

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 740 109**

51 Int. Cl.:

**C12Q 1/68** (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **27.01.2017 PCT/EP2017/051793**

87 Fecha y número de publicación internacional: **03.08.2017 WO17129756**

96 Fecha de presentación y número de la solicitud europea: **27.01.2017 E 17702824 (8)**

97 Fecha y número de publicación de la concesión europea: **15.05.2019 EP 3408408**

54 Título: **Amplificación selectiva de las regiones de ácido nucleico deseadas en una secuencia diana**

30 Prioridad:

**27.01.2016 EP 16152951**

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

**05.02.2020**

73 Titular/es:

**DEVYSER HOLDING AB (100.0%)  
Instrumentvägen 19  
126 53 Hägersten, SE**

72 Inventor/es:

**KLANGBY, ULF;  
HEDRUM, ANDERS y  
HEIM, STEFFEN**

74 Agente/Representante:

**PADIAL MARTÍNEZ, Ana Belén**

ES 2 740 109 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

Amplificación selectiva de las regiones de ácido nucleico deseadas en una secuencia diana.

Campo de la divulgación

5 La presente divulgación está en el campo del análisis de ácidos nucleicos, en particular en relación con las reacciones en cadena de la polimerasa y la posterior secuenciación masiva en paralelo.

Antecedentes

10 Durante varias décadas, la secuenciación de dideoxi-ADN (secuenciación de "Sanger", Sanger et al (1977), Proc Natl Acad Sci USA 74: 5463-5467) se ha usado como la tecnología de secuenciación estándar en la mayoría de los laboratorios genéticos. La productividad es limitada cuando se usa la secuenciación de Sanger, puesto que el proceso se realiza en moléculas individuales de la secuencia diana. La búsqueda de un procedimiento de secuenciación más rápido, más barato y más preciso, que añada la ventaja de la alta productividad, ha conducido al desarrollo de varias técnicas nuevas actualmente designadas "secuenciación paralela masiva" (MPS) o "secuenciación de nueva generación" (NGS).

15 La introducción de procedimientos de MPS ha cambiado el paradigma de la secuenciación del ADN (Mardis (2008), Trends Genet 24: 133-141; Shendure y Ji (2008), Nat. Biotechnol 26: 1135-1145). Los procedimientos de MPS posibilitan el procesamiento paralelo de cientos de miles a millones de moldes de ADN en paralelo, lo que da como resultado una alta productividad y un bajo coste por base de la secuencia generada. La introducción de procedimientos de MPS además permite al usuario secuenciar genomas complejos completos tales como el genoma humano.

20 No obstante, la secuenciación rutinaria a gran escala de todo el complejo genoma humano en su totalidad aún no es factible para uso diagnóstico debido a que el coste y el tiempo aún son demasiado grandes. La secuenciación rutinaria del genoma humano para uso diagnóstico requiere una cobertura de cada nucleótido de al menos 100-1000 veces, lo que da como resultado la necesidad de secuenciar y procesar 300-3000 Gb de datos de secuenciación por paciente. Con los procedimientos de MPS disponibles, esto requeriría varias tandas de secuenciación/paciente y costaría decenas de miles de dólares. Además de la carga económica, requeriría un procesamiento masivo de datos y capacidades de almacenamiento de datos que supondrían una carga sustancial para la infraestructura informática de un laboratorio genético.

25 Como consecuencia, se han desarrollado varios procedimientos para simplificar el proceso y enriquecer las regiones de interés para, p. ej., la secuenciación, es decir, procedimientos de enriquecimiento de diana. Los procedimientos de enriquecimiento de diana se utilizan para definir regiones genómicas que se pueden capturar y enriquecer de forma selectiva a partir de una muestra de ADN antes de la secuenciación. Resequenciar solo aquellas regiones genómicas que están enriquecidas es mucho más rápido y rentable que la secuenciación del genoma completo, y los datos resultantes son considerablemente menos complejos de analizar y requieren menos capacidades de almacenamiento de datos.

30 Los procedimientos de enriquecimiento de diana conocidos incluyen sondas de inversión molecular (MIP), captura de híbridos en matriz y en solución y reacción en cadena de la polimerasa (PCR). Cada uno se discute por separado a continuación.

35 *Sondas de inversión molecular (MIP):* Las MIP son sondas candado modificadas. Cuando una sonda se hibrida con una diana genómica correspondiente, hay un hueco en una o más posiciones de nucleótidos. Un diseño donde la sonda carece de un nucleótido complementario, p. ej., en una localización de polimorfismo de un solo nucleótido (SNP), se puede usar para la detección e identificación del polimorfismo. Si el hueco de la sonda hibridada es más de un solo nucleótido, las sondas en general se denominan "sondas de inversión de conector" (CIP). La ventaja de usar este tipo de sondas es que se pueden usar para el genotipado de SNP, el análisis de variación del número de copias (CNV) y para la detección de desequilibrios alélicos. Las sondas pueden diseñarse de tal manera que contengan secuencias marcadoras para la identificación de la sonda, así como secuencias cebadoras para la secuenciación de la región diana. En comparación con otras técnicas de enriquecimiento de diana, las sondas de inversión molecular demuestran una buena especificidad, pero pueden mostrar cierta variabilidad en el rendimiento entre diferentes sondas dentro de los ensayos.

40 *Captura de híbrido en matriz y en solución:* Si se está interesado en capturar regiones genómicas utilizando la técnica de captura en solución, pueden usarse una serie de oligonucleótidos (sondas) e hibridarlos con ADN genómico fragmentado en solución (a diferencia de los procedimientos híbridos o basados en matrices). Las sondas, que pueden enlazarse con perlas paramagnéticas, hibridan con el ADN de interés. Después de la hibridación, las perlas que contienen sondas y fragmentos de ADN complementarios se pueden separar, y el material de ADN no unido se retira lavando las perlas. Después de la extracción de las perlas, el ADN puede secuenciarse usando la secuenciación de Sanger o procedimientos de MPS. Una ventaja de esta técnica en comparación con las técnicas basadas en matrices es el enriquecimiento de diana mejorado debido a la alta relación de sonda/diana. Sin embargo, esto puede tener implicaciones en los intentos de reducir los costes de MPS.

45 *Reacción en cadena de la polimerasa (PCR):* La PCR se ha usado ampliamente para la preparación de muestras

previa a la secuenciación (Saiki et al (1988), Science 239: 487-491), ya que es muy compatible con un enfoque tradicional basado en la secuenciación de Sanger. La PCR también es compatible con cualquier plataforma de MPS actual. Sin embargo, con el fin de aprovechar al máximo la alta productividad posibilitada por la tecnología de MPS, se deben procesar y secuenciar conjuntamente un gran número de productos de amplificación por PCR ("amplicones").

5 La PCR multiplexada permite al usuario generar múltiples amplicones de PCR diferentes a partir de una única reacción de PCR, y es particularmente útil durante el enriquecimiento de diana para MPS. La PCR multiplexada puede ser difícil de realizar, porque el uso simultáneo de múltiples pares de cebadores a menudo genera un alto nivel de amplificación no específica, causada por una interacción entre los cebadores (Cho et al (1999), Nat Genet 23: 203-207; Wang et al (1998), Science 280: 1077-1082). Se han desarrollado diversos procedimientos para superar la amplificación no

10 específica en la PCR multiplexada (Fredriksson y otros (2007), Nucleic Acids Res 35: e47; Meuzelaar y otros (2007), Nat Methods 4: 835-837; Varley y Mitra (2008), Genome Res 18: 1844-1850; patente de EE. UU. 5677152).

Para muchas plataformas de MPS, hay un límite superior a la longitud del fragmento de ADN que se puede secuenciar en una sola tanda. Por esta razón, muchos fragmentos de ADN deben dividirse en amplicones más cortos, cada uno representado por un par de cebadores específico, con el fin de obtener una secuencia continua de ADN después del

15 análisis de los datos. Con el fin de maximizar la productividad, a menudo también es necesario no usar amplicones de PCR demasiado cortos, ya que estos reducirán el número total de lecturas de base potenciales en una tanda de secuenciación.

Con el fin de obtener una cobertura de secuenciación completa de una región definida dentro del genoma humano, los amplicones pueden diseñarse para superponerse usando un diseño de solapamiento de cebadores. Con el fin de

20 obtener una amplificación por PCR eficiente y longitudes de amplicón óptimas, y para evitar una amplificación no específica extensa, actualmente es necesario dividir las reacciones de PCR multiplexadas en varias reacciones separadas cuando se usa un diseño de superposición de cebadores. De esta manera, los diseños de superposición reales se separan en diferentes reacciones de PCR. Si una muestra clínica necesita dividirse en varias reacciones de PCR multiplexadas diferentes, aumenta el riesgo de confusión de la muestra y contaminación de la muestra. El uso

25 de múltiples reacciones de PCR para el análisis de una muestra clínica también puede ser un problema cuando la cantidad de ADN disponible para el análisis es limitada.

La PCR digital se ha propuesto como un procedimiento adecuado para la preparación de colecciones de MPS. La PCR digital se basa en la amplificación clonal de ácidos nucleicos y requiere un equipo altamente especializado con el fin de que se realice de manera eficiente. La PCR digital también se considera propensa a errores en manos de

30 usuarios inexpertos (Sykes et al (1992), BioTechniques 13 (3): 444-9; Perkel (2015), BioTechniques 58 (5): 217-21, Pekin et al (2011), Lab on a Chip 11 (13): 2156-66). La preparación de muestras mediante PCR digital requiere más ADN disponible para el análisis en comparación con una PCR de reacción única.

Existe una necesidad en la técnica de procedimientos que mejoren las reacciones de PCR multiplexadas en general. En particular, existe la necesidad de procedimientos que aborden el problema de la amplificación no específica y/o no deseada en las reacciones de PCR multiplexadas en varias etapas, por ejemplo, durante la amplificación de diana

35 para aplicaciones de diagnóstico genético, preparación de colecciones, resecuenciación y otras situaciones.

#### Sumario de la divulgación

Un objeto de la divulgación es abordar esta necesidad posibilitando la amplificación selectiva de las regiones de ácido nucleico deseadas en una secuencia diana.

40 Otro objetivo es proporcionar un procedimiento que introduzca diferencias en las propiedades de autohibridación de los fragmentos de ácido nucleico como una herramienta para seleccionar entre los productos de amplificación deseados y no deseados.

Otro objetivo es proporcionar un procedimiento que permita la amplificación selectiva por PCR de amplicones cortos superpuestos en una sola reacción.

45 Otro objetivo es proporcionar un procedimiento que proporcione amplicones cortos superpuestos que cubran largos tramos continuos de ADN en una sola reacción.

Estos objetos, y otros que son evidentes para la persona experta a partir de las enseñanzas de este documento, se satisfacen por los diferentes aspectos de la divulgación como se definen en el presente documento y en las reivindicaciones adjuntas.

50 Por tanto, en un primer aspecto, la divulgación proporciona un procedimiento para la amplificación selectiva de los amplicones deseados en una reacción en cadena de la polimerasa en dos etapas, que comprende las etapas de:

- proporcionar un material genético diana;
- desnaturalizar el material genético diana;
- añadir un conjunto de cebadores diana en condiciones que permitan la reasociación de cebadores con material

genético diana;

\* comprendiendo cada cebador en dicho conjunto de cebadores diana

- ◆ una secuencia específica de diana,
- ◆ una secuencia de control de desemparejamiento, y
- ◆ una secuencia de hibridación genérica;

\* siendo dicha secuencia de hibridación genérica idéntica en cada cebador diana;

\* estando dicha secuencia de control de desemparejamiento y secuencia de hibridación genérica próximas entre sí; y

\* comprendiendo dicho conjunto de cebadores diana al menos dos pares de cebadores adaptados para la amplificación de al menos dos amplicones deseados superpuestos en el material genético diana, de tal modo que existe la posibilidad de obtener al menos un amplicón no deseado en la región de superposición;

\* estando caracterizado además dicho conjunto de cebadores diana porque cualquier par de cebadores que flanquean una secuencia de amplicón deseada tiene secuencias de control de desemparejamiento desemparejadas, mientras que cualquier par de cebadores que flanquean una secuencia de amplicón no deseada tiene secuencias de control de desemparejamiento emparejadas;

• llevar a cabo una primera reacción en cadena de la polimerasa multiplexada, dando como resultado amplicones superpuestos correspondientes a diferentes combinaciones posibles de pares de cebadores en el primer conjunto de cebadores, manteniéndose dichos amplicones superpuestos en condiciones tales que se autohibridan a través del acoplamiento de pares de bases entre las secuencias complementarias resultantes de la secuencia de hibridación genérica en cada cebador;

• tratar dichos amplicones superpuestos en condiciones de selección que son tales que los amplicones deseados se desnaturalizan debido a secuencias no complementarias resultantes de las secuencias de control de desemparejamiento desemparejadas, mientras que los amplicones no deseados permanecen autohibridados debido a secuencias complementarias adicionales resultantes de las secuencias de control de desemparejamiento emparejadas;

• añadir cebadores genéricos, cada uno de los cuales comprende una secuencia que es complementaria de la secuencia de hibridación genérica;

• llevar a cabo una segunda reacción en cadena de la polimerasa, que da como resultado la amplificación selectiva de los amplicones deseados.

En un segundo aspecto, la divulgación proporciona un kit para llevar a cabo el procedimiento, que comprende los reactivos clave necesarios y las instrucciones para poner el procedimiento en práctica. En este aspecto, se proporciona un kit para llevar a cabo el procedimiento según el primer aspecto, que comprende:

• un conjunto de cebadores diana;

\* comprendiendo cada cebador en dicho conjunto de cebadores diana

- ◆ una secuencia específica de diana,
- ◆ una secuencia de control de desemparejamiento, y
- ◆ una secuencia de hibridación genérica;

\* siendo dicha secuencia de hibridación genérica idéntica en cada cebador diana;

\* estando dicha secuencia de control de desemparejamiento y secuencia de hibridación genérica próximas entre sí; y

\* comprendiendo dicho conjunto de cebadores diana al menos dos pares de cebadores adaptados para la amplificación de al menos dos amplicones deseados superpuestos en un material genético diana, de tal modo que existe la posibilidad de obtener al menos un amplicón no deseado en la región de superposición;

\* estando caracterizado además dicho conjunto de cebadores diana porque cualquier par de cebadores que flanquean una secuencia de amplicón deseada tiene secuencias de control de desemparejamiento desemparejadas, mientras que cualquier par de cebadores que flanquean una secuencia de amplicón no deseada tiene secuencias de control de desemparejamiento emparejadas;

- cebadores genéricos, cada uno de los cuales comprende una secuencia que es complementaria de la secuencia de hibridación genérica; e
- instrucciones para llevar a cabo el procedimiento según el primer aspecto usando dichos cebadores.

5 Por tanto, la divulgación se refiere a situaciones en las que se lleva a cabo una primera reacción de PCR, y los fragmentos amplificados resultantes, o amplicones, se usan como moldes en una segunda reacción de PCR. Específicamente, dicha primera reacción de PCR es tal que tiene la posibilidad de generar al menos dos amplicones alternativos, uno de los cuales es menos deseado que el otro con respecto al uso como molde para la segunda reacción. Una de las ideas subyacentes de esta divulgación es que es posible promover el uso del amplicón deseado como molde y suprimir el uso del amplicón no deseado. Esto se hace potenciando la tendencia intrínseca de los amplicones de la primera reacción de PCR a autohibridar en regiones de complementariedad de secuencia. Todos los amplicones de la primera reacción de PCR tendrán secuencias complementarias debido a la presencia de una secuencia de hibridación genérica (que se usa para la reasociación de cebadores en la segunda PCR). En condiciones normales de reacción de PCR, esto conducirá a que los amplicones formen estructuras de tallo-bucle después de la primera reacción de PCR. Con el fin de que tales estructuras de tallo-bucle sirvan como moldes para una segunda reacción de PCR, el tramo de secuencia autohibridado necesita desnaturalizarse con el fin de permitir la reasociación de cebadores genéricos en la misma secuencia. En condiciones normales, tal desnaturalización ocurriría en la misma medida, independientemente de si el amplicón es deseado o no deseado como molde. Esto es especialmente así en caso de que el recipiente de reacción ya contenga los cebadores genéricos, porque estos competirán por el apareamiento de bases con la secuencia genérica. La divulgación del presente documento permite ajustar la fuerza de la autohibridación en las estructuras de tallo-bucle que se forman por los amplicones de la primera reacción de PCR, asegurando que los amplicones no deseados forman estructuras de tallo-bucle que son menos fácilmente desnaturalizadas que las correspondientes estructuras de tallo-bucle formadas por amplicones deseados. Esto se hace insertando, en los cebadores para la primera reacción de PCR, una secuencia relativamente corta diseñada para conferir una autocomplementariedad extra en los amplicones no deseados, mientras que no se proporciona tal autocomplementariedad adicional a los amplicones deseados. Esta secuencia corta se designa "secuencia de control de desemparejamiento" en la presente divulgación. Las regiones de la secuencia diana que no se desean como resultado final después de la segunda PCR están flanqueadas por cebadores (para la primera PCR) que contienen secuencias de control de desemparejamiento emparejadas, porque tales secuencias emparejadas potenciarán la autohibridación de los amplicones resultantes y los harán menos accesibles a la reasociación de cebadores al comienzo de la segunda PCR. A la inversa, las regiones de secuencia diana que se desean como resultado final, p. ej., para secuenciación, detección o cualquier otro propósito, están flanqueadas por cebadores de la primera PCR que contienen secuencias de control de desemparejamiento desemparejadas. Una realización del procedimiento divulgado se presenta esquemáticamente en la Figura 9.

35 El experto en la materia se dará cuenta de que es posible beneficiarse del principio general de la divulgación en varias configuraciones diferentes, todas las cuales se pretenden englobadas en las reivindicaciones adjuntas. Esto se aplica, por ejemplo, al orden de los pasos en el procedimiento. Es posible llevar a cabo las dos etapas de PCR en diferentes recipientes de reacción, pero puede ser igualmente preferible llevarlas a cabo en el mismo recipiente de reacción. En una realización, los cebadores genéricos (es decir, los cebadores para la segunda reacción de PCR) se añaden después de la primera reacción de PCR. En otra realización, en la que las reacciones se llevan a cabo en el mismo recipiente de reacción, los cebadores genéricos se añaden antes de la primera reacción de PCR, por ejemplo, al mismo tiempo que se añaden los cebadores diana (es decir, los cebadores para la primera reacción de PCR).

45 El mismo razonamiento se aplica a las condiciones de reacción de las reacciones de PCR. El experto en la materia puede adaptarlos de manera tal que estén presentes las condiciones apropiadas de reasociación y desnaturalización para los pasos del procedimiento divulgado. En particular, las condiciones al comienzo de la segunda reacción de PCR necesitan ser tales que la separación de la hebra (desnaturalización) y la reasociación del cebador genérico sean posibles en los amplicones con secuencias de control de desemparejamiento desemparejadas, es decir, aquellas que tienen un tramo autocomplementario más corto de secuencia, mientras que la reasociación del cebador genérico se inhibe en los amplicones con secuencias de control de desemparejamiento emparejadas, es decir, aquellos que tienen un tramo autocomplementario más largo de secuencia. La calibración de tales condiciones, con respecto a factores tales como temperaturas para desnaturalización, reasociación y temperaturas de extensión, así como la concentración de aditivos tales como iones  $Mg^{2+}$ , DMSO, betaína y formamida, dependería de las circunstancias y fines específicos en cada caso, pero encontrarlos estaría dentro de las capacidades de una persona experta en la técnica establecida de tecnología de PCR (véase para una vista general de la tecnología de PCR, por ejemplo, Bartlett y Stirling (2003), PCR Protocols). Methods in Molecular Biology 226: 3-6, o Sambrook y Russel (2001), Molecular Cloning: A Laboratory Manual (3ª ed.). Cold Spring Harbor, N.Y.: Cold Spring Harbor Laboratory Press. ISBN 0-879-69576-5. Capítulo 8: "In vitro Amplification of DNA by the Polymerase Chain Reaction").

60 Con respecto a los cebadores diana, estos comprenden tres tramos diferentes de secuencia con diferentes funciones en el contexto de la divulgación. Estos tres pueden ser distintos y seguir uno a otro, ya sea directamente o próximos entre sí. Por tanto, la secuencia específica de diana puede estar seguida por la secuencia de control de desemparejamiento, que a su vez entonces está seguida por la secuencia de hibridación genérica. En algunas realizaciones, sin embargo, estos tres tramos de secuencia definidos funcionalmente pueden superponerse. Este es, por ejemplo, el caso si una o más bases de la secuencia de control de desemparejamiento también son

complementarias de la secuencia diana, de modo que uno de los extremos de la secuencia específica de diana cumple la función de control de desemparejamiento así como la unión de diana. También es posible que el control de desemparejamiento esté proporcionado por desemparejamientos introducidos por cualquiera o ambas de la secuencia específica de diana y la secuencia de hibridación genérica, lo que significa que los tres elementos de secuencia diferentes están realmente total o parcialmente intercalados a lo largo del oligonucleótido cebador diana. La definición precisa de un cebador diana para uso en el procedimiento y el kit divulgados no es crítica, siempre que permita una fuerza de autohibridación diferente dependiendo de si la secuencia diana flanqueada es deseada o no deseada para amplificación. Para los propósitos de la siguiente discusión sobre la longitud de las diferentes secuencias, es más fácil considerar estas como segmentos de secuencia distintos, pero, de nuevo, también es posible usar cebadores en los que la función de control de desemparejamiento está entremezclada dentro de las secuencias específicas de diana o de hibridación genérica.

En una realización, la longitud total de un cebador diana en el conjunto es de aproximadamente 7 a aproximadamente 200 nucleótidos, siempre a condición de que estén presentes las funciones necesarias de unión a diana, control de desemparejamiento e hibridación genérica.

En una realización en la que la secuencia de control de desemparejamiento es una secuencia distinta, la longitud de la secuencia de control de desemparejamiento es de 1 a 7 nucleótidos, tal como de uno a cinco, de uno a tres o por ejemplo de uno a dos nucleótidos. La longitud no es crítica, pero depende de otros factores, en particular las condiciones de reasociación y desnaturalización en las etapas iniciales de la segunda reacción de PCR.

La longitud de la secuencia de hibridación genérica tampoco es crítica, siempre que permita la reasociación y cebado por los cebadores genéricos correspondientes en la segunda reacción de PCR. En una realización, la secuencia de hibridación genérica tiene una longitud de 1 a 75 nucleótidos, por ejemplo de 3 a 75 nucleótidos, de 5 a 60 nucleótidos, de 5 a 40 nucleótidos o de 5 a 20 nucleótidos, tal como de 10 a 20 o de 7 a 15 nucleótidos. El experto en la materia puede diseñar secuencias de hibridación genéricas adecuadas basándose, p. ej., en la fuerza de apareamiento de bases, etc., en el que el uso de un apareamiento de bases más fuerte podría permitir una secuencia de hibridación más corta, mientras que el uso de un apareamiento de bases más débil podría permitir una secuencia de hibridación más larga.

Con respecto a la secuencia específica de diana, esta necesita ser lo suficientemente larga para asegurar la hibridación específica en la región deseada del ácido nucleico diana. De nuevo, el experto en la materia es consciente de las consideraciones de diseño implicadas, y puede diseñar tales cebadores sin una experimentación indebida. En una realización, la secuencia específica de diana tiene una longitud de 5 a 100 nucleótidos, por ejemplo de 8 a 70 nucleótidos, de 10 a 50 nucleótidos o de 10 a 40 nucleótidos, tal como de 15 a 30 nucleótidos, por ejemplo de 18 a 25 nucleótidos, por ejemplo de 20 a 25 nucleótidos.

Típicamente, pero no siempre, la secuencia específica de diana es complementaria de una secuencia de diagnóstico en el material genético diana, por lo que la detección o secuenciación de esta secuencia es de interés diagnóstico o clínico.

En una realización, al menos algunos de los cebadores implicados en las diferentes reacciones están marcados de forma detectable, por ejemplo, con el fin de seguir las reacciones y/o detectar los productos finales. El experto en la materia es consciente de las diferentes opciones para marcar moléculas de nucleótidos, p. ej., con marcadores fluorescentes, y puede hacerlo sin una experimentación indebida.

En una realización, el procedimiento divulgado se usa en la preparación de una colección de secuencias a partir de un material genético diana que consiste en ADN genómico para secuenciar. Esta colección de secuencias puede usarse posteriormente para la secuenciación de ácidos nucleicos. En tales aplicaciones, los cebadores genéricos, usados en la segunda reacción de PCR del procedimiento divulgado y contenidos en el kit divulgado, pueden comprender además secuencias adaptadoras para el procedimiento de secuenciación deseado usado. Además, en una de tales realizaciones, el procedimiento divulgado comprende además el paso de secuenciar los amplicones deseados resultantes de las dos etapas de la amplificación por PCR. En una realización específica, dicha secuenciación se lleva a cabo usando una secuenciación paralela masivamente. Adecuadamente, en una realización, los cebadores genéricos también comprenden secuencias de índice para que los productos de amplicón puedan identificarse si se agrupan con amplicones de reacciones separadas realizadas en otro material genético diana o en muestras de diferentes individuos. Por tanto, en estas realizaciones, es posible usar el procedimiento y el kit divulgados para obtener una recopilación de amplicones específicos de diana específicos de paciente, marcados en diana, indexados, purificados y estandarizados en forma de una colección de dianas específicas de paciente, que se pueden agrupar entonces junto con colecciones específicas de diana de otros pacientes o analizar solo antes de realizar la secuenciación, p. ej., una secuenciación paralela masiva. Se puede hacer un agrupamiento de colecciones a partir de varias muestras de pacientes para la misma diana o múltiples dianas para el mismo paciente o una combinación de ambos, siempre y cuando las secuencias de indexación usadas en la segunda reacción de PCR sean únicas para cada paciente.

Si bien la invención se ha descrito con referencia a diversos aspectos y realizaciones ejemplares, los expertos en la técnica entenderán que pueden realizarse diversos cambios y pueden sustituirse por equivalentes elementos de la

misma sin apartarse del alcance de la divulgación del presente documento. Además, pueden hacerse muchas modificaciones para adaptar una situación particular a las enseñanzas de la invención sin apartarse del alcance esencial de la misma. Por lo tanto, se pretende que la invención no se limite a ninguna realización particular contemplada, sino que incluya todas las realizaciones que estén dentro del alcance de las reivindicaciones adjuntas.

5 Breve descripción de las figuras

La Figura 1 es una ilustración esquemática de los constructos de cebadores divulgados en el presente documento y usados en diversas realizaciones del procedimiento.

La Figura 2 es una ilustración esquemática de los pares de cebadores usados en el Ejemplo 1, así como sus posiciones en relación con la secuencia diana, el exón 14 del gen *CFTR* humano.

10 La Figura 3 muestra los cromatogramas de la detección de A) los fragmentos de PCR de TT y B) los fragmentos de PCR de ID, usando electroforesis capilar después del uso del par de cebadores TT-EX14\_02F/TT-EX14\_02R, seguido del par de cebadores ID-lon P1 e ID-lon A.

15 La Figura 4 muestra los cromatogramas de la detección de A) los fragmentos de PCR de TT y B) los fragmentos de PCR de ID, usando electroforesis capilar después del uso del par de cebadores TT-EX14\_01F/TT-EX14\_02R, seguido del par de cebadores ID-lon P1 e ID-lon A.

La Figura 5 muestra los cromatogramas de la detección de A) los fragmentos de PCR de TT y B) los fragmentos de PCR de ID, usando electroforesis capilar después del uso del par de cebadores TT-EX14\_02F/TT-EX14\_03R, seguido del par de cebadores ID-lon P1 e ID-lon A.

20 La Figura 6 es una ilustración esquemática de estructuras de autorreasociación de productos de PCR de TT amplificadas por A) el par de cebadores TT-EX14\_02F/TT-EX14\_02R y B) el par de cebadores TT-EX14\_02F/TT-EX14\_03R. La secuencia de hibridación genérica de 13 pb se indica en negrita. La secuencia de control de desemparejamiento de 2 pb está subrayada.

La Figura 7 es una ilustración esquemática del uso de cebadores de PCR superpuestos en la amplificación de un tramo de ADN genómico diana.

25 La Figura 8 es una ilustración esquemática de uno de los problemas superados por el presente procedimiento. Se produce una multitud de amplicones de PCR no deseados durante una segunda etapa de PCR cuando se usan cebadores superpuestos sin secuencias de control de desemparejamiento durante una primera etapa de PCR.

30 La Figura 9 es una ilustración esquemática resumida del procedimiento divulgado. Los amplicones generados durante la PCR de TT y que tienen secuencias de control de desemparejamiento emparejadas en ambos lados (1-1 o 2-2) se representan encima de la secuencia genética diana, y generan las estructuras de bucle autocomplementarias más largas representadas en la esquina inferior izquierda durante la PCR de TT. Los amplicones que tienen secuencias de control de desemparejamiento desemparejadas (1-2, 2-1) se representan debajo de la secuencia genética diana, y generan las estructuras de bucle autocomplementarias más cortas representadas en la esquina inferior derecha durante la PCR de TT.

35 La Figura 10 es una ilustración esquemática del resultado del presente procedimiento en el problema ilustrado en la Figura 8. La inhibición de los amplicones de PCR de ID cortos no deseados durante la PCR de ID se indica mediante X.

40 La Figura 11 muestra los cromatogramas del análisis de electroforesis capilar después de A) PCR multiplexada de TT (2 carriles superiores) con un conjunto de cebadores de PCR de TT sin secuencias de control de desemparejamiento, y posterior PCR de ID (carril inferior) que muestra una multitud de productos de PCR de ID cortos no deseados; y B) PCR multiplexada de TT (2 carriles superiores) con un conjunto de cebadores de PCR de TT que incluye secuencias de control de desemparejamiento (2 carriles superiores), y la posterior PCR de ID (carril inferior) que no muestra productos de PCR de ID no deseados.

45 La Figura 12 es una ilustración esquemática de los posibles productos de la amplificación por PCR multiplexada de un tramo de ADN diana, que muestra el producto corto no deseado eliminado por el procedimiento divulgado en el presente documento, los amplicones más largos deseados y los amplicones obtenidos tras translectura.

La Figura 13 ilustra esquemáticamente las regiones de los productos de PCR que se cuentan al calcular los parámetros de cobertura de amplicón (CA) y cobertura de superposición (CO) en el procedimiento descrito en el Ejemplo 3.

Ejemplos

50 Los siguientes ejemplos ilustran realizaciones de la divulgación como se pusieron en práctica en diferentes entornos.

Ejemplo 1

Establecimiento de los principios básicos del procedimiento divulgado

*Materiales y procedimientos*

5 Este ejemplo implicaba llevar a cabo dos experimentos separados, cada uno de los cuales comprende una primera reacción de PCR de marcaje de diana (TT) realizada en ADN genómico humano purificado. La PCR de TT fue seguida por una segunda etapa de reacción de PCR de indexación (ID). Los amplicones de PCR generados en la reacción de PCR de TT se diluyeron 1:50 en agua antes de que se realizara la PCR de ID.

10 Los cebadores de PCR de TT se diseñaron para incluir una parte específica de diana en 3' de aproximadamente 20-25 nucleótidos y una secuencia de hibridación genérica en el extremo 5' del oligonucleótido. Los cebadores de PCR de ID se diseñaron para incluir una secuencia de oligonucleótidos en el extremo 3' que fuera capaz de reasociar específicamente con la secuencia de hibridación genérica presente en todos los cebadores de PCR de TT. Los cebadores de ID incluían además adaptadores de secuenciación específicos y secuencias de índice para permitir MPS y reconocimiento de índice. En este ejemplo, las secuencias adaptadoras de secuenciación usadas fueron las secuencias adaptadoras "lon A" e "lon P1" diseñadas específicamente para permitir la secuenciación en el instrumento Thermo Fisher lon PGM MPS.

15 La Figura 1 representa los cebadores usados esquemáticamente, y la Figura 2 muestra, de nuevo esquemáticamente, su relación con la secuencia diana, el exón 14 del gen *CFTR*. Las secuencias de nucleótidos de los cebadores usados se presentan a continuación en la Tabla 1, en la que la parte de hibridación del oligonucleótido respectivo está en letra subrayada y negrita.

Tabla 1: Cebadores oligonucleotídicos para PCR multiplexada en dos etapas

SEQ ID NO:	Denominación	Secuencia	Marcaje 5'
1	TT-EX14_02F (P51)	<b><u>GCTCTTCCGATCT</u></b> GCGAGTCTGTCC TGAACCTGAT	-
2	TT-EX14_02R (P52)	<b><u>GCTCTTCCGATCT</u></b> GCCAGTTTCTT GAGATAACCTTCTTGA	FAM
3	TT-EX14_01F (P57)	<b><u>GCTCTTCCGATCT</u></b> TTCCATTGTGC AAAAGACTCCCTTA	-
4	TT-EX14_03R (P60)	<b><u>GCTCTTCCGATCT</u></b> GGAAGAGATA TGTCCATTGCAAAAAGAA	HEX
5	ID-lon P1 (P43)	CCACTACGCCTCCGCTTTCCTCTC TATGGGCAGTCGGTGAT <b><u>GCTCTTC</u></b> <b><u>CGATCT</u></b>	HEX
6	ID-lon A (P44)	CCATCTCATCCCTGCGTGTCTCCG ACTCACAGAAGGAACGAT <b><u>GCTCTT</u></b> <b><u>CCGATCT</u></b>	FAM

20 En el primer experimento, se usó la combinación de cebadores TT-EX14\_02F/TT-EX14\_02R, específica de secuencias dentro del exón 14 de *CFTR*, para la PCR de TT como reacción de PCR de primera etapa. Después de completar este programa de PCR, la solución resultante se diluyó 50 veces y se añadieron 2 µl de la dilución como molde en la siguiente reacción multiplexada de PCR de ID de segunda etapa. Se realizó la PCR de ID usando los cebadores ID-lon P1 e ID-lon A. Se analizaron ambas reacciones PCR de TT e ID mediante detección por electroforesis capilar de fragmentos de PCR amplificados marcados con fluorescencia, y los cromatogramas obtenidos se presentan como la Figura 3.

25 Se realizó un segundo experimento similar, pero usando en cambio las dos parejas de cebadores TT-EX14\_01F/TT-EX14\_02R y TT-EX14\_02F/TT-EX14\_03R en dos reacciones de PCR de TT de primera etapa separadas, simulando una PCR multiplexada con secuencias diana superpuestas. Usando estos dos pares de cebadores, y con referencia a la Figura 2, los dos amplicones deseados son las regiones flanqueadas por i) el cebador directo TT-EX14\_01F y el cebador inverso TT-EX14\_02R y ii) el cebador directo TT-EX14\_02F y el cebador inverso TT-EX14\_03R, respectivamente. En una verdadera reacción de PCR multiplexada, teóricamente también sería posible que la PCR de TT generara un amplicón no deseado en la región de superposición, es decir, la región flanqueada por el cebador directo TT-EX14\_02F y el cebador inverso TT-EX14\_02R, es decir, el mismo producto obtenido por estos dos cebadores en el primer experimento.

30 Una vez completadas las reacciones de PCR de TT, las soluciones resultantes se diluyeron 50 veces y se añadieron 2 µl de la dilución como molde a la reacción de PCR de ID de segunda etapa correspondiente. Se realizó la PCR de ID usando los cebadores ID-lon P1 e ID-lon A. Las reacciones de PCR de TT e ID se analizaron mediante detección por electroforesis capilar de fragmentos de PCR amplificados marcados con fluorescencia, y los cromatogramas se presentan como las Figuras 4 y 5.



*Resultados*

Para el primer experimento, la electroforesis capilar de la reacción PCR de TT que usa el conjunto de cebadores TT-EX14\_02F/TT-EX14\_02R daba como resultado la detección de un fragmento de PCR con la longitud esperada de 180 pb en el canal FAM (Figura 3A). Sin embargo, no se detectaron amplicones de PCR después de la reacción de PCR de ID de segunda etapa que usa el producto de PCR de TT resultante como molde (Figura 3B). La ausencia de amplicones de PCR detectables después de las reacciones combinadas de PCR de TT e ID indicaba que no había ocurrido ninguna amplificación durante la PCR de ID.

En el segundo experimento, donde se usaron los conjuntos de cebadores de PCR de TT TT-EX14\_01F/TT-EX14\_02R y TT-EX14\_02F/TT-EX14\_03R en reacciones de PCR de TT separadas, se pudieron observar fragmentos con las longitudes de fragmento esperadas en el canal FAM para la combinación de cebadores de TT TT-EX14\_01F/TT-EX14\_02R (Figura 4A) y en el canal HEX para la combinación de cebadores de TT TT-EX14\_02F/TT-EX14\_03R (Figura 5A).

Es importante que también pudieron detectarse dos fragmentos de PCR diferentes en ambos canales FAM y HEX después de la reacción de PCR de ID usando los cebadores ID-Ion P1 e ID-Ion A (Figuras 4B y 5B). La aparición de dos fragmentos de PCR en cada canal de detección después de la electroforesis capilar de las dos reacciones de PCR de ID está de acuerdo con los cuatro fragmentos de PCR esperados resultantes de un cebado genérico utilizando dos cebadores de PCR de ID diferentes (ID-Ion P1 e ID-Ion A), que cada uno individualmente puede reasociarse con ambos extremos de los amplicones de PCR de TT.

*Conclusiones*

Sin querer limitarse por ninguna teoría, se interpretó que la ausencia de un producto de PCR de ID después del uso del conjunto de cebadores TT-EX14\_02F/TT-EX14\_02R estaba causada por la formación de una estructura de autorreasociación de 15 pares de bases de largo (que incluye la secuencia de hibridación genérica y al menos dos bases complementarias adicionales) en los amplicones de PCR de TT. Bajo las condiciones de PCR usadas en este conjunto de experimentos, la estructura de autorreasociación obstruye la reasociación de los cebadores de PCR de ID con los amplicones de molde de PCR de TT durante la PCR de ID. A la inversa, cuando la longitud de la estructura de autorreasociación se limita a solo la secuencia de hibridación genérica de 13 pares de bases de largo en los amplicones de PCR de TT, es posible la reasociación de los cebadores de PCR de ID con la estructura de hibridación genérica. Para explicar con más detalle la falta de un producto de PCR de ID del primer experimento cuando la combinación de cebadores era TT-EX14\_02F/TT-EX14\_02R, se observa que estos cebadores de PCR de TT consisten en una parte específica de diana (subrayada) y una secuencia de hibridación genérica (negrita):

Cebador TT-EX14\_02F: **GCTCTTCGATCT** GCAGTCTGTCCTGAACCTGAT

Cebador TT-EX14\_02R: **GCTCTTCGATCT** GCCAGTTTCTTGAGATAACCTTCTTGA

La PCR de TT que usa TT-EX14\_02F y TT-EX14\_02R genera amplicones de PCR de TT como se esboza a continuación (con una flecha que simboliza la secuencia diana amplificada que se encuentra entre las secuencias flanqueantes del cebador):

**GCTCTTCGATCT**GCAGTCTGTCCTGAACCTGAT————▶TCAAGAAGGT  
TATCTCAAGAAACTGGCAGATCGGAAGGC

(SEQ ID NO: 1 —————▶ SEQ ID NO: 7)

Este amplicón de PCR de TT resultante incluye una secuencia autocomplementaria en los extremos 5' y 3'. Bajo condiciones de PCR estándar, los extremos 5' y 3' autocomplementarios de cada amplicón se autorreasocian y generan una estructura de bucle autocomplementario dentro del amplicón. Cuando se usa el conjunto de cebadores TT-EX14\_02F/TT-EX14\_02R, la región complementaria tiene una longitud de 15 pares de bases, incluyendo la secuencia de hibridación genérica de 13 pb más dos bases complementarias adicionales de la secuencia específica de diana de los cebadores TT-EX14\_02F y TT-EX14\_02R (Figura 6A). La estructura de bucle autocomplementario de 15 pares de bases es lo suficientemente estable como para obstruir una amplificación adicional por PCR del amplicón durante la siguiente PCR de ID, al bloquear el acceso del cebador de PCR de ID a la secuencia de hibridación genérica.

Las regiones complementarias para los conjuntos de cebadores TT-EX14\_02F/TT-EX14\_03R y TT-EX14\_01F/TT-EX14\_02R carecen de la secuencia complementaria adicional y solo tienen una longitud de 13 pares de bases (por ejemplo, el conjunto de cebadores TT-EX14\_02F/TT-EX14\_03R como se muestra en Figura 6B). Bajo las mismas condiciones, este tramo no es suficiente para generar una estructura de bucle autocomplementaria lo suficientemente estable como para obstruir la amplificación por PCR adicional a partir del amplicón durante la PCR de ID.

En otras palabras, cuando la región de autorreasociación es lo suficientemente larga bajo las condiciones del experimento de PCR (en este caso 15 pares de bases), la estructura de autorreasociación se favorece frente a la reasociación de un cebador de PCR de ID (probablemente debido a la proximidad espacial de las secuencias

complementarias en los extremos 5' y 3' del producto de PCR). En este caso, los cebadores de PCR de ID no pueden hibridar y, por lo tanto, no ocurrirá una amplificación del producto de PCR de TT de autorreasociación (Figuras 3B y 6A). A la inversa, cuando la región de autorreasociación es más corta (en este caso, 13 pares de bases), los cebadores de PCR de ID son capaces de competir con la estructura de autorreasociación, accediendo al molde de amplicón de PCR de TT para reasociarse con el mismo. Como resultado, puede ocurrir la amplificación por PCR de ID (Figuras 4B y 6B). Las dos bases autocomplementarias adicionales de las secuencias específicas de diana se pueden considerar como una secuencia intermedia entre la secuencia de hibridación genérica y la región de unión a diana. Debido al efecto que esta secuencia intermedia tiene en la amplificación en la reacción de PCR de segunda etapa, se designa en el presente documento como una "secuencia de control de desemparejamiento", que puede diseñarse para permitir o bloquear el emparejamiento de pares de bases adicionales y, por tanto, la formación de una estructura más larga y con autorreasociación más fuerte que la estructura de autorreasociación creada por la propia secuencia de hibridación genérica.

### Ejemplo 2

#### Selección de amplicones deseados en PCR multiplexada

15 A continuación, se planteó la hipótesis de que en una reacción de PCR multiplexada que usa cebadores superpuestos, pueden incluirse secuencias de control de desemparejamiento o secuencias ligadoras en los cebadores de PCR de primera etapa para permitir la formación selectiva de una estructura autocomplementaria, y por tanto, puede asegurarse la formación de amplicones de PCR deseados, al mismo tiempo que se inhibe la formación de amplicones de PCR no deseados.

20 Para probar esta hipótesis de estructuras autocomplementarias selectivas en un experimento, se diseñaron dos conjuntos de siete pares de cebadores cada uno siguiendo los principios generales presentados en el Ejemplo 1. Por tanto, cada par de cebadores de PCR de TT se diseñó para superponerse con el primer cebador del siguiente par de cebadores de amplicón, para cubrir todo el exón 14 del gen *CFTR* (Figura 7).

25 En el primer conjunto de siete pares de cebadores de PCR de TT, los cebadores se diseñaron como se esboza en la Figura 1 y 7. Estos cebadores no incluían un control de desemparejamiento adicional o una secuencia ligadora entre la secuencia de hibridación genérica y la secuencia específica de diana. Por tanto, se supondrá que todos los amplicones que surgen de la PCR con este conjunto de cebadores incluyen una región de autorreasociación de solo 13 pares de bases. Cuando se combinan en una reacción de PCR de TT multiplexada de un solo tubo, se espera una multitud de amplicones de PCR de TT cortos y largos (como se representa en la Figura 8).

30 En el segundo conjunto de siete pares de cebadores de PCR de TT, los cebadores eran idénticos al primer conjunto, pero se añadió una secuencia de control de desemparejamiento de 3 pb entre la secuencia de hibridación genérica y la secuencia específica de diana en cada cebador. Las secuencias de control de desemparejamiento en cada cebador se diseñaron (basándose, por ejemplo, en cebadores vecinos y superpuestos) para crear selectivamente estructuras de bucle autocomplementarias de 16 pares de bases de largo en los amplicones de PCR de TT no deseados más cortos, y estructuras de bucle autocomplementarias de 13 pares de bases de largo en los amplicones de PCR de TT deseados más largos. Las estructuras de bucle autocomplementarias de 16 pb se construyeron mediante la inserción de secuencias autocomplementarias de 3 pb además de las secuencias de estructura de bucle autocomplementarias de 13 pb. Las dos secuencias de control de desemparejamiento de 3 bp o ligadora usadas en este experimento eran:

Ligador 1: 5'-GAC-3'

40 Ligador 2: 5'-TCA-3'

Las secuencias ligadoras de 3 pb se colocaron para generar estructuras de bucle autocomplementarias más largas en los amplicones de PCR cortos no deseados y para generar estructuras de bucle autocomplementarias más cortas en los amplicones de PCR más largos y deseados (Figura 9).

45 Bajo condiciones adecuadas, la formación de la estructura de bucle autocomplementaria más larga en un amplicón de PCR de TT obstruye su uso como molde en la reacción de PCR de ID posterior, mientras que la formación de la estructura de bucle autocomplementaria más corta en un amplicón de PCR de TT permite su uso como molde en la reacción de PCR de ID posterior. Como consecuencia, solo se generan amplicones de PCR de ID predefinidos específicos a partir de la reacción de PCR de ID (Figura 10). El principio representado en las Figuras 7-10 se probó experimentalmente.

50 Los dos conjuntos de cebadores de TT descritos anteriormente se usaron en reacciones de PCR de TT multiplexada de primera etapa de un solo tubo, para ambos conjuntos por separado. Se realizó una PCR de TT multiplexada de un solo tubo en ADN genómico humano para ambos conjuntos de cebadores. Los amplicones de PCR de TT resultantes se diluyeron 1:50 en agua y se sometieron entonces a una PCR de ID multiplexada de segunda etapa usando los cebadores de PCR de ID ID-Ion P1/ID-Ion A.

55 Cuando se usó el conjunto de cebadores sin las secuencias de control de desemparejamiento diseñadas, los 7 amplicones de PCR deseados esperados se detectaron mediante electroforesis capilar después de PCR de TT.

Además, se detectó una serie de productos de PCR no deseados pequeños (Figura 11A, dos carriles superiores). Los amplicones de PCR de TT detectados servían como moldes durante la reacción de PCR de ID, generando una multitud de productos de PCR de ID cortos no deseados. La formación de amplicones de PCR de ID no deseados se confirmó mediante electroforesis capilar (Figura 11A, carril inferior).

- 5 Cuando se usó el conjunto de cebadores que incorporaba secuencias de control de desemparejamiento diseñadas, los 7 amplicones de PCR deseados esperados se detectaron mediante electroforesis capilar después de PCR de TT multiplexada. Además, se detectó una serie de productos de PCR no deseados pequeños (Figura 11B, dos carriles superiores). En este caso, sin embargo, solo los amplicones de PCR de TT más largos deseados servían como moldes durante la PCR de ID, lo que generó por tanto solo los productos de PCR de ID deseados. La formación de solo los amplicones de PCR de ID deseados se confirmó mediante electroforesis capilar (Figura 11B, carril inferior).

El cálculo de áreas bajo la curva (AUC) para todos los picos en los experimentos de electroforesis capilar que se muestran en las Figuras 11A y 11B, paneles inferiores, procuró los siguientes números:

	<i>Sin secuencias de control de desemparejamiento (Fig. 11A)</i>	<i>Con secuencias de control de desemparejamiento (Fig. 11B)</i>
AUC total:	313457	101921
Picos de producto deseado:	156330	98517
Picos de producto no deseado:	157127	3404
<i>Porcentaje de productos no deseados:</i>	<i>50 %</i>	<i>3 %</i>

### Ejemplo 3

#### Valoración de la relación de productos de amplificación deseados frente a no deseados

- 15 Para valorar la cantidad de productos cortos no deseados en comparación con la cantidad de productos de amplicón más largos deseados en una PCR multiplexada para la preparación de amplicones de molde para la secuenciación de nueva generación, se comparó el número de lecturas generadas en la reacción de secuenciación para la región de amplicón deseada con el número de lecturas de la región superpuesta. La Figura 12 representa un ejemplo con cuatro amplicones que cubren una región de interés. En este caso, durante la primera PCR multiplexada o "de TT", se generan tres productos cortos no deseados, cuatro productos más largos deseados y dos productos de translectura. Los productos de translectura se generan mediante la interacción de PCR del primer cebador directo con el tercer cebador inverso y del segundo cebador directo con el cuarto cebador inverso, respectivamente. La Figura 13 muestra la localización de las lecturas de secuencias relativas entre sí después del recorte de cebador. El recuento de amplicones (o recuento de cobertura de amplicones, CA) se generó contando el número de lecturas en las posiciones marcadas para cada amplicón. Todas las lecturas que cubren estas posiciones son específicas para el amplicón valorado. El recuento de la región de superposición (o el recuento de cobertura para la región de superposición, CO) se generó contando el número de lecturas dentro de las regiones de superposición, marcadas con círculos en la Figura 13.

- Si se usan los productos cortos no deseados generados durante la PCR de TT como moldes durante la PCR de ID, la cobertura de la región superpuesta supera la suma de las coberturas de los amplicones adyacentes. Esta situación se expresa de la siguiente manera:

$$CA_n + CA_{n+1} < CO_n$$

- Si, por otro lado, el uso de secuencias de control de desemparejamiento en la PCR de TT, como se divulga en el presente documento, evita que se usen los productos cortos no deseados como moldes durante la segunda PCR multiplexada o "ID", la cobertura de la región superpuesta debería ser igual a la suma de las coberturas de los amplicones adyacentes. Esta situación se expresa de la siguiente manera:

$$CA_n + CA_{n+1} = CO_n$$

- Para determinar la cantidad de cobertura en las regiones superpuestas debido a las lecturas de secuencia de productos de PCR de TT no deseados cortos, es decir, la incapacidad de formación de estructuras autocomplementarias y, por tanto, la formación de amplicones cortos no deseados en PCR de ID, se calculó una relación de cobertura (CR).

$$CR = \frac{CO_n}{CA_n + CA_{n+1}}$$

Una relación de cobertura de 1 indica que la cobertura en las regiones superpuestas solo deriva de los amplicones

que forman la superposición. Un índice de cobertura mayor que 1 indica que los productos de PCR de TT no deseados cortos contribuyen a la cobertura de la región de solapamiento. Esto significa que la formación de amplicones de PCR no deseados durante la PCR de ID no se inhibe como se desea. En los siguientes ejemplos, la relación de cobertura CR se usa como medida de si el procedimiento divulgado en el presente documento tiene o no el efecto deseado.

5 Ejemplo 4

39 amplicones superpuestos del gen BRCA 2

10 Se repitió el Ejemplo 2, pero utilizando el exón 11 del gen BRCA 2 como diana. Se amplificó la diana con 39 amplicones superpuestos usando 78 cebadores de PCR de TT diseñados según la divulgación para incluir secuencias de control de desemparejamiento de 3 pb como en el Ejemplo 2. Los 39 amplicones de PCR de TT se sometieron a PCR de ID como se describe en el Ejemplo 2. La colección resultante se secuenció en un instrumento Illumina MiSeq MPS. La relación de cobertura para cada uno de los 39 amplicones se calculó como se describe en el Ejemplo 3, y los resultados se presentan en la Tabla 2.

Tabla 2

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones (CA <sub>n</sub> + CA <sub>n+1</sub> )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 1	690	1166	1154	0,99
A 2	476	1254	1252	1,00
A 3	778	1708	1670	0,98
A 4	930	1430	1360	0,95
A 5	500	1492	1446	0,97
A 6	992	1522	1452	0,95
A 7	530	1388	1226	0,88
A 8	858	1474	1566	1,06
A 9	616	1132	1120	0,99
A 10	516	1098	1092	0,99
A 11	582	1596	1588	0,99
A 12	1014	1732	1738	1,00
A 13	718	1508	1462	0,97
A 14	790	1496	1462	0,98
A 15	706	1656	1610	0,97
A 16	950	1866	1780	0,95
A 17	916	1740	1628	0,94
A 18	824	1378	1342	0,97
A 19	554	1052	990	0,94
A 20	498	1456	1376	0,95
A 21	958	1842	1718	0,93
A 22	884	1456	1322	0,91
A 23	572	1316	1140	0,87
A 24	744	1598	1390	0,87
A 25	854	1872	1744	0,93
A 26	1018	2204	2218	1,01
A 27	1186	1992	1940	0,97
A 28	806	1530	1498	0,98
A 29	724	1720	1658	0,96
A 30	996	1682	1552	0,92
A 31	686	1392	1314	0,94
A 32	706	1618	1450	0,90
A 33	912	1856	1794	0,97

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones (CA <sub>n</sub> + CA <sub>n+1</sub> )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 34	944	2018	1992	0,99
A 35	1074	1568	1542	0,98
A 36	494	1266	1228	0,97
A 37	772	1782	1774	1,00
A 38	1010	1674	1652	0,99
A 39	664			

La Tabla 2 muestra las relaciones de cobertura para las 38 regiones superpuestas de los 39 amplicones usados para crear una colección de secuenciación para el exón 11 del gen BRCA 2. Solo dos de las 38 regiones superpuestas tienen relaciones de cobertura superiores a 1,00. La relación de cobertura más alto, 1,06, se encontró para la región de superposición 8. Esto significa que solo el 6 % de las lecturas que cubren esta región de solapamiento provenían de productos de PCR de TT cortos no deseados.

La razón por la que la mayoría de las relaciones de cobertura son menores que 1 es la formación de productos de translectura que contribuyen de manera desigual a los recuentos de lectura de los amplicones y las regiones de superposición, respectivamente. Esto se debe a que un producto de translectura se cuenta dos veces para el recuento de lecturas de amplicón CA (una vez para el amplicón n y una vez para el amplicón n + 1), pero solo una vez para el recuento de lectura de superposición CO.

#### Ejemplo 5

##### 31 amplicones superpuestos del gen BRCA 1

Se repitió el Ejemplo 2, pero utilizando los exones 10 y 11 del gen BRCA 1 como diana. Se amplificó la diana con 31 amplicones superpuestos usando 62 cebadores de PCR de TT diseñados según la divulgación para incluir secuencias de control de desemparejamiento de 3 pb como en el Ejemplo 2. Los 31 amplicones de PCR de TT se sometieron a PCR de ID como se describe en el Ejemplo 2. La colección resultante se secuenció en un instrumento Illumina MiSeq MPS. La relación de cobertura para cada uno de los 31 amplicones se calculó como se describe en el Ejemplo 3, y los resultados se presentan en la Tabla 3.

Tabla 3

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones (CA <sub>n</sub> + CA <sub>n+1</sub> )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 1	470	1188	1204	1,01
A 2	718	1414	1226	0,87
A 3	696	1538	1368	0,89
A 4	842	1266	1186	0,94
A 5	424	1016	1002	0,99
A 6	592	1214	1200	0,99
A 7	622	1334	1306	0,98
A 8	712	1472	1424	0,97
A 9	760	1404	1364	0,97
A 10	644	1378	1270	0,92
A 11	734	1494	1450	0,97
A 12	760	1414	1380	0,98
A 13	654	1260	1148	0,91
A 14	606	1372	1310	0,95
A 15	766	1244	1198	0,96
A 16	478	1220	1186	0,97
A 17	742	1502	1424	0,95
A 18	760	1536	1448	0,94

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones (CA <sub>n</sub> + CA <sub>n+1</sub> )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 19	776	1582	1492	0,94
A 20	806	1738	1656	0,95
A 21	932	1820	1812	1,00
A 22	888	1518	1452	0,96
A 23	630	1526	1484	0,97
A 24	896	1722	1680	0,98
A 25	826	1252	1278	1,02
A 26	426	1218	1178	0,97
A 27	792	1312	1268	0,97
A 28	520	1246	1156	0,93
A 29	726	1590	1500	0,94
A 30	864	1584	1474	0,93
A 31	720			

La Tabla 3 muestra las relaciones de cobertura para las 30 regiones superpuestas de los 31 amplicones usados para crear una colección de secuenciación para los exones 10 y 11 del gen BRCA1. Solo dos de las 30 regiones superpuestas tienen relaciones de cobertura superiores a 1,00. La relación de cobertura más alta, 1,02, se encontró para la región de superposición 25. Esto significa que solo el 2 % de las lecturas que cubren esta región de solapamiento provenían de productos de PCR de TT cortos no deseados.

5

Ejemplo 6

7 amplicones superpuestos del gen CFTR

Se repitió el ejemplo 2. Se amplificó la diana con 7 amplicones superpuestos usando 14 cebadores de PCR de TT diseñados según la divulgación para incluir secuencias de control de desemparejamiento de 3 pb como en el Ejemplo 2. Los 7 amplicones de TT PCR se sometieron a PCR de ID como se describe en el Ejemplo 2. La colección resultante se secuenció en un instrumento Illumina MiSeq MPS. La relación de cobertura para cada uno de los 7 amplicones se calculó como se describe en el Ejemplo 3, y los resultados se presentan en la Tabla 4.

10

Tabla 4

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones (CA <sub>n</sub> + CA <sub>n+1</sub> )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 1	148	218	220	1,01
A 2	70	958	940	0,98
A 3	888	1348	1250	0,93
A 4	460	966	922	0,95
A 5	506	936	902	0,96
A 6	430	830	824	0,99
A 7	400			

La Tabla 4 muestra las relaciones de cobertura para las 6 regiones superpuestas de los 7 amplicones usados para crear una colección de secuenciación para el exón 14 del gen CFTR. Solo dos de las 7 regiones superpuestas tienen relaciones de cobertura superiores a 1,00. La relación de cobertura más alta, 1,01, se encontró para la región de superposición 1. Esto significa que solo el 1 % de las lecturas que cubren esta región de solapamiento provienen de productos de PCR de TT cortos no deseados.

15

Ejemplo 7

7 amplicones superpuestos del gen TP53

Se repitió el Ejemplo 2, pero usando los exones 4 y 5 del gen TP53 como diana. Se amplificó la diana con 7 amplicones

20

superpuestos usando 14 cebadores de PCR de TT diseñados según la divulgación para incluir secuencias de control de desemparejamiento de 3 pb como en el Ejemplo 2. Los 7 amplicones de TT PCR se sometieron a PCR de ID como se describe en el Ejemplo 2. La colección resultante se secuenció en un instrumento Illumina MiSeq MPS. La relación de cobertura para cada uno de los 7 amplicones se calculó como se describe en el Ejemplo 3, y los resultados se presentan en la Tabla 5.

5

Tabla 5

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones ( $CA_n + CA_{n+1}$ )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 1	1798	5276	4486	0,85
A 2	3478	6612	5622	0,85
A 3	3134	5462	3850	0,70
A 4	2328	4584	3256	0,71
A 5	2256	4846	4404	0,91
A 6	2590	4208	4008	0,95
A 7	1618			

La Tabla 5 muestra las relaciones de cobertura para las 6 regiones superpuestas de los 7 amplicones usados para crear una colección de secuenciación para los exones 4 y 5 del gen TP53. Ninguna de las 7 regiones superpuestas tiene una relación de cobertura superior a 1,00, lo que indica que no se formaron productos de PCR de TT cortos no deseados.

10

Ejemplo 8

5 amplicones superpuestos del gen KIT

Se repitió el Ejemplo 2, pero usando los exones 12 y 13 del gen KIT como diana. Se amplificó la diana con 5 amplicones superpuestos usando 10 cebadores de PCR de TT diseñados según la divulgación para incluir secuencias de control de desemparejamiento de 3 pb como en el Ejemplo 2. Los 5 amplicones de PCR de TT se sometieron a PCR de ID como se describe en el Ejemplo 2. La colección resultante se secuenció en un instrumento Illumina MiSeq MPS. La relación de cobertura para cada uno de los 5 amplicones se calculó como se describe en el Ejemplo 3, y los resultados se presentan en la Tabla 6.

15

Tabla 6

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones ( $CA_n + CA_{n+1}$ )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 1	706	1702	1638	0,96
A 2	996	1936	1728	0,89
A 3	940	1952	1792	0,92
A 4	1012	1772	1724	0,97
A 5	760			

20

La Tabla 6 muestra las relaciones de cobertura para las 5 regiones superpuestas de los 6 amplicones usados para crear una colección de secuenciación para los exones 12 y 13 del gen TP53. Ninguna de las 5 regiones superpuestas tiene una relación de cobertura superior a 1,00, lo que indica que no se formaron productos de PCR TT cortos no deseados.

Ejemplo 9

25

3 amplicones superpuestos del gen PDGFRA

Se repitió el Ejemplo 2, pero usando el exón 12 del gen PDGFRA como diana. Se amplificó la diana con 3 amplicones superpuestos usando 6 cebadores de PCR de TT diseñados según la divulgación para incluir secuencias de control de desemparejamiento de 3 pb como en el Ejemplo 2. Los 3 amplicones de PCR de TT se sometieron a PCR de ID como se describe en el Ejemplo 2. La colección resultante se secuenció en un instrumento Illumina MiSeq MPS. La relación de cobertura para cada uno de los 3 amplicones se calculó como se describe en el Ejemplo 3, y los resultados

30

se presentan en la Tabla 7.

Tabla 7

Amplicón	Cobertura de amplicón (CA)	Cobertura de la suma de amplicones ( $CA_n + CA_{n+1}$ )	Cobertura de solapamiento (CO)	Relación de cobertura $\left(\frac{CO_n}{CA_n + CA_{n+1}}\right)$
A 1	451	1068	1006	0,94
A 2	617	1539	1472	0,96
A 3	922			

5 La Tabla 7 muestra las relaciones de cobertura para las 2 regiones superpuestas de los 3 amplicones usados para crear una colección de secuenciación para el exón 12 del gen PDGFRA. Ninguna de las 2 regiones superpuestas tiene una relación de cobertura superior a 1,00, lo que indica que no se formaron productos de PCR TT cortos no deseados.



**REIVINDICACIONES**

1. Procedimiento para la amplificación selectiva de amplicones deseados en una reacción en cadena de la polimerasa en dos etapas, que comprende los pasos de:

- proporcionar un material genético diana;

5 • desnaturalizar el material genético diana;

- añadir un conjunto de cebadores diana en condiciones que permitan la reasociación de cebadores con material genético diana;

- \* comprendiendo cada cebador en dicho conjunto de cebadores diana

- ◆ una secuencia específica de diana,.

10 ◆ una secuencia de control de desemparejamiento, y

- ◆ una secuencia de hibridación genérica;

- \* siendo dicha secuencia de hibridación genérica idéntica en cada cebador diana;

- \* estando dicha secuencia de control de desemparejamiento y secuencia de hibridación genérica próximas entre sí; y

15 \* comprendiendo dicho conjunto de cebadores diana al menos dos pares de cebadores adaptados para la amplificación de al menos dos amplicones deseados superpuestos en el material genético diana, de tal modo que existe la posibilidad de obtener al menos un amplicón no deseado en la región de superposición;

- \* estando **caracterizado** además dicho conjunto de cebadores diana **porque** cualquier par de cebadores que flanquean una secuencia de amplicón deseada tiene secuencias de control de desemparejamiento desemparejadas, mientras que cualquier par de cebadores que flanquean una secuencia de amplicón no deseada tiene secuencias de control de desemparejamiento emparejadas;

20 • llevar a cabo una primera reacción en cadena de la polimerasa multiplexada, dando como resultado amplicones superpuestos correspondientes a diferentes combinaciones posibles de pares de cebadores en el primer conjunto de cebadores, manteniéndose dichos amplicones superpuestos en condiciones tales que se autohibridan a través del acoplamiento de pares de bases entre las secuencias complementarias resultantes de la secuencia de hibridación genérica en cada cebador;

25 • tratar dichos amplicones superpuestos en condiciones de selección que son tales que los amplicones deseados se desnaturalizan debido a secuencias no complementarias resultantes de las secuencias de control de desemparejamiento desemparejadas, mientras que los amplicones no deseados permanecen autohibridados debido a secuencias complementarias adicionales resultantes de las secuencias de control de desemparejamiento emparejadas;

30 • añadir cebadores genéricos, cada uno de los cuales comprende una secuencia que es complementaria de la secuencia de hibridación genérica;

35 • llevar a cabo una segunda reacción en cadena de la polimerasa, que da como resultado la amplificación selectiva de los amplicones deseados.

2. Procedimiento según la reivindicación 1, en el que dicha primera y segunda reacciones en cadena de la polimerasa se llevan a cabo en diferentes recipientes de reacción.

3. Procedimiento según la reivindicación 1, en el que dicha primera y segunda reacciones en cadena de la polimerasa se llevan a cabo en el mismo recipiente de reacción.

40 4. Procedimiento según la reivindicación 2 o 3, en el que dichos cebadores diana se añaden antes de dicha primera reacción en cadena de la polimerasa multiplexada y dichos cebadores genéricos se agregan después de dicha primera reacción en cadena de la polimerasa multiplexada y antes de dicha segunda reacción en cadena de la polimerasa.

5. Procedimiento según la reivindicación 3, en el que dichos cebadores diana y cebadores genéricos se añaden antes de dicha primera reacción en cadena de la polimerasa multiplexada.

45 6. Procedimiento según cualquier reivindicación precedente, en el que la longitud total de un cebador diana en dicho conjunto de cebadores diana está en el intervalo de aproximadamente 7 a aproximadamente 200 nucleótidos.

7. Procedimiento según cualquier reivindicación precedente, en el que la longitud de dicha secuencia de control de desemparejamiento está en el intervalo de 1 a 7 nucleótidos.

8. Procedimiento según cualquier reivindicación precedente, en el que la longitud de dicha secuencia de hibridación genérica está en el intervalo de 1 a 75 nucleótidos.
9. Procedimiento según cualquier reivindicación precedente, en el que la longitud de dicha secuencia específica de diana está en el intervalo de 5 a 100 nucleótidos.
- 5 10. Procedimiento según cualquier reivindicación precedente, en el que al menos un cebador en al menos un par está marcado de forma detectable.
11. Procedimiento según cualquier reivindicación precedente, en el que dichos cebadores genéricos comprenden además secuencias adaptadoras para la secuenciación de dichos amplicones deseados.
- 10 12. Procedimiento según la reivindicación 11, que comprende además el paso de secuenciar dichos amplicones deseados.
13. Procedimiento según la reivindicación 12, en el que dicha secuenciación se lleva a cabo usando una secuenciación paralela masiva.
14. Procedimiento según cualquier reivindicación precedente, en el que dichos cebadores genéricos comprenden además secuencias de índice.
- 15 15. Procedimiento según cualquier reivindicación precedente, en el que al menos un cebador diana es complementario de una secuencia de diagnóstico en el material genético diana.
16. Kit para llevar a cabo el procedimiento según la reivindicación 1, que comprende:
- un conjunto de cebadores diana;
    - \* comprendiendo cada cebador en dicho conjunto de cebadores diana
      - 20 ♦ una secuencia específica de diana,.
      - ♦ una secuencia de control de desemparejamiento, y
      - ♦ una secuencia de hibridación genérica;
    - \* siendo dicha secuencia de hibridación genérica idéntica en cada cebador diana;
    - \* estando dicha secuencia de control de desemparejamiento y secuencia de hibridación genérica próximas entre sí; y
    - 25 \* comprendiendo dicho conjunto de cebadores diana al menos dos pares de cebadores adaptados para la amplificación de al menos dos amplicones deseados superpuestos en un material genético diana, de tal modo que existe la posibilidad de obtener al menos un amplicón no deseado en la región de superposición;
    - \* estando **caracterizado** además dicho conjunto de cebadores diana porque cualquier par de cebadores que flanquean una secuencia de amplicón deseada tiene secuencias de control de desemparejamiento desemparejadas, mientras que cualquier par de cebadores que flanquean una secuencia de amplicón no deseada tiene secuencias de control de desemparejamiento emparejadas;
    - 30
  - cebadores genéricos, cada uno de los cuales comprende una secuencia que es complementaria de la secuencia de hibridación genérica; e
  - 35 • instrucciones para llevar a cabo el procedimiento según la reivindicación 1 usando dichos cebadores.
17. Kit según la reivindicación 15, en el que dichos cebadores son como se definen en una cualquiera de las reivindicaciones 6-11 y 14-15.

Figura 1

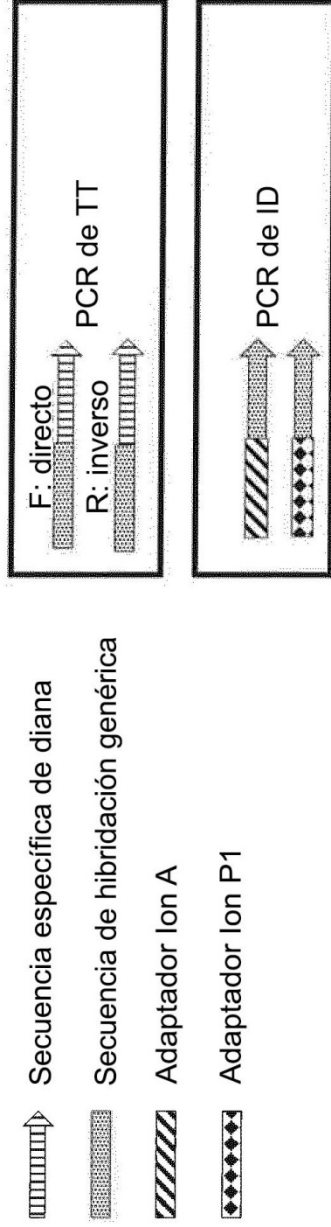


Figura 2

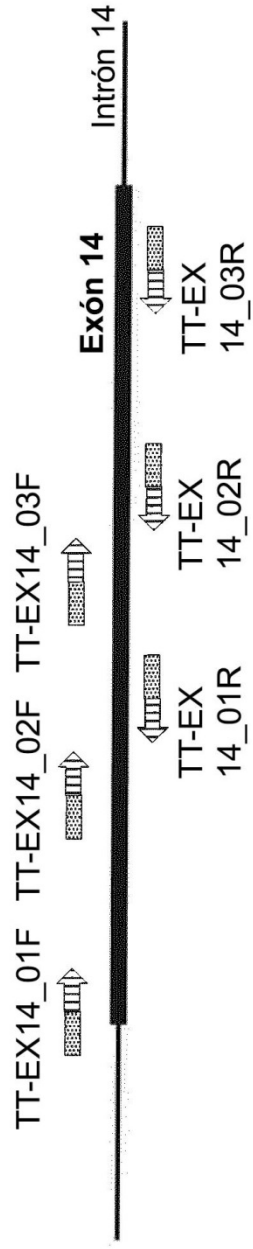


Figura 3

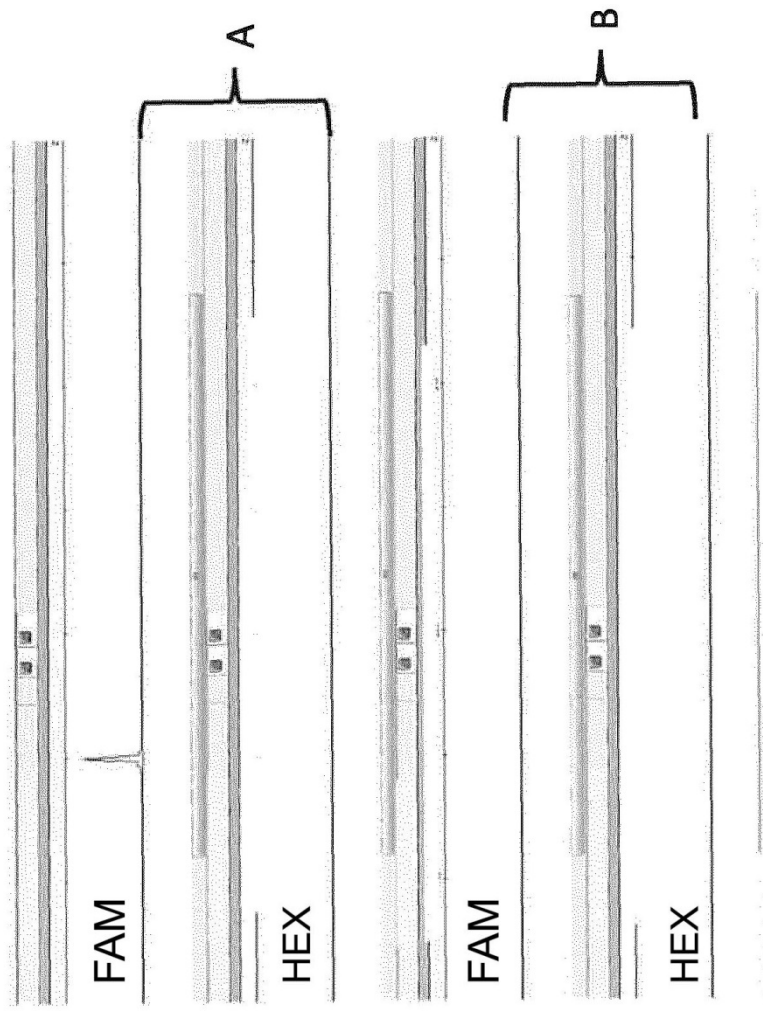


Figura 4

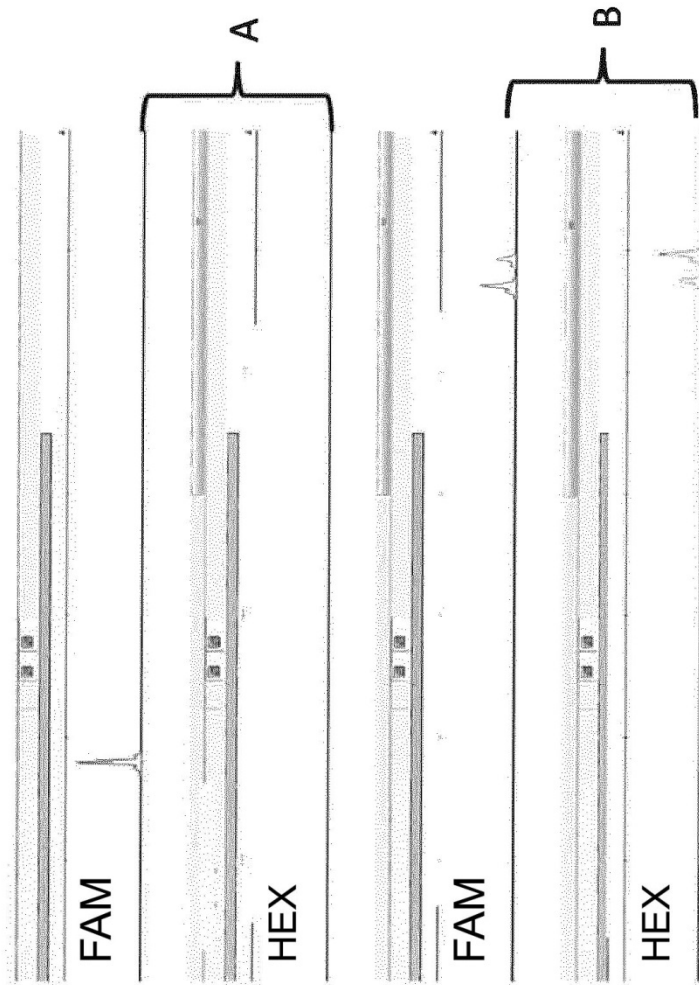


Figura 5



**Figura 6**

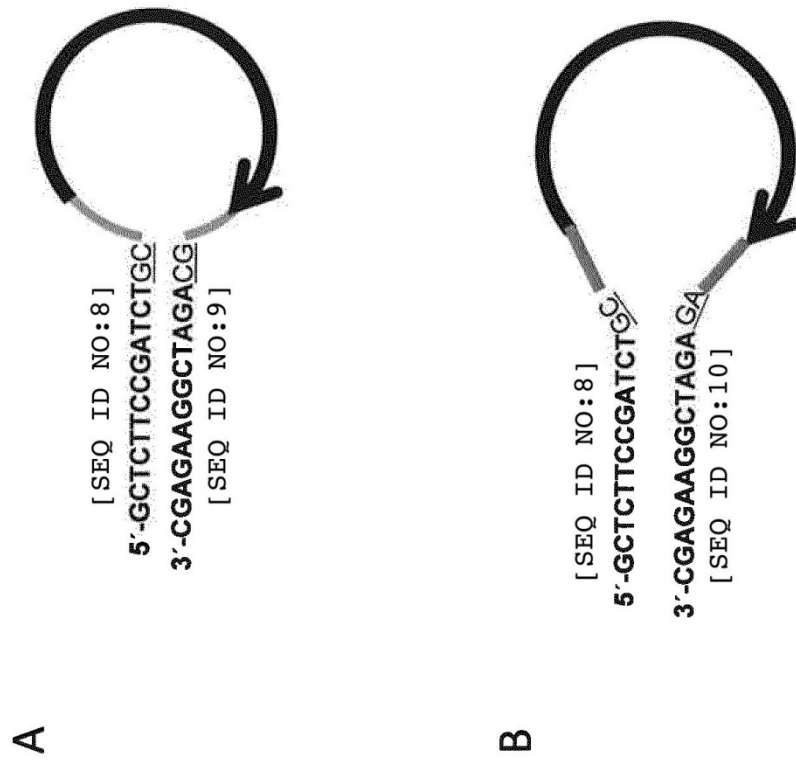


Figura 7



Figura 8

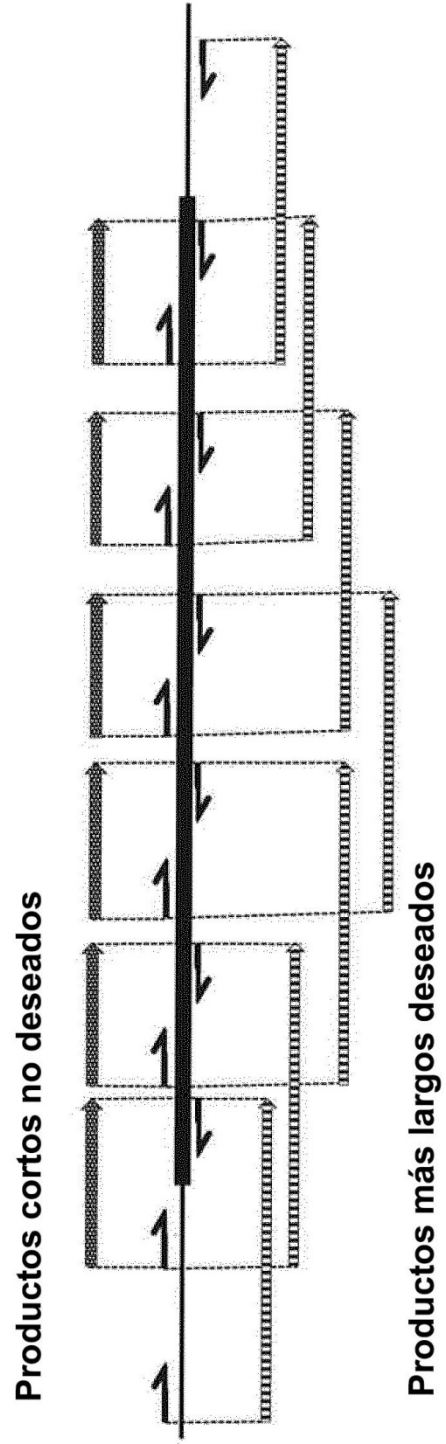




Figura 9

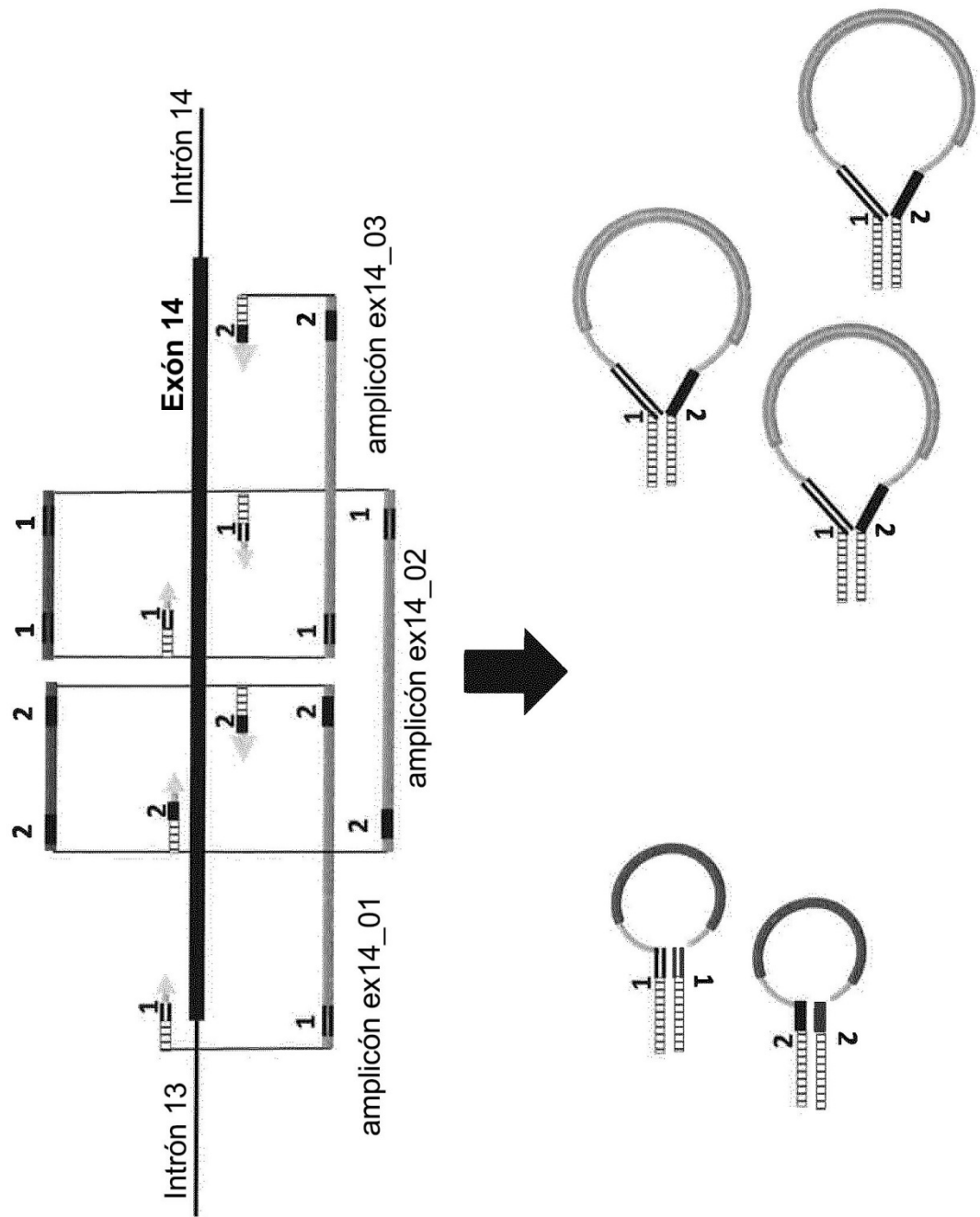


Figura 10

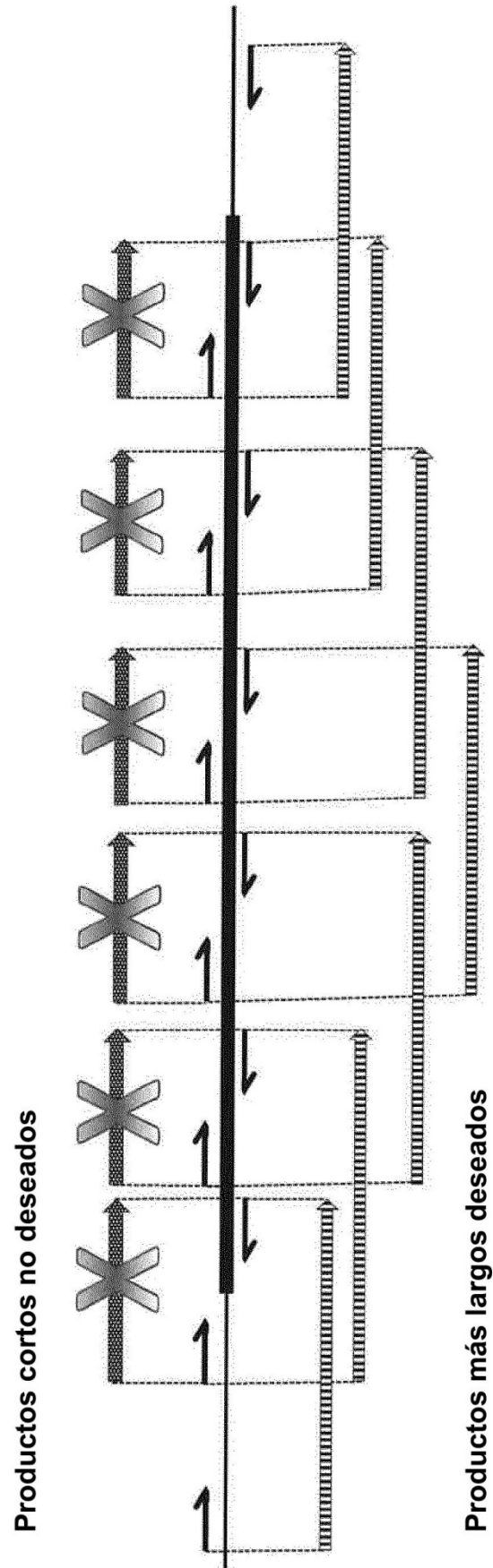


Figura 11A

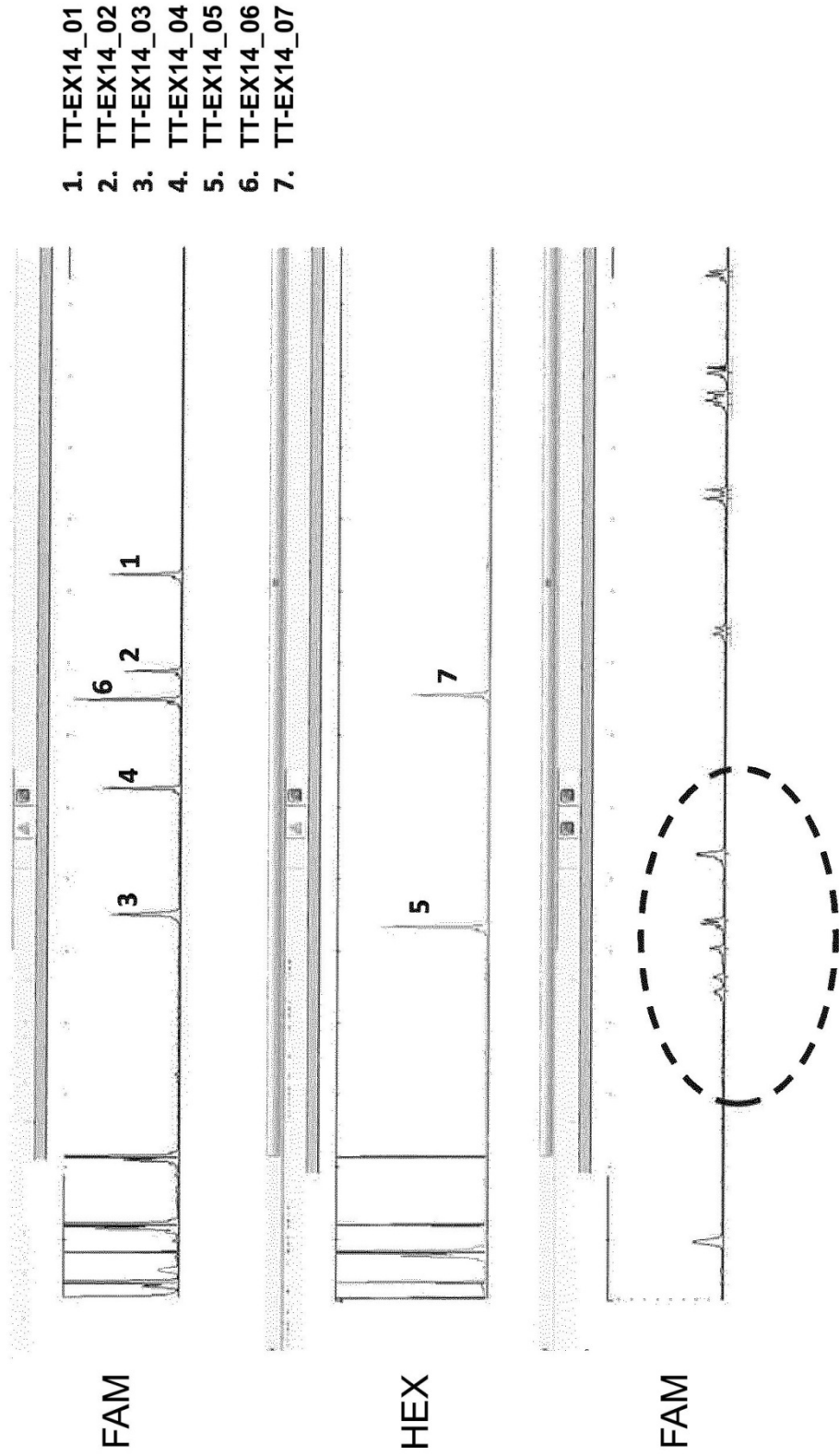


Figura 11B



Figura 12

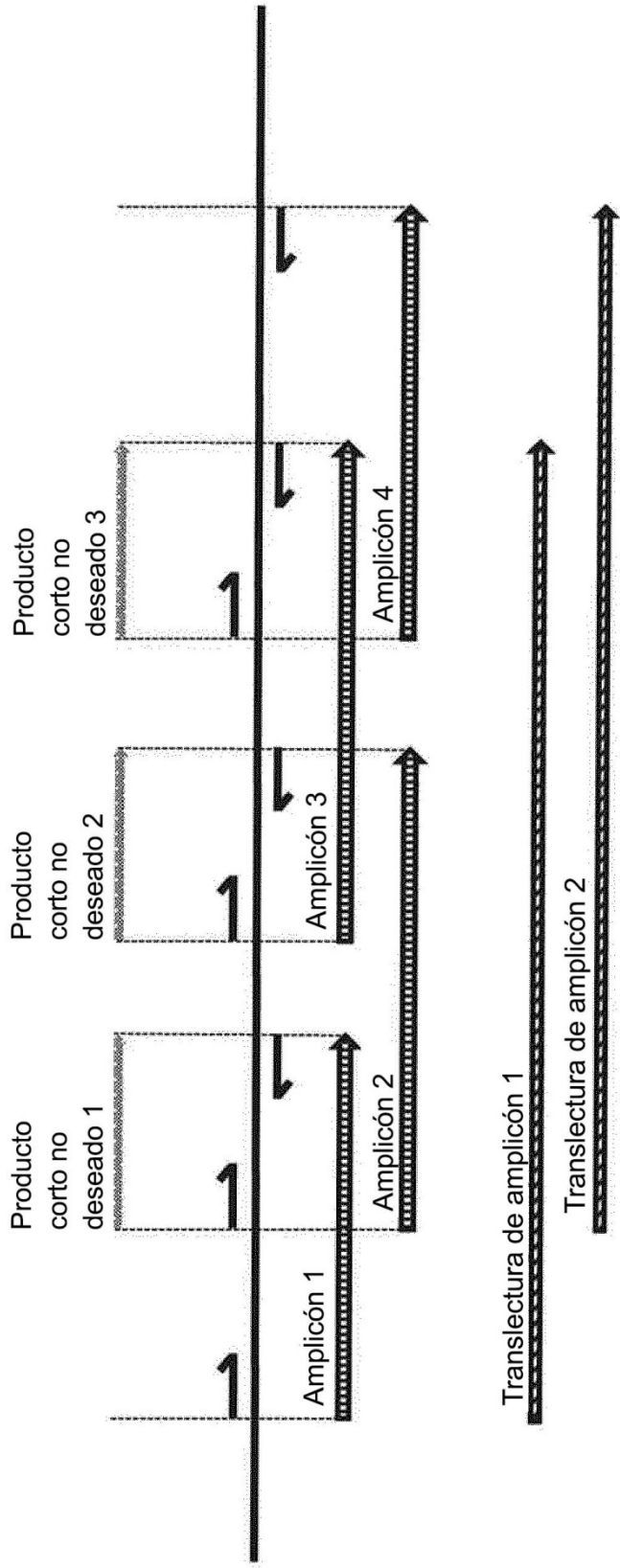


Figura 13

