fig. 13

E indica
vector vacío;
Z colonias azules
(que contienen *lacZ*)

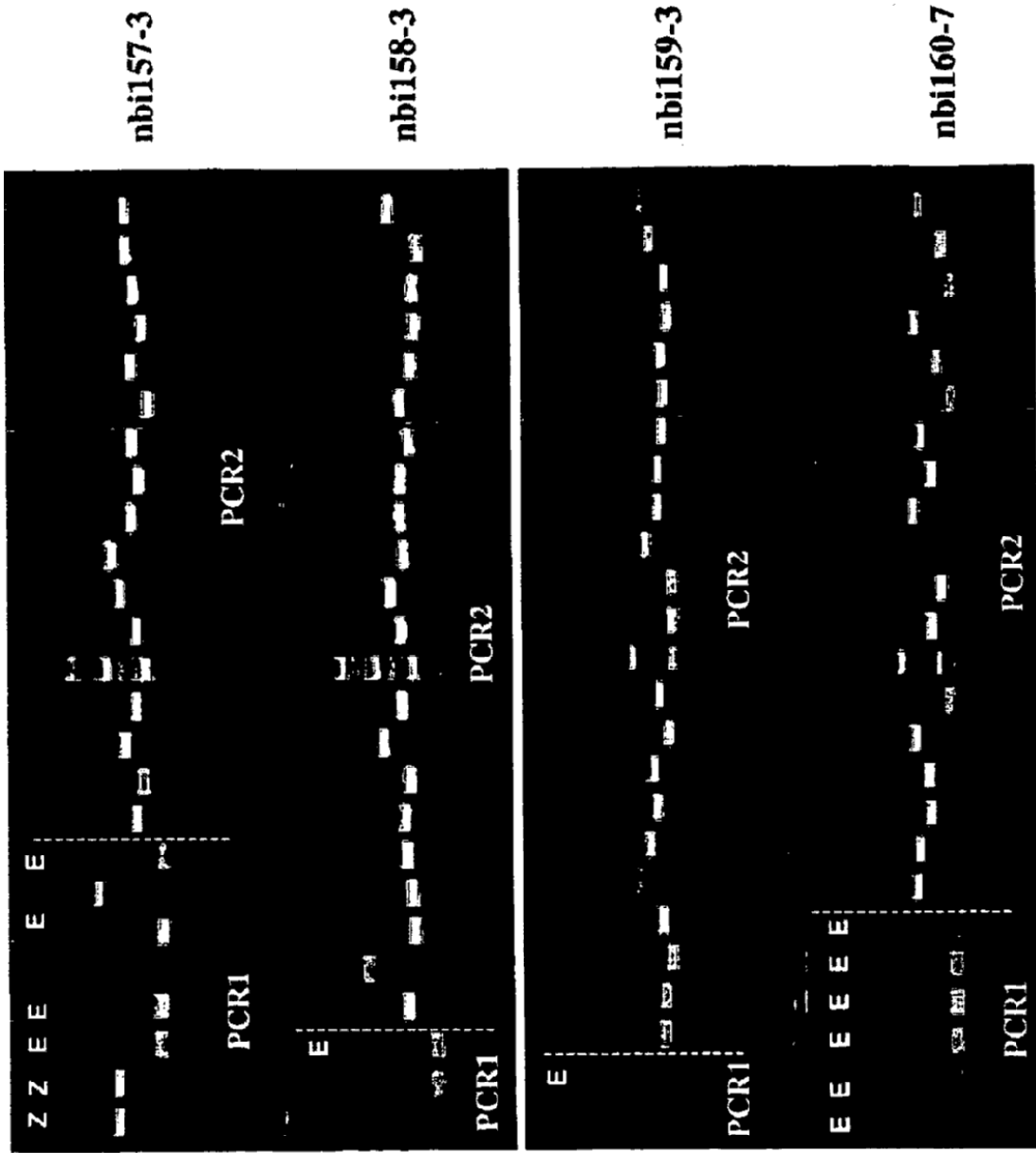


Fig. 14

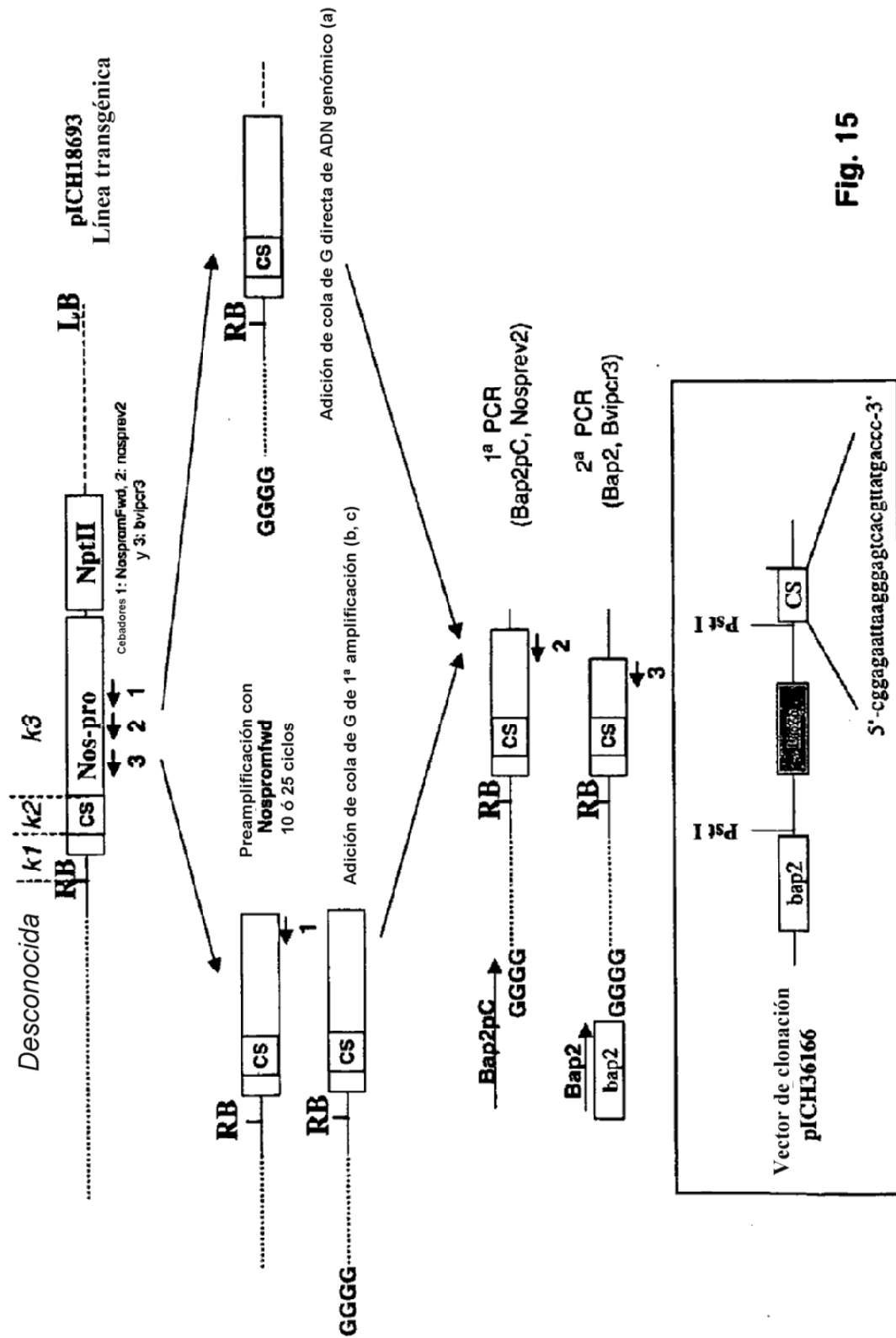


Fig. 15

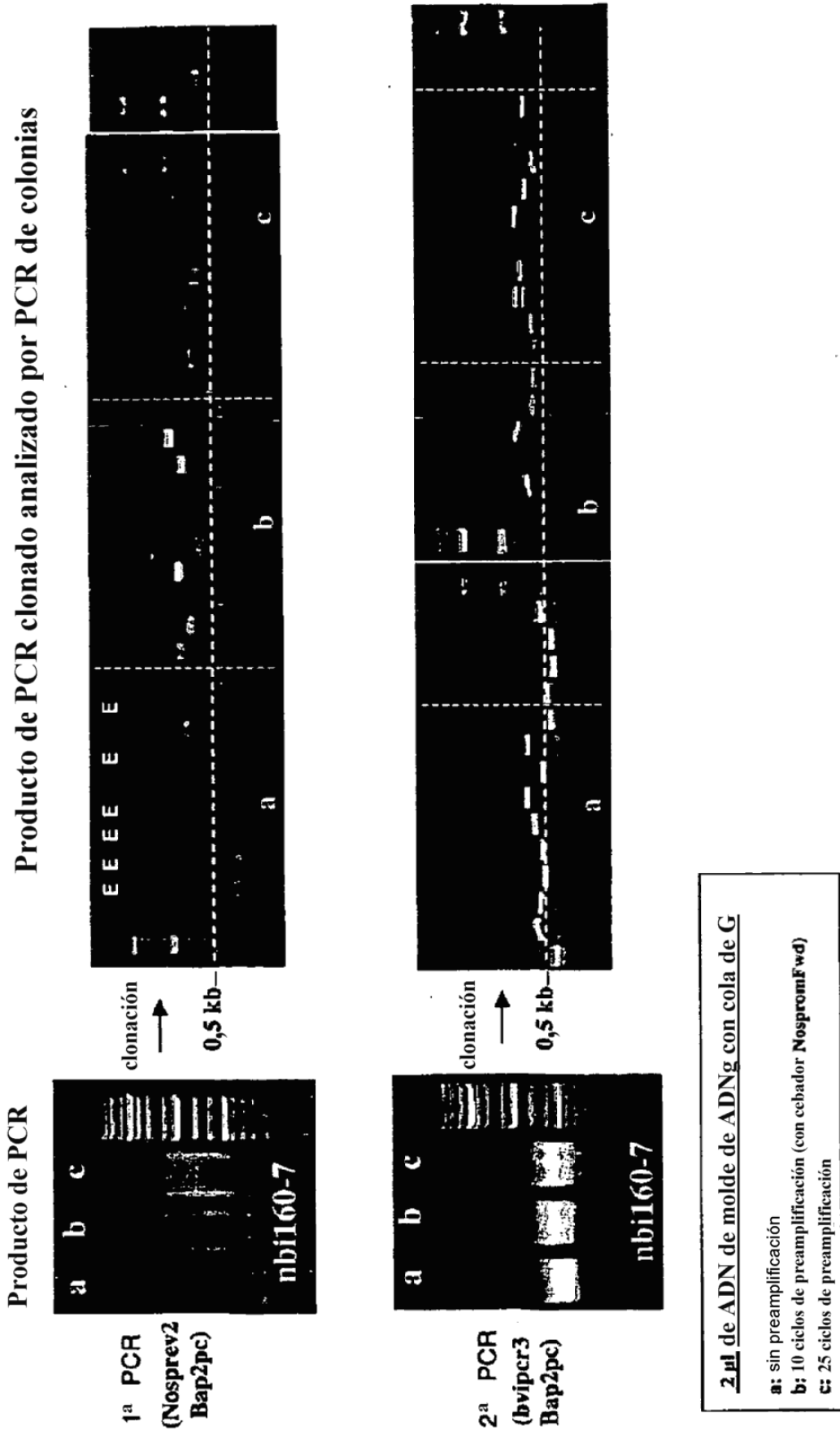


Fig. 16

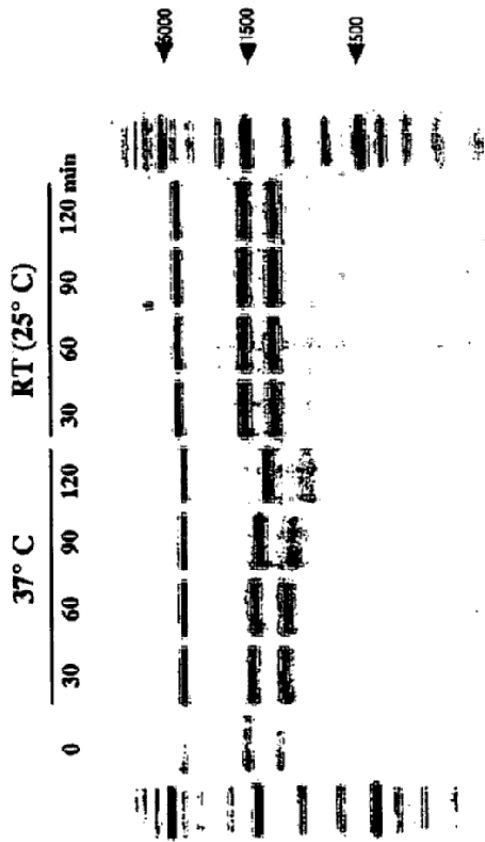
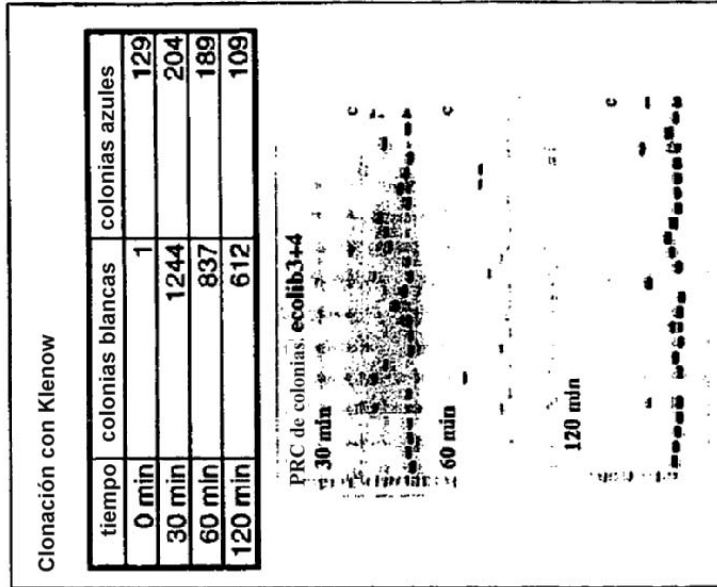


Fig. 17