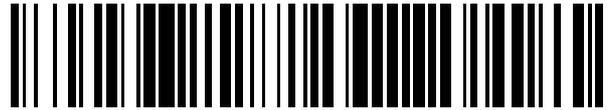


19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 705 427**

51 Int. Cl.:

**C12Q 1/68**

(2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **11.09.2015 PCT/GB2015/052626**

87 Fecha y número de publicación internacional: **17.03.2016 WO16038381**

96 Fecha de presentación y número de la solicitud europea: **11.09.2015 E 15766213 (1)**

97 Fecha y número de publicación de la concesión europea: **31.10.2018 EP 3191603**

54 Título: **Método de obtención de información de secuenciación de extremos emparejados**

30 Prioridad:

**11.09.2014 GB 201416106**

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

**25.03.2019**

73 Titular/es:

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

72 Inventor/es:

**RIGATTI, ROBERTO y  
BOUTELL, JONATHAN**

74 Agente/Representante:

**ELZABURU, S.L.P**

**ES 2 705 427 T3**

Aviso: En el plazo de nueve meses a contar desde la fecha de publicación en el Boletín Europeo de Patentes, de la mención de concesión de la patente europea, cualquier persona podrá oponerse ante la Oficina Europea de Patentes a la patente concedida. La oposición deberá formularse por escrito y estar motivada; sólo se considerará como formulada una vez que se haya realizado el pago de la tasa de oposición (art. 99.1 del Convenio sobre Concesión de Patentes Europeas).

**DESCRIPCIÓN**

Método de obtención de información de secuenciación de extremos emparejados

5 La invención se refiere en general al campo de la secuenciación de ácidos nucleicos. En particular, la invención se refiere a un método de obtención de información de secuenciación de extremos emparejados. La invención también se refiere a un kit para el uso en el método y a un sistema que incorpora el método.

La secuenciación de extremos emparejados (EE) o por pares es una técnica conocida que implica secuenciar ambos extremos de los fragmentos de ácidos nucleicos de una biblioteca de secuenciación y alinear las lecturas directa e inversa como pares de lectura.

10 Además de producir el doble de lecturas para el mismo tiempo y esfuerzo en la preparación de la biblioteca, las secuencias alineadas en forma de pares de lectura posibilitan una alineación de lecturas más exacta.

15 La secuenciación de extremos emparejados permite a los usuarios secuenciar los dos extremos de un fragmento y generar datos de secuencia alineables de alta calidad. Facilita la detección de reordenamientos genómicos y elementos de secuencias repetitivas, así como fusiones de genes y transcritos nuevos. Debido a que las lecturas de extremos emparejados son más susceptibles a alinearse respecto de una referencia, la calidad del conjunto de datos completo mejora mediante el uso de esta técnica.

Sería deseable proporcionar un método mejorado de obtención de información de secuenciación de extremos emparejados para reducir el coste, incrementar el rendimiento, y/o simplificar el proceso.

20 Ming, Y. *et al.*, "Dual Primer Emulsion PCR for Next-Generation DNA Sequencing", *Biotechniques*, 48(5), pág. 409, 2010 describe un método de PCR de emulsión con cebadores dobles en el que ambos cebadores están unidos a una microesfera.

El documento WO 2007/086935 A2 describe un método de secuenciación de ADN que comprende preparar una muestra de ADN, amplificar el ADN preparado y llevar a cabo múltiples reacciones de secuenciación con el ADN amplificado solamente con una etapa de hibridación de cebadores.

25 Satterfield, B. *et al.*, "2.5 Million-Fold Improvement in the Reduction of Nonspecific Amplification", *the Journal of Molecular Diagnostics*, Vol. 16, Nº 2, 2014, describe un método para reducir la propagación de dímeros de cebadores mediante el uso de cebadores cooperativos.

**Declaración de la Invención**

Un aspecto de la invención proporciona un método de obtención de información de secuenciación de extremos emparejados en una reacción de secuenciación que comprende las etapas de:

30 proporcionar un ácido nucleico objetivo inmovilizado en un soporte sólido;

determinar la secuencia del ácido nucleico objetivo generando una cadena de ácido nucleico complementaria, y la cadena de ácido nucleico complementaria comprende como mínimo un resto unible;

determinar la secuencia de la cadena de ácido nucleico complementaria; y

35 comparar las secuencias del ácido nucleico objetivo y de la cadena de ácido nucleico complementaria para proporcionar la información de secuenciación de extremos emparejados.

De manera ventajosa, el como mínimo un resto unible se puede incorporar en la cadena complementaria durante la síntesis de la cadena.

El como mínimo un resto unible se puede proporcionar en un primer cebador de secuenciación.

El como mínimo un resto unible se puede proporcionar en un oligonucleótido de prolongación.

40 El como mínimo un resto unible se puede proporcionar en un nucleótido modificado.

En una realización, el soporte sólido comprende como mínimo un resto unible correspondiente para unir de manera covalente o no covalente selectivamente el como mínimo un resto unible de la cadena de ácido nucleico complementaria en condiciones predeterminadas.

45 En una realización, el como mínimo un resto unible correspondiente comprende un oligonucleótido de superficie en el soporte sólido, y el como mínimo un resto unible comprende una secuencia de ácido nucleico sustancialmente complementaria al oligonucleótido de superficie.

En una realización, el ácido nucleico objetivo se puede inmovilizar en el soporte sólido mediante un oligonucleótido de superficie que comprende una primera secuencia, y el oligonucleótido de superficie con resto unible

correspondiente del soporte sólido puede tener una segunda secuencia que difiere de la primera secuencia.

Esto proporciona la ventaja de minimizar la hibridación indeseable durante la etapa de hibridación, ya que se pueden utilizar oligonucleótidos de superficie diferentes para una etapa de amplificación y una etapa de hibridación.

5 En una realización, una temperatura incrementada en una etapa de hibridación puede inhibir la hibridación del resto unible a un oligonucleótido de superficie del soporte sólido.

Esto proporciona la ventaja de minimizar la hibridación indeseable durante la etapa de hibridación.

La invención proporciona la ventaja de poder anclar la cadena complementaria a un cebador de superficie de un soporte sólido del tipo usado para las reacciones de secuenciación de última generación.

10 En una realización, el como mínimo un resto unible correspondiente comprende una porción de la secuencia del ácido nucleico objetivo inmovilizada en el soporte sólido.

En ciertas realizaciones, se puede usar una polimerasa con desplazamiento de cadena en una etapa de prolongación.

Opcionalmente, el método comprende la etapa de activar el como mínimo un resto unible para unir la cadena de ácido nucleico complementaria al soporte sólido.

15 De manera ventajosa, se puede llevar a cabo la activación durante una etapa de secuenciación. Por ejemplo, mediante la adición de una molécula de activación durante una secuenciación mediante una reacción de síntesis. De manera alternativa, o además, la activación puede ser mediante el cambio de una o más condiciones de la reacción.

En una realización, el como mínimo un resto unible comprende un resto biotilado en la cadena complementaria y estreptavidina en el soporte sólido.

20 Preferiblemente, el método comprende la etapa de eliminar el ácido nucleico objetivo tras la etapa de determinación de su secuencia.

La eliminación del ácido nucleico objetivo puede ser mediante una etapa de linealización.

De manera ventajosa, la eliminación del ácido nucleico objetivo se puede llevar a cabo durante una etapa de secuenciación. Por ejemplo, durante los ciclos de secuenciación por síntesis.

25 En ciertas realizaciones, solamente hay unidos ácidos nucleicos idénticos al soporte sólido en cada momento. La cadena objetivo o complementaria se inmoviliza en el soporte sólido durante una lectura de secuenciación.

En una realización preferida, la etapa de determinación de la secuencia del ácido nucleico objetivo comprende hibridar un primer cebador de secuenciación al ácido nucleico objetivo y prolongar el primer cebador de secuenciación para obtener una primera lectura de secuenciación.

30 Opcionalmente, el primer cebador de secuenciación se prolonga hasta sustancialmente la longitud completa del ácido nucleico objetivo tras la primera lectura de secuenciación.

De manera ventajosa, la cadena de ácido nucleico complementaria se puede prolongar hasta sustancialmente la longitud completa del ácido nucleico objetivo.

35 Preferiblemente, la etapa de determinación de la secuencia de la cadena de ácido nucleico complementaria comprende hibridar un segundo cebador de secuenciación a la cadena de ácido nucleico complementaria y prolongar el segundo cebador de secuenciación para obtener una segunda lectura de secuenciación.

La lectura de secuenciación se puede obtener mediante el uso de secuenciación por síntesis o secuenciación por ligadura.

40 La etapa de comparar las secuencias del ácido nucleico objetivo y la cadena de ácido nucleico complementaria para proporcionar la información de secuenciación de extremos emparejados puede utilizar un algoritmo y/o procesador informático.

El ácido nucleico objetivo puede consistir en un ácido nucleico monocatenario.

En una realización, el ácido nucleico objetivo puede consistir en un ácido nucleico bicatenario.

45 En una realización, el ácido nucleico objetivo consiste en un ácido nucleico que es en parte monocatenario y en parte bicatenario.

El ácido nucleico objetivo se puede inmovilizar en el soporte sólido en forma de un ácido nucleico bicatenario antes de proporcionar las cadenas individuales para la secuenciación.

El ácido nucleico objetivo se puede inmovilizar en el soporte sólido en forma de un ácido nucleico bicatenario.

Una recombinasa puede facilitar la hibridación de un cebador de secuenciación en el ácido nucleico bicatenario.

En una realización, el método comprende la pre-etapa de amplificación clonal de la molécula de ácido nucleico objetivo.

- 5 De manera ventajosa, el método se puede aplicar a un agrupamiento o población clonal.

El ácido nucleico puede comprender ADN o ARN.

El soporte sólido puede comprender una celda de flujo o una microesfera.

- 10 Otro aspecto de la invención proporciona un cebador y un oligonucleótido unido a la superficie para el uso en el método, que comprende una secuencia para la hibridación a una secuencia de ácido nucleico objetivo y un resto unible para la inmovilización de una cadena de ácido nucleico complementaria a una secuencia de ácido nucleico objetivo a un soporte sólido, en el que el resto unible comprende una secuencia de ácido nucleico sustancialmente complementaria a un oligonucleótido de superficie del soporte sólido, y en el que la secuencia para la hibridación a una secuencia de ácido nucleico objetivo está unida en su extremo 5' al extremo 5' de la secuencia de ácido nucleico del resto unible.

- 15 La secuencia para la hibridación a una secuencia de ácido nucleico objetivo está unida en su extremo 5' al extremo 3' de la secuencia de ácido nucleico del resto unible.

Otro aspecto de la invención proporciona un sistema para obtener información de secuenciación de extremos emparejados que comprende un aparato de secuenciación que tiene un soporte sólido para inmovilizar un ácido nucleico objetivo, una entrada y una salida para los reactivos de secuenciación, y

- 20 medios para determinar la secuencia del ácido nucleico objetivo generando una cadena de ácido nucleico complementaria, y la cadena de ácido nucleico complementaria comprende como mínimo un resto unible;

determinar la secuencia de la cadena de ácido nucleico complementaria; y

comparar las secuencias del ácido nucleico objetivo y de la cadena de ácido nucleico complementaria para proporcionar la información de secuenciación de extremos emparejados.

- 25 En una realización, el aparato comprende un soporte sólido revestido de estreptavidina.

Otro aspecto de la invención proporciona un kit para obtener información de secuenciación de extremos emparejados en una reacción de secuenciación, que comprende uno o más reactivos de secuenciación y el resto unible descrito anteriormente.

El kit puede comprender instrucciones para llevar a cabo el método de la invención.

- 30 **Breve descripción de los dibujos**

En las figuras, que ilustran los aspectos de la invención a modo de ejemplo solamente:

las Figuras 1A y 1B ilustran vistas laterales de una región de un sustrato de secuenciación para la secuenciación de alto rendimiento;

- 35 las Figuras 2A a 2E ilustran vistas laterales del sustrato de secuenciación de la Figura 1B y muestran un ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido;

las Figuras 3A a 3E ilustran vistas laterales del sustrato de secuenciación de la Figura 1B y muestran otro ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido; y

- 40 las Figuras 4A a 4E ilustran vistas laterales del sustrato de secuenciación de la Figura 1B y muestran otro ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido.

Las Figuras 5A y 5B ilustran una comparación de las etapas del método de la técnica anterior mostrado en 5A respecto de las etapas del método de la presente invención, resumido en la Figura 5B.

- 45 Las Figuras 6A a 6E ilustran vistas laterales del sustrato de secuenciación de la Figura 1B y muestran un ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados.

Las Figuras 7A y 7B ilustran vistas laterales del sustrato de secuenciación de la Figura 1B y muestran un ejemplo de

un proceso de obtención de información de secuenciación de extremos emparejados.

### Descripción

5 La invención proporciona métodos de obtención de información de secuenciación de extremos emparejados cuando solamente un tipo de cadena de ADN está unido a la superficie de un soporte sólido. Por ejemplo, se puede unir y amplificar de manera clonal un tipo de cadena de ADN en un soporte sólido (p.ej., una celda de flujo) mediante el uso de un único tipo de sonda de captura de oligonucleótido. Un tipo de cadena de ADN se puede denominar "A" y su cadena complementaria se puede denominar "A prima" (A'). En diversas realizaciones, un tipo de cadena de ADN (p.ej., la cadena A) se une a la superficie de un soporte sólido y se usa para generar una primera lectura de secuenciación. Posteriormente, se genera una cadena complementaria de longitud sustancialmente completa (p.ej., la cadena A') y se une a la superficie del soporte sólido. El primer tipo de cadena de ADN (p.ej., la cadena A) se elimina, y se obtiene una segunda lectura de secuenciación de la cadena complementaria unida a la superficie (p.ej., la cadena A').

15 En un ejemplo, se usa un primer cebador de secuenciación que incluye un resto unible para obtener una primera lectura de secuenciación. Tras la finalización de la primera lectura de secuenciación, se lleva a cabo una prolongación completa del cebador de secuenciación en presencia de dNTPs (naturales o no naturales) y ADN polimerasa para generar una cadena A' complementaria de longitud sustancialmente completa. La cadena A' complementaria se une posteriormente a la superficie del soporte sólido por medio del resto unible del primer cebador de secuenciación completamente prolongado.

20 En otro ejemplo, se elimina un primer cebador de secuenciación que se usó para obtener una primera lectura de secuenciación de la cadena A, y se hibrida un oligonucleótido de prolongación que incluye un resto unible y se usa para generar una cadena A' complementaria de longitud sustancialmente completa. La cadena A' complementaria se ancla después a la superficie del soporte sólido por medio del resto unible.

25 En otro ejemplo, se pueden incorporar nucleótidos modificados que incluyen un resto unible durante los ciclos de una primera lectura de secuenciación. Estos nucleótidos se pueden incorporar durante los primeros ciclos, o se pueden incorporar durante ciclos posteriores, especialmente si las moléculas son mucho más largas que la longitud de lectura.

30 Tras la finalización de la primera lectura de secuenciación, se lleva a cabo una prolongación completa del primer cebador de secuenciación en presencia de dNTPs y ADN polimerasa para generar una cadena A' complementaria de longitud sustancialmente completa. La cadena A' complementaria se une posteriormente a la superficie del soporte sólido por medio de los restos unibles de los nucleótidos modificados incorporados.

La unión y/o secuenciación posterior de la cadena A' complementaria puede ser en el mismo sustrato/soporte, en sustancialmente la misma región o una región diferente de la primera lectura de secuenciación. De manera alternativa o además, puede ser en un sustrato/soporte diferente.

35 Pueden ser necesarias ciertas condiciones o la activación de los restos unibles para que el resto unible asociado con la cadena A' se una a un resto asociado de la superficie del soporte sólido, de forma que se pueda controlar el anclaje de la cadena complementaria a la superficie.

40 Por ejemplo, en cierta realización, se usan métodos químicos de múltiples etapas de forma que el resto unible no se acople al resto del soporte sólido hasta que se añada un tercer resto u otra condición (p.ej., cambio del pH o la temperatura). Esto permite el control y, si se desea, el aplazamiento del acoplamiento del resto unible durante la secuenciación. Los ejemplos no limitantes de la química de unión adecuada para el uso en la invención se describen en el número de solicitud de EE.UU. 13/784.368.

De manera alternativa o además, se puede llevar a cabo la activación controlada del resto o restos unibles mediante interacciones ligando-receptor / anticuerpo-antígeno, tales como la unión de un anticuerpo a un antígeno

45 Los métodos descritos en la presente memoria se pueden aplicar a moléculas de ADN individuales (p.ej., la cadena A) o a poblaciones clonales de moléculas de ADN (p.ej., poblaciones clonales de la cadena A) unidas a la superficie de un soporte sólido.

Por lo tanto, en comparación con los protocolos actuales de secuenciación de extremos emparejados, la invención para la obtención de información de secuenciación de extremos emparejados proporciona métodos más simples de obtención de información de secuenciación de extremos emparejados.

50 Las Figuras 1A y 1B ilustran vistas laterales de una región de un sustrato de secuenciación 100 para la secuenciación de alto rendimiento. El sustrato de secuenciación 100 incluye un soporte sólido 110. En un ejemplo, el soporte sólido 110 es un sustrato plano, tal como una celda de flujo. En otro ejemplo, el soporte sólido 110 es una microesfera (no mostrada). Con respecto a la Figura 1A, un único tipo de cadena de ADN 115 puede estar unido a la superficie del soporte sólido 110. La cadena de ADN 115 se puede denominar cadena A. El complemento inverso de la cadena A se puede denominar cadena A' (no mostrada).

Con respecto a la Figura 1B, una población clonal 120 de moléculas de ADN puede estar unida a la superficie del soporte sólido 110. La población clonal 120 es una población de moléculas de ADN idénticas que se obtienen mediante amplificación clonal de una única molécula de ADN que se debe secuenciar. Todas las copias idénticas de la población clonal 120 están unidas al sustrato sólido 110. En un ejemplo, la población clonal 120 comprende un tipo de cadena de ADN (p.ej., la cadena A). En otro ejemplo (no mostrado), una primera población clonal 120 puede comprender la cadena A y una segunda población clonal 120 puede comprender la cadena A'. En este ejemplo, se muestra una única población clonal 120, pero se puede unir cualquier número de poblaciones clonales a la superficie del soporte sólido 110.

Las Figuras 2A a 2E ilustran vistas laterales del sustrato de secuenciación 100 de la Figura 1B y muestran un ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido. En este ejemplo, un primer cebador de secuenciación incluye un resto unible que se puede usar para anclar una cadena A' complementaria sustancialmente completamente prolongada a la superficie del soporte sólido. Un ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido puede incluir, pero sin limitación, las etapas siguientes.

En una etapa, la Figura 2A muestra la hibridación de una diversidad de primeros cebadores de secuenciación 210 a la población clonal 120. En este ejemplo, la población clonal 120 comprende la cadena A. Cada primer cebador de secuenciación 210 incluye un resto unible 215 que se puede anclar de manera covalente o no covalente al soporte sólido 110 en ciertas condiciones. En un ejemplo, el resto unible 215 puede ser un resto biotinilado que se puede anclar a un soporte sólido 110 revestido de estreptavidina por medio de un complejo biotina-estreptavidina. En otro ejemplo, el resto unible 215 puede ser cualquier otro grupo funcional adecuado que se puede anclar de manera covalente o no covalente al soporte sólido 110. Los ejemplos de grupos funcionales que se pueden usar para anclar un oligonucleótido a un soporte sólido se describen en "Strategies for Attaching Oligonucleotides to Solid Supports" (Integrated DNA Technologies 2014 (v5)). Se pueden usar los primeros cebadores de secuenciación 210 para obtener la primera lectura en una reacción de secuenciación, tal como secuenciación por síntesis (SBS), secuenciación por ligadura, o cualquier otro método de secuenciación adecuado.

En otra etapa, la Figura 2B muestra la prolongación completa de los primeros cebadores de secuenciación 210 tras la finalización de la primera lectura de secuenciación. Los primeros cebadores de secuenciación 210 se pueden prolongar en presencia de dNTPs naturales y ADN polimerasa para generar una diversidad de cadenas A' de ADN 220 de longitud sustancialmente completa. Las cadenas A' de ADN 220 son el complemento a las cadenas A de la población clonal 120.

En otra etapa, la Figura 2C muestra el anclaje de las cadenas A' de ADN 220 completamente prolongadas al soporte sólido 110 por medio de restos unibles 215. La población clonal 120 (cadena A) se ha eliminado. En un ejemplo, la población clonal 120 se eliminó mediante una etapa de linealización (no mostrada).

En otra etapa, la Figura 2D muestra la hibridación de una diversidad de segundos cebadores de secuenciación 225 a las cadenas A' de ADN 220.

En otra etapa, la Figura 2E muestra la secuenciación de las cadenas A' de ADN 220 mediante la prolongación de los segundos cebadores de secuenciación 225 para obtener una segunda lectura de secuenciación 230 y producir información de secuenciación de extremos emparejados.

Las Figuras 3A a 3E ilustran vistas laterales del sustrato de secuenciación 100 de la Figura 1B y muestran otro ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido. En este ejemplo, se elimina un primer cebador de secuenciación que se usó para obtener una primera lectura de secuenciación de la cadena A, y se hibrida un oligonucleótido de prolongación que incluye un resto unible y se usa para generar una cadena A' complementaria de longitud sustancialmente completa. La cadena A' complementaria se ancla a la superficie del soporte sólido por medio del resto unible. En este ejemplo, un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido puede incluir, pero sin limitación, las etapas siguientes.

En una etapa, la Figura 3A muestra la hibridación de una diversidad de primeros cebadores de secuenciación 310 a la población clonal 120. En este ejemplo, la población clonal 120 comprende la cadena A.

En otra etapa, la Figura 3B muestra la prolongación de los primeros cebadores de secuenciación 310 para obtener una primera lectura de secuenciación en una reacción de secuenciación, tal como SBS, secuenciación por ligadura, o cualquier otro método de secuenciación adecuado.

En otra etapa, la Figura 3C muestra la eliminación de los primeros cebadores de secuenciación 310 y la primera lectura prolongada tras la finalización de la primera reacción de secuenciación con lectura.

En otra etapa, la Figura 3D muestra la hibridación de una diversidad de oligonucleótidos de prolongación 315 a la población clonal 120 (cadenas A). Cada oligonucleótido de prolongación 315 incluye un resto unible 320 que se

puede anclar (unir) de manera covalente o no covalente al soporte sólido 110 en ciertas condiciones.

5 En otra etapa, la Figura 3E muestra la prolongación de los oligonucleótidos de prolongación 315. Los oligonucleótidos de prolongación 315 se pueden prolongar en presencia de dNTPs y una ADN polimerasa para generar una diversidad de cadenas A' de ADN 325 de longitud sustancialmente completa. Las cadenas A' de ADN 325 son el complemento a las cadenas A de la población clonal 120.

En otra etapa, la Figura 3F muestra las cadenas A' de ADN 325 ancladas a la superficie del soporte sólido 110 por medio de los restos unibles 320. La población clonal 120 (cadenas A) se ha eliminado. En un ejemplo, la población clonal 120 se puede eliminar mediante una etapa de linealización (no mostrada).

10 En otra etapa, la Figura 3G muestra la hibridación de una diversidad de segundos cebadores de secuenciación 330 a las cadenas A' de ADN 325 de longitud completa que se anclan por medio de los restos unibles 320 a la superficie del soporte sólido 110.

En otra etapa, la Figura 3H muestra la prolongación de los cebadores de secuenciación 330 para obtener una segunda lectura de secuenciación. La primera lectura de secuenciación (Figura 3B) y la segunda lectura de secuenciación proporcionan la información de secuenciación de extremos emparejados.

15 Las Figuras 4A a 4E ilustran vistas laterales del sustrato de secuenciación 100 de la Figura 1B y muestran otro ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente está unido un tipo de cadena de ADN a la superficie de un soporte sólido. En este ejemplo, los nucleótidos modificados que incluyen un resto unible se incorporan en la porción más en 5' de la cadena A' complementaria. En otro ejemplo (no mostrado), los nucleótidos con restos unibles se pueden incorporar a una distancia más lejana del extremo 5'. La posición puede depender del tamaño del inserto y de la longitud de lectura. Por ejemplo, si el inserto es muy largo, es posible incorporar los nucleótidos modificados en una porción sustancialmente central, ya que las lecturas no alcanzarían la región central del inserto.

20

25 En ciertas realizaciones, puede ser deseable incluir los nucleótidos modificados cerca del extremo de la prolongación (extremo 3'), ya que esto puede mejorar la retención de las cadenas de ácido nucleico (especialmente en casos en los que las condiciones de la secuenciación pueden provocar una ruptura). Un método de colocación de los nucleótidos en el extremo 3' comprende incorporar dNTPs naturales para producir la cadena A' de longitud completa, y después añadir una polimerasa que tenga actividad 3'-5' exonucleasa en presencia de los nucleótidos modificados, de forma que los nucleótidos modificados se pueden incorporar en el propio extremo 3'.

30 En el ejemplo ilustrado en las Figuras 4A-4E, un proceso de obtención de información de secuenciación de extremos emparejados cuando solamente hay unido un tipo de cadena de ADN a una superficie sólida puede incluir, pero sin limitación, las etapas siguientes.

En una etapa, la Figura 4A muestra la hibridación de una diversidad de primeros cebadores de secuenciación 410 a la población clonal 120. En este ejemplo, la población clonal 120 comprende la cadena A.

35 En otra etapa, la Figura 4B muestra la prolongación de los primeros cebadores de secuenciación 410 y la incorporación de nucleótidos modificados que incluyen un resto unible 415 en una reacción de secuenciación, tal como SBS, secuenciación por ligadura, o cualquier otro método de secuenciación adecuado. Se pueden incorporar nucleótidos modificados que incluyen restos unibles 415, por ejemplo, durante los primeros ciclos de secuenciación. Las rondas posteriores de secuenciación se pueden llevar a cabo con nucleótidos que no incluyen un resto unible 415.

40 En otra etapa, la Figura 4C muestra la prolongación completa de los primeros cebadores de secuenciación 410 tras la finalización de la primera lectura de secuenciación. Los primeros cebadores de secuenciación 410 se pueden prolongar completamente en presencia de dNTPs naturales (p.ej., no incluyen un resto unible) y ADN polimerasa para generar una diversidad de cadenas A' de ADN 420 de longitud sustancialmente completa. Las cadenas A' de ADN 420 son el complemento a las cadenas A de la población clonal 120.

45 En otra etapa, la Figura 4D muestra la activación del resto unible 415 y la unión posterior de las cadenas A' de ADN 420 a la superficie del soporte sólido 110. El resto unible 415 se puede activar, por ejemplo, mediante métodos químicos, físicos o enzimáticos. Debido a que el resto unible 415 requiere una etapa de activación, se impide la unión de los nucleótidos modificados libres y la saturación de los sitios de unión de la superficie del soporte sólido 110. Entonces se puede eliminar la población clonal 120 (cadenas A). Entonces se unen las cadenas A' de ADN 420 al sustrato sólido 110. Después se hibrida una diversidad de segundos cebadores de secuenciación 425 a las cadenas A' de ADN 420.

50

En otra etapa, la Figura 4E muestra la prolongación de los segundos cebadores de secuenciación 425 para obtener una segunda lectura de secuenciación 430 y producir información de secuenciación de extremos emparejados.

55 Las Figuras 5A y 5B ilustran el método simplificado de la presente invención (mostrado en la Figura 5B) en comparación con los métodos de la técnica anterior. El método de la técnica anterior ilustrado en la Figura 5A

requiere un gran número de etapas para obtener la información de secuenciación de extremos emparejados.

Es evidente a partir de las etapas del método ilustrado en la Figura 5B que la presente invención proporciona un método mejorado, más simple y que consume menos tiempo. De hecho, las etapas de la Figura 5B mostradas en líneas discontinuas ("unión" y "eliminación de la cadena A") se pueden incorporar en ciclos de reacciones de secuenciación, lo que reduce adicionalmente el número de etapas y el tiempo necesario para llevar a cabo el método. La etapa de unión se podría llevar a cabo incluyendo una molécula de activación en los últimos ciclos de SBS, por ejemplo. En este caso, la SBS y la unión se llevan a cabo de manera simultánea, y de ese modo se elimina la necesidad de tiempo extra para llevar a cabo la etapa de unión como una etapa diferente. De forma similar, la eliminación de la cadena A mediante una etapa de linealización se puede llevar a cabo mediante el uso de una enzima de linealización o un producto químico incluido en los ciclos posteriores de SBS. El método mejorado proporciona por tanto ventajas significativas en el proceso de trabajo en comparación con los protocolos de la técnica anterior.

Las Figuras 6A a 6E ilustran vistas laterales del sustrato de secuenciación 100 de la Figura 1B y muestran un ejemplo de un proceso de obtención de información de secuenciación de extremos emparejados. En este ejemplo, el resto unible 615 del primer cebador de secuenciación 610 comprende una secuencia complementaria a un cebador de superficie 130 en la superficie del soporte sólido.

Como se muestra en la Figura 6A, la secuencia (resto unible 615) está unida en su extremo 5' al extremo 5' del cebador de secuenciación y orientada en la dirección opuesta, de forma que el oligonucleótido cebador 640 resultante, que comprende tanto el cebador de secuenciación 610 como la secuencia del resto unible 615, tiene dos extremos 3'.

El oligonucleótido cebador 640 se puede construir mediante la unión química del extremo 5' del cebador de secuenciación al extremo 5' de la secuencia complementaria al cebador de superficie, o por medio de uniones estreptavidina/biotina, por ejemplo, dos oligonucleótidos biotinilados en 5' unidos a estreptavidina.

La Figura 6A ilustra la hibridación de una diversidad de estos primeros cebadores oligonucleotídicos 640 "de doble final" por medio de la hibridación del componente de cebador de secuenciación 610 a la población clonal 120. En este ejemplo, la población clonal 120 comprende la cadena A. Cada cadena A está anclada al soporte sólido 110 mediante un cebador de superficie 130 en la superficie del soporte sólido. Como se muestra, varios cebadores de superficie 130 siguen estando disponibles para la hibridación tras la amplificación.

En general, un gran número de cebadores de superficie 130 siguen estando disponibles para la hibridación. En ciertas realizaciones, sin embargo, se pueden optimizar las condiciones de reacción para la etapa de hibridación para favorecer la hibridación de los cebadores oligonucleotídicos 640 a los cebadores de superficie 130. Por ejemplo, una temperatura mayor durante la hibridación favorece la hibridación del componente de cebador de secuenciación 610 al sitio asociado en la cadena A, ya que el sitio del cebador de secuenciación tiene una temperatura de fusión mayor que las secuencias más cortas de los cebadores de superficie 130.

En ciertas condiciones, los cebadores oligonucleotídicos 640 pueden hibridar a los cebadores de superficie 130 antes de la hibridación del componente de cebador de secuenciación 610 a un sitio asociado en una cadena A. Cualquier cebador oligonucleotídico 640 que hibride a cebadores de superficie 130 pero que no sea capaz de hibridar también a la cadena A no se prolongará durante la etapa de secuenciación y se podrá eliminar posteriormente, por ejemplo mediante el uso de una etapa de calentamiento.

De forma similar, se puede usar una polimerasa con desplazamiento de cadena en la etapa de prolongación para asegurar que se corrige cualquier hibridación indeseable de la secuencia del resto unible 615 a parte de la molécula que se está secuenciando durante la prolongación, de forma que la secuencia 615 se separa de la cadena de ácido nucleico y por lo tanto es libre de hibridar a un cebador de superficie 130 disponible.

En una realización, la unión indeseable del resto unible 615 a los cebadores de superficie 130 y/o las secuencias de población clonal se reduce proporcionando cebadores de superficie adicionales que tienen una secuencia complementaria al resto unible 615, pero cuya secuencia difiere de la de los cebadores de superficie 130 que unen la cadena A de la población clonal al soporte sólido.

Como se muestra en la Figura 6B, se pueden usar primeros cebadores de secuenciación 610, que tienen la secuencia del resto unible 615 unida al extremo 5', para obtener la primera lectura en una reacción de secuenciación, tal como secuenciación por síntesis (SBS), secuenciación por ligadura, o cualquier otro método de secuenciación adecuado.

La Figura 6C muestra la prolongación completa de los primeros cebadores de secuenciación 610 tras la finalización de la primera lectura de secuenciación. Los primeros cebadores de secuenciación 610 se pueden prolongar en presencia de dNTPs y ADN polimerasa para generar una diversidad de cadenas A' de ADN 220 de longitud sustancialmente completa. Las cadenas A' de ADN 220 son el complemento a las cadenas A de la población clonal 120.

5 Como se ilustra en la Figura 6D, la secuencia del resto unible 615 del cebador oligonucleotídico 640 se hibrida a un cebador de superficie 130 disponible del soporte sólido, por lo que se forma un puente tal como se muestra, y se ancla la cadena A' al soporte sólido. Una etapa de linealización (indicada en 650) linealiza el cebador de superficie 130 cerca de su extremo 3' para generar un extremo 5' en la cadena A 120. Después se puede eliminar la cadena A mediante el uso de una 5'-3' exonucleasa.

Las cadenas A' de ADN 220 completamente prolongadas se anclan así al soporte sólido 110 por medio de la hibridación de los restos unibles 615 a los cebadores de superficie 130. La población clonal 120 (cadena A) se ha eliminado. La Figura 6E muestra la hibridación de una diversidad de segundos cebadores de secuenciación 225 a las cadenas A' de ADN 220.

10 Las cadenas A' de ADN 220 se secuencian después mediante la prolongación de los segundos cebadores de secuenciación 225 para obtener una segunda lectura de secuenciación y proporcionar información de secuenciación de extremos emparejados.

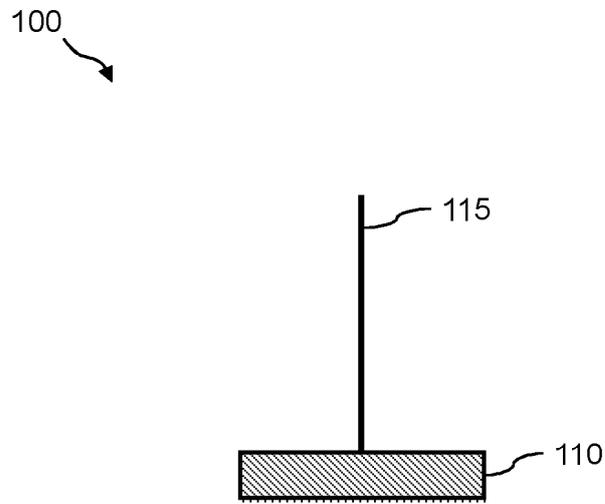
15 La realización de la invención ilustrada en las Figuras 6A a 6E proporciona la ventaja de poder usar un soporte sólido convencional, sin que sea necesario ningún procedimiento químico adicional en la superficie para anclar la cadena A' a la superficie.

20 Las Figuras 7A y 7B ilustran una realización en la que el método ilustrado en las Figuras 6A a 6E se puede llevar a cabo mediante el uso de un cebador oligonucleotídico 740 alternativo que comprende una secuencia complementaria al cebador de superficie (resto unible 615) unida en su extremo 3' al extremo 5' de un cebador de secuenciación y orientada en la misma dirección, de forma que el cebador oligonucleotídico 240 resultante, que comprende tanto el cebador de secuenciación 610 como la secuencia del resto unible 615, tiene un extremo 5' y un extremo 3'.

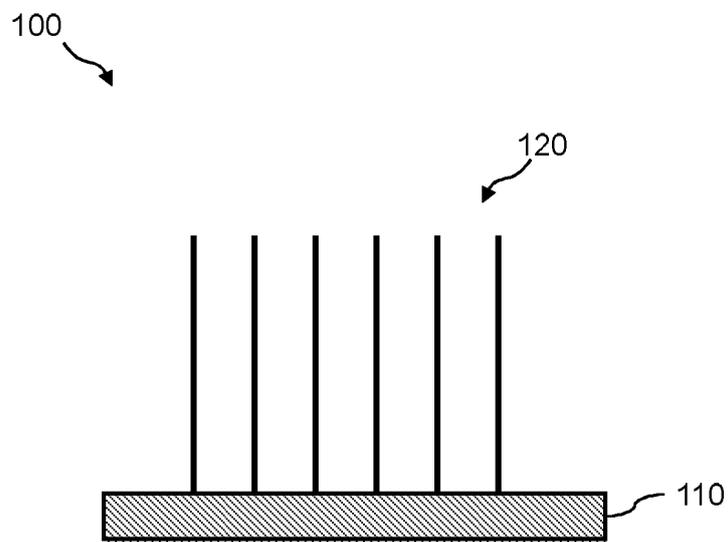
La Figura 7B ilustra con más exactitud la etapa del método mostrado y descrito con respecto a la Figura 6D, cuando se usa el oligonucleótido cebador alternativo de la Figura 7A.

**REIVINDICACIONES**

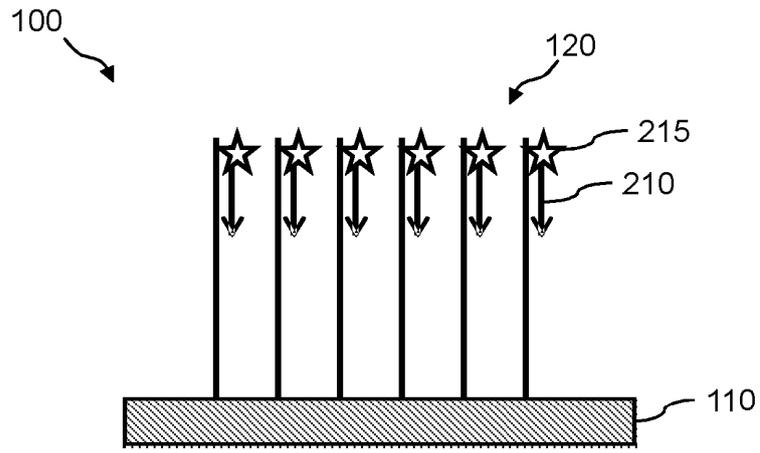
- 5 1. Un método de obtención de información de secuenciación de extremos emparejados en una reacción de secuenciación que comprende las etapas de: proporcionar un ácido nucleico objetivo inmovilizado en un soporte sólido; determinar la secuencia del ácido nucleico objetivo generando una cadena de ácido nucleico complementaria, y la cadena de ácido nucleico complementaria comprende como mínimo un resto unible adecuado para la inmovilización; determinar la secuencia de la cadena de ácido nucleico complementaria; y comparar las secuencias del ácido nucleico objetivo y de la cadena de ácido nucleico complementaria para proporcionar la información de secuenciación de extremos emparejados.
- 10 2. El método de la reivindicación 1, en el que el como mínimo un resto unible se proporciona en un primer cebador de secuenciación.
3. El método de la reivindicación 1 o 2, en el que el como mínimo un resto unible se proporciona en un oligonucleótido de prolongación.
4. El método de cualquier reivindicación precedente, en el que el como mínimo un resto unible se proporciona en un nucleótido modificado.
- 15 5. El método de cualquier reivindicación precedente, en el que el soporte sólido comprende como mínimo un resto unible correspondiente para unir selectivamente de manera covalente o no covalente el como mínimo un resto unible de la cadena de ácido nucleico complementaria en condiciones predeterminadas, opcionalmente en el que el como mínimo un resto unible correspondiente comprende un oligonucleótido de superficie en el soporte sólido y el como mínimo un resto unible comprende una secuencia de ácido nucleico sustancialmente complementaria al oligonucleótido de superficie, que comprende opcionalmente la etapa de activar el como mínimo un resto unible para unir la cadena de ácido nucleico complementaria al soporte sólido.
- 20 6. El método de cualquier reivindicación precedente, en el que el método comprende la etapa de eliminación del ácido nucleico objetivo tras la etapa de determinación de su secuencia.
- 25 7. El método de cualquier reivindicación precedente, en el que la etapa de determinación de la secuencia del ácido nucleico objetivo comprende hibridar un primer cebador de secuenciación al ácido nucleico objetivo y prolongar el primer cebador de secuenciación para obtener una primera lectura de secuenciación.
8. El método de la reivindicación 7, en el que el primer cebador de secuenciación se prolonga hasta sustancialmente la longitud completa del ácido nucleico objetivo tras la primera lectura de secuenciación.
- 30 9. El método de cualquier reivindicación precedente, en el que la etapa de determinación de la secuencia de la cadena de ácido nucleico complementaria comprende hibridar un segundo cebador de secuenciación a la cadena de ácido nucleico complementaria y prolongar el segundo cebador de secuenciación para obtener una segunda lectura de secuenciación.
10. El método de las reivindicaciones 7 a 9, en el que la lectura de secuenciación se obtiene mediante el uso de secuenciación por síntesis o secuenciación por ligadura.
- 35 11. El método de cualquier reivindicación precedente, en el que la etapa de comparación de las secuencias del ácido nucleico objetivo y de la cadena de ácido nucleico complementaria para proporcionar la información de secuenciación de extremos emparejados utiliza un algoritmo y/o un procesador informático.
12. El método de cualquier reivindicación precedente, que comprende la pre-etapa de amplificación clonal de la molécula de ácido nucleico objetivo.
- 40 13. Un sistema para obtener información de secuenciación de extremos emparejados según cualquiera de las reivindicaciones 1 a 12, que comprende un aparato de secuenciación que tiene un soporte sólido para inmovilizar un ácido nucleico objetivo, una entrada y una salida para los reactivos de secuenciación, y medios para la determinación de la secuencia del ácido nucleico objetivo generando una cadena de ácido nucleico complementaria, y la cadena de ácido nucleico complementaria comprende como mínimo un resto unible adecuado para la inmovilización; determinar la secuencia de la cadena de ácido nucleico complementaria; y comparar las secuencias del ácido nucleico objetivo y de la cadena de ácido nucleico complementaria para proporcionar la información de secuenciación de extremos emparejados.
- 45



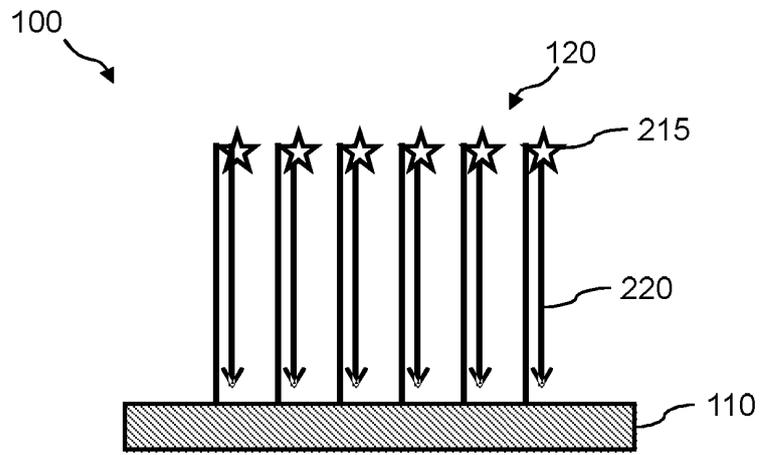
**Figura 1A**



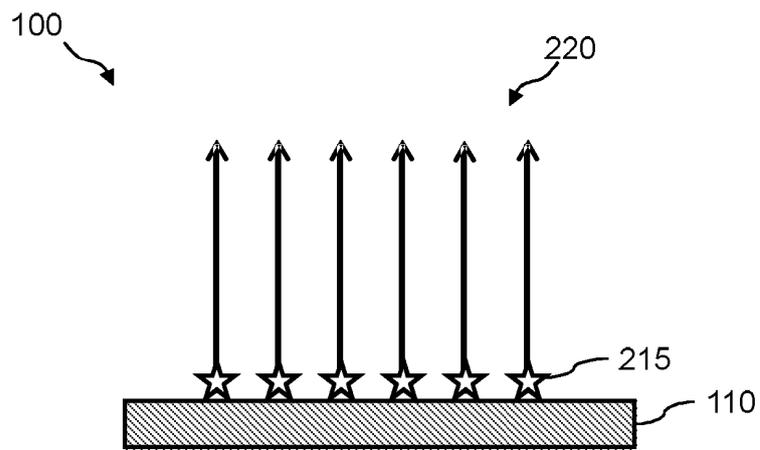
**Figura 1B**



**Figura 2A**



**Figura 2B**



**Figura 2C**

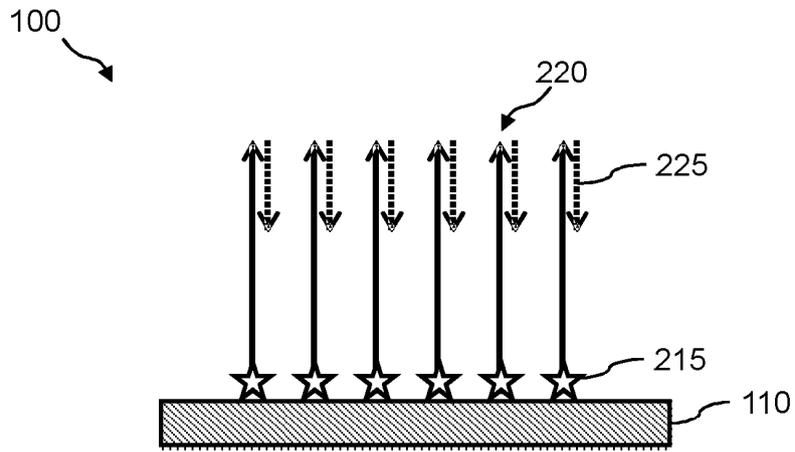


Figura 2D

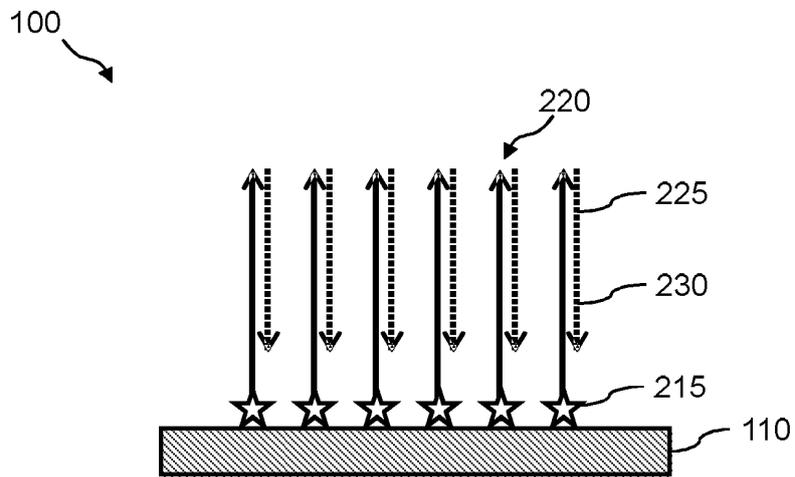
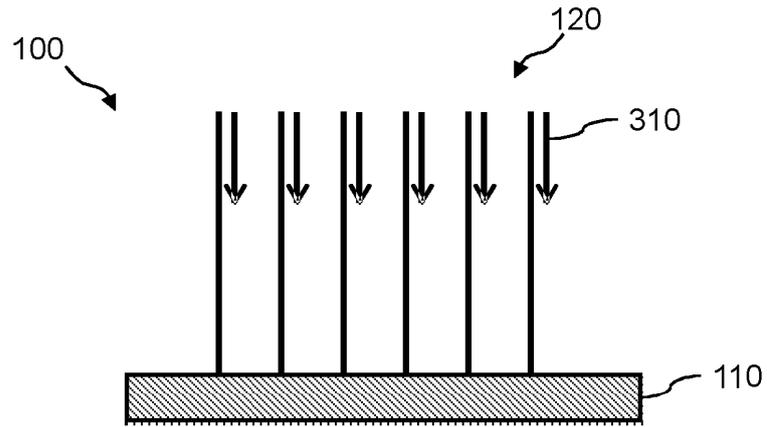
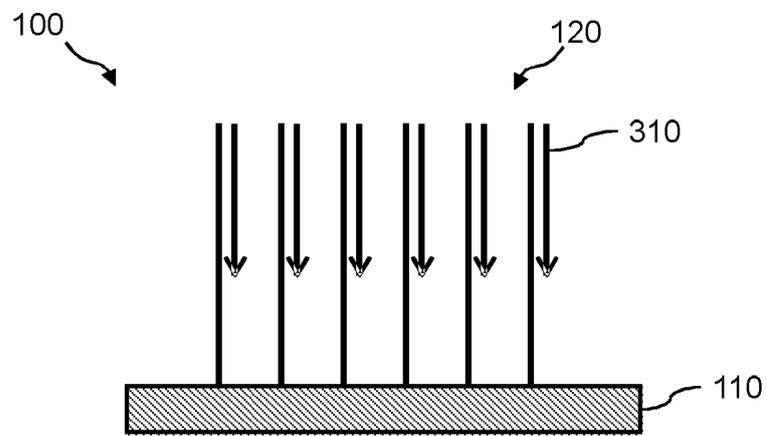


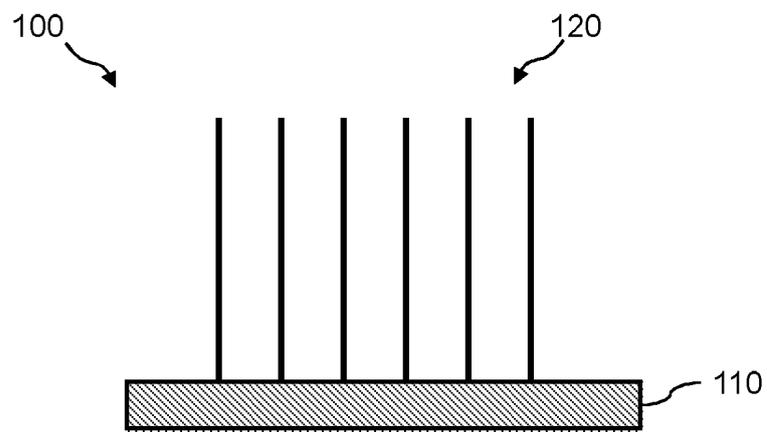
Figura 2E



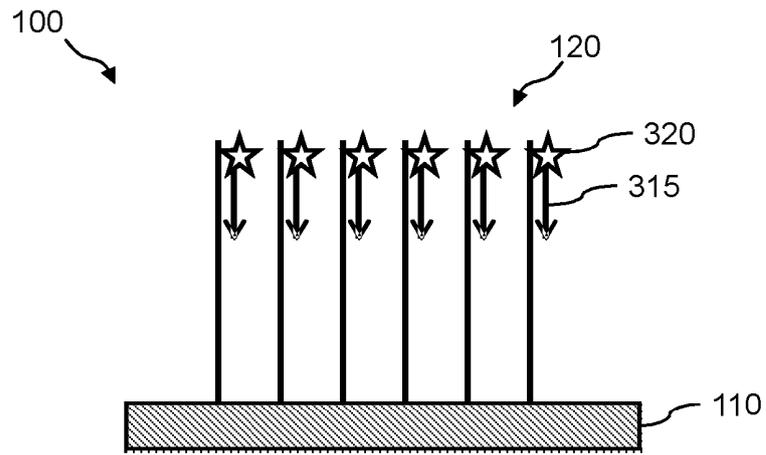
**Figura 3A**



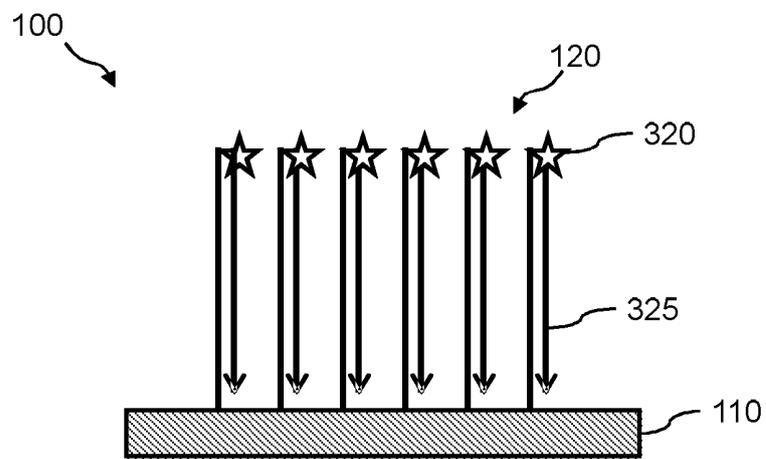
**Figura 3B**



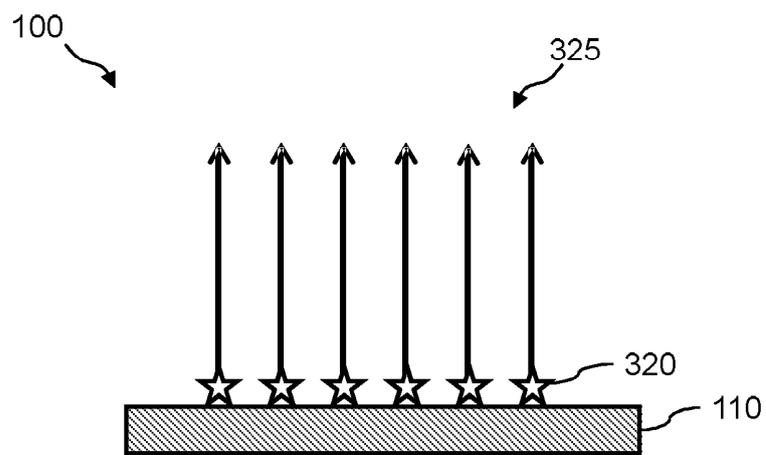
**Figura 3C**



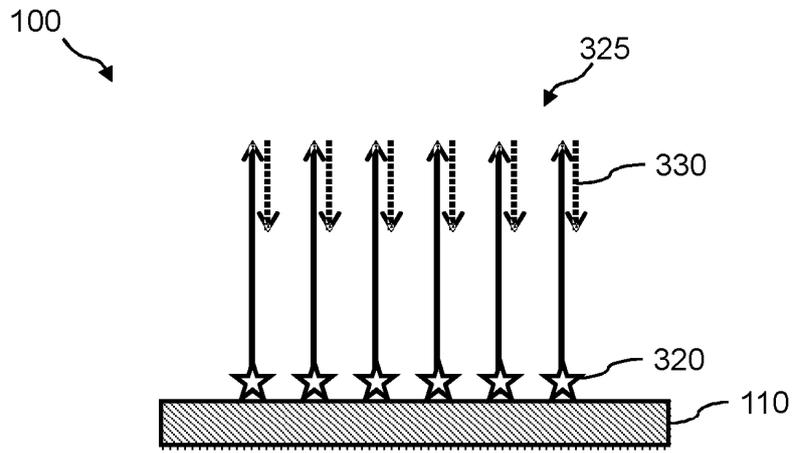
**Figura 3D**



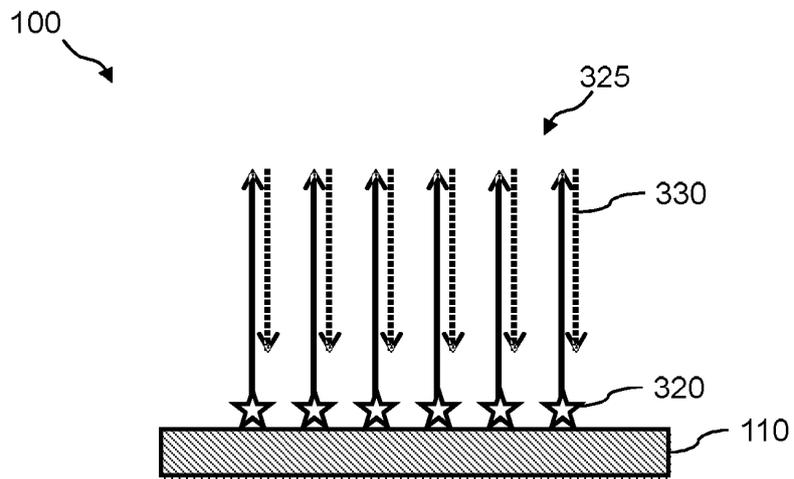
**Figura 3E**



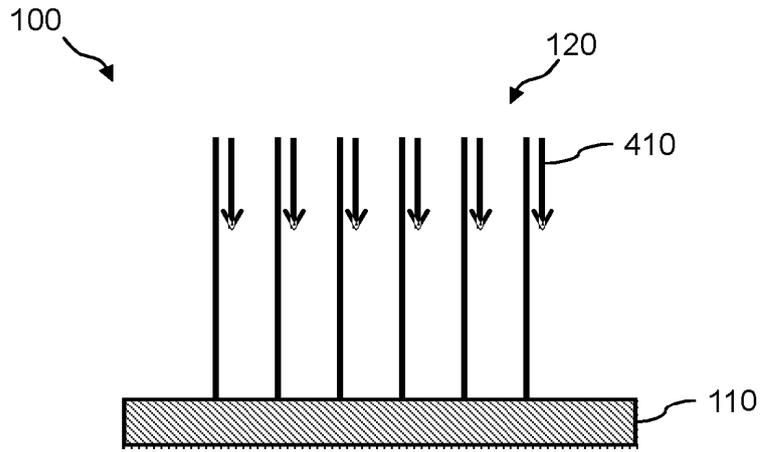
**Figura 3F**



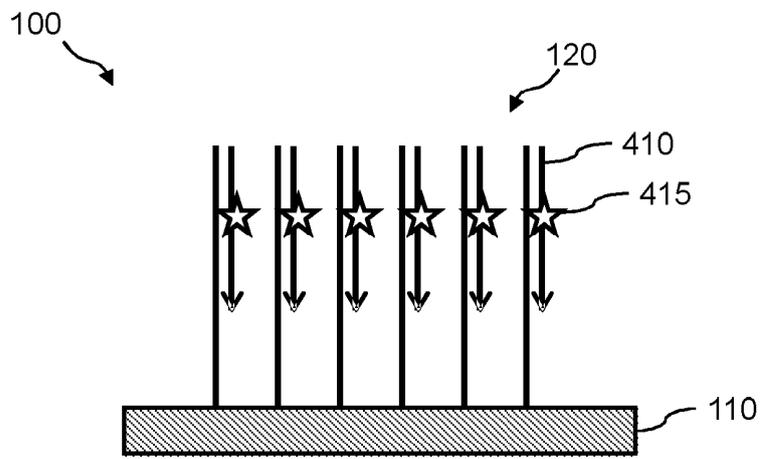
**Figura 3G**



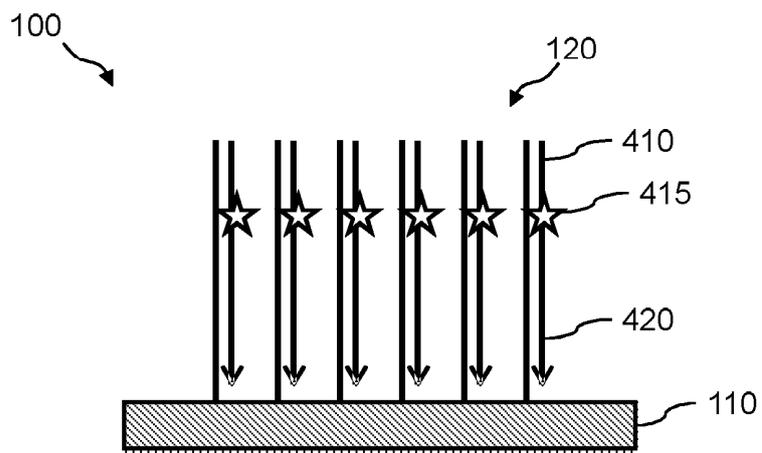
**Figura 3H**



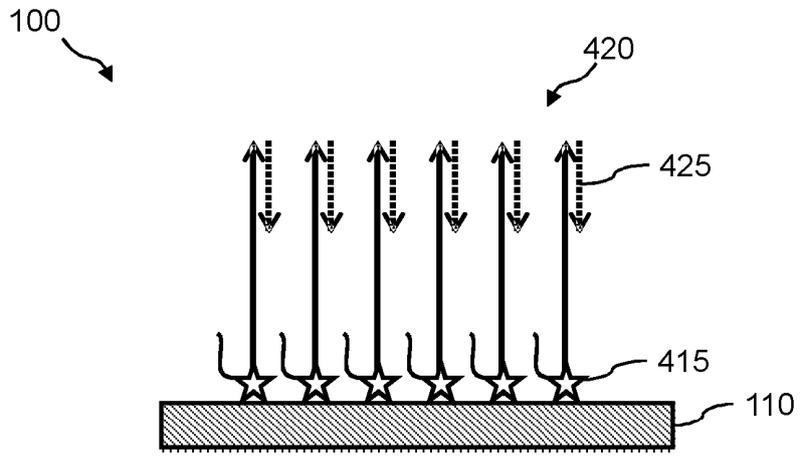
**Figura 4A**



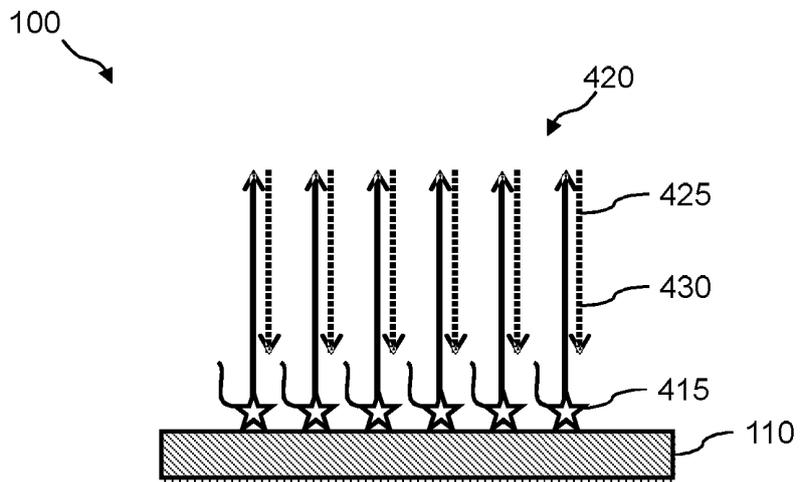
**Figura 4B**



**Figura 4C**



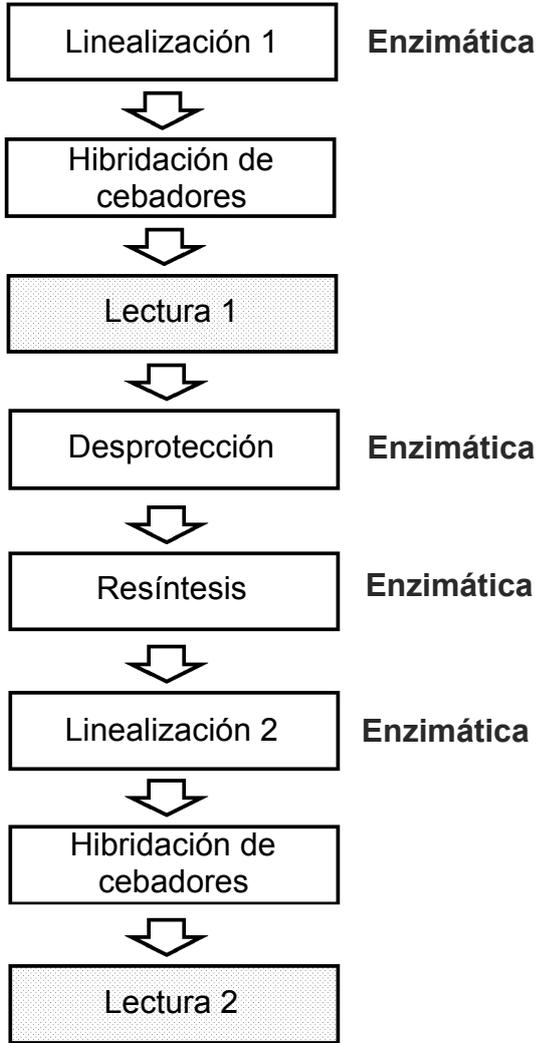
**Figura 4D**



**Figura 4E**

**Fig 5 A**

**Técnica anterior**



**Fig 5B**

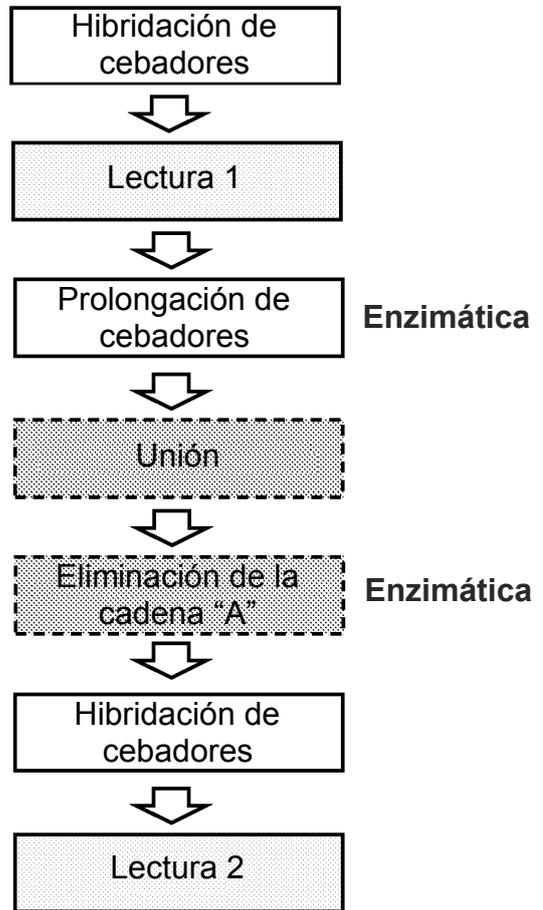
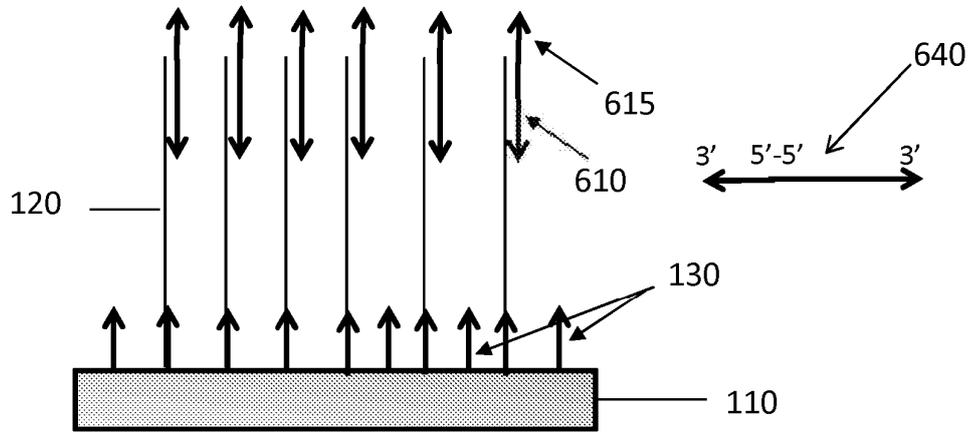
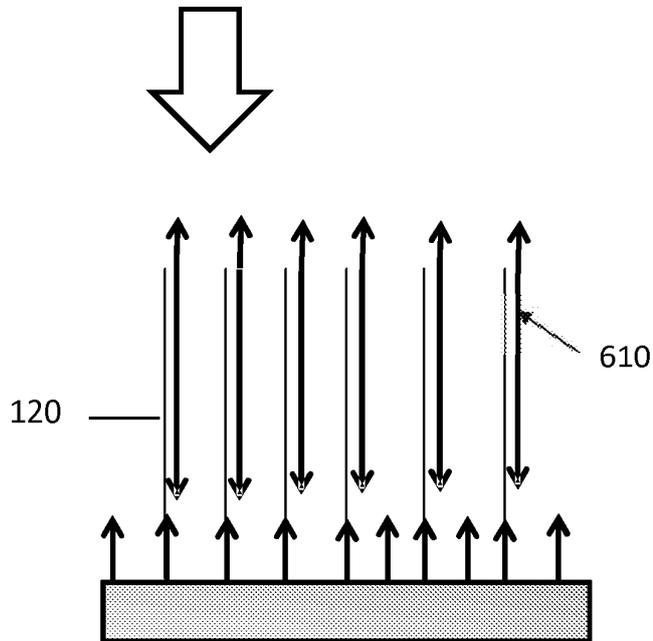


Figura 6

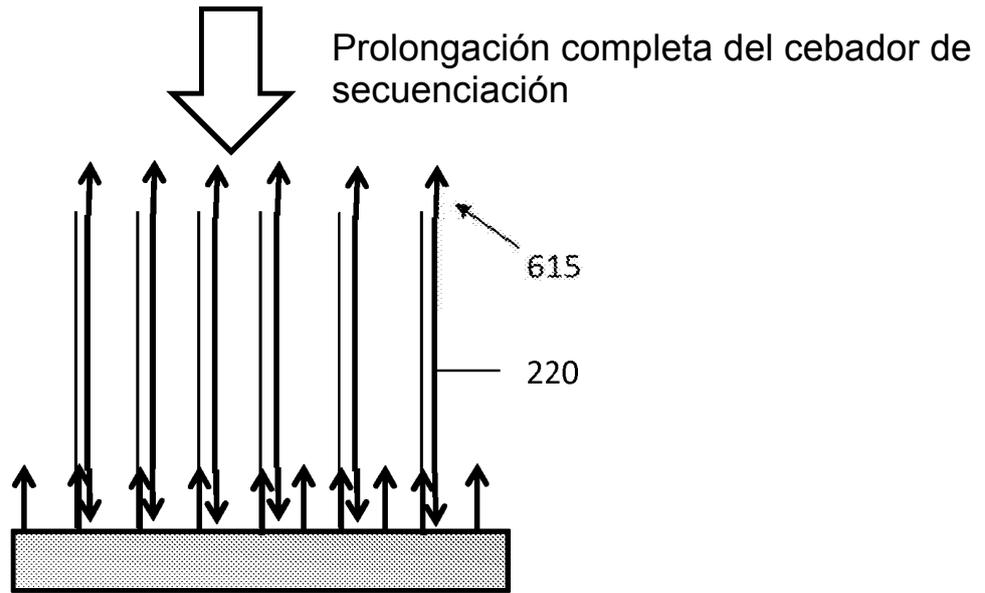
**A**



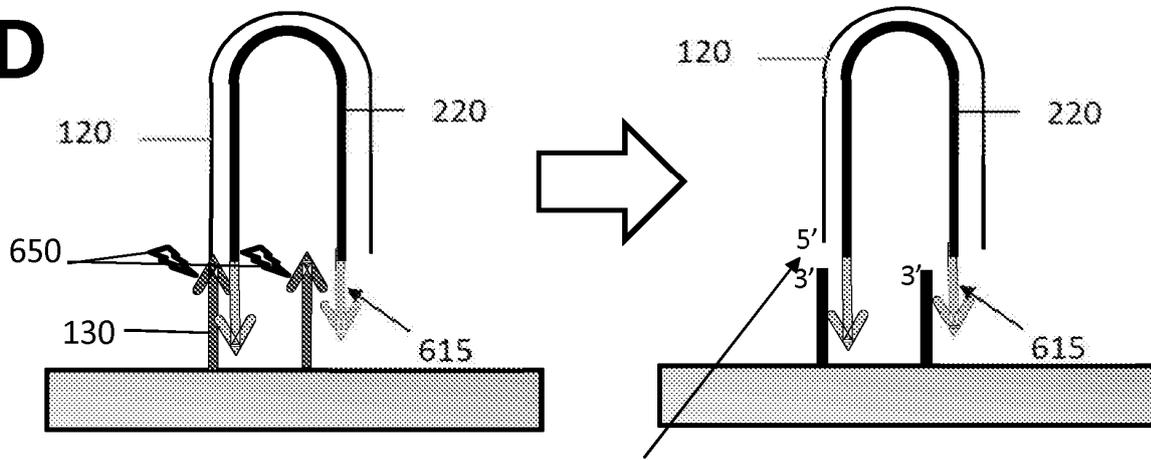
**B**



**C**



**D**



**E**

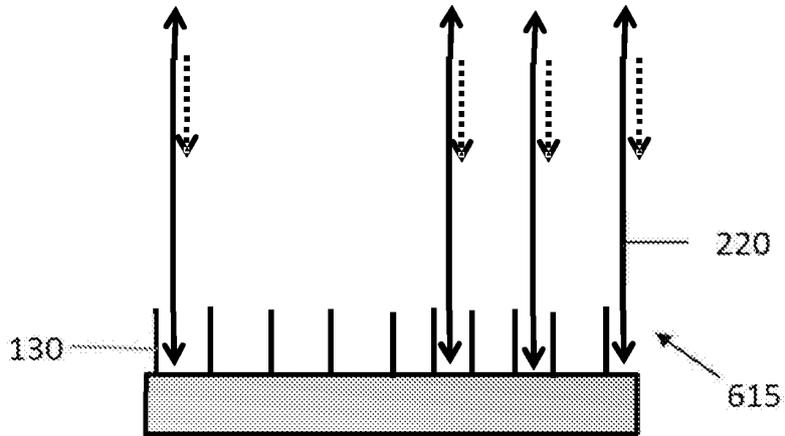
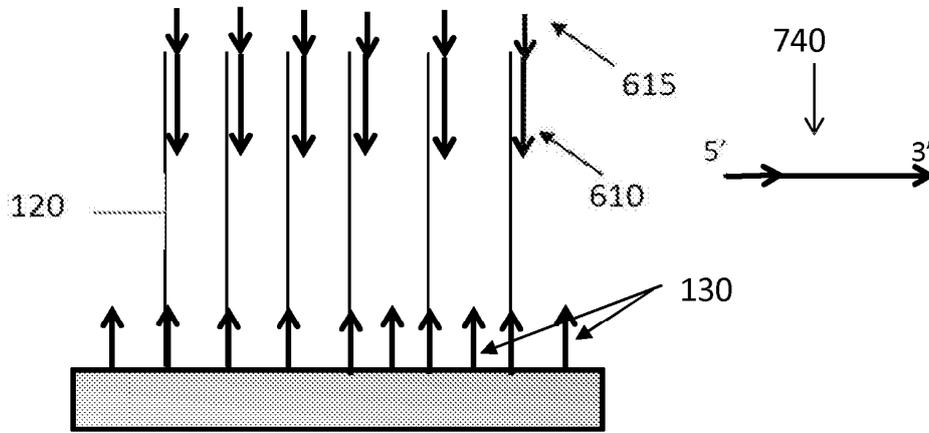


Figura 7

**A**



**B**

