



①9



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA

①1 Número de publicación: **2 310 514**

⑤1 Int. Cl.:
C12Q 1/68 (2006.01)

①2

TRADUCCIÓN DE PATENTE EUROPEA

T3

- ⑨6 Número de solicitud europea: **00909464 .0**
⑨6 Fecha de presentación : **10.03.2000**
⑨7 Número de publicación de la solicitud: **1159453**
⑨7 Fecha de publicación de la solicitud: **05.12.2001**

⑤4 Título: **Un método para la secuenciación directa de ácido nucleico.**

③0 Prioridad: **10.03.1999 US 266187**

④5 Fecha de publicación de la mención BOPI:
16.01.2009

④5 Fecha de la publicación del folleto de la patente:
16.01.2009

⑦3 Titular/es: **ASM Scientific, Inc.**
240 Norfolk Street
Cambridge, Massachusetts 02139, US

⑦2 Inventor/es: **Stemple, Derek Lyle y**
Armes, Niall Antony

⑦4 Agente: **Elzaburu Márquez, Alberto**

ES 2 310 514 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

Un método para la secuenciación directa de ácido nucleico.

5 **Campo de la invención**

La presente invención se refiere a métodos para secuenciar muestras de ácido nucleico. Más específicamente, la presente invención se refiere a métodos para secuenciar sin necesidad de amplificación; el conocimiento anterior de algunas de la secuencia de nucleótidos para generar los cebadores de secuenciación; y las largas técnicas de electroforesis.

10 **Antecedentes de la invención**

La secuenciación de muestras de ácido nucleico es una importante técnica analítica en la moderna biología molecular. El desarrollo de métodos fiables para la secuenciación de ADN ha sido crucial para entender la función y el control de genes y para aplicar muchas de las técnicas básicas de biología molecular. Estos métodos también se han hecho cada vez más importantes como herramientas en el análisis genómico y en muchas aplicaciones no relacionadas con la investigación, tales como la identificación genética, el análisis forense, el asesoramiento genético, el diagnóstico médico y muchas otras. En estas últimas aplicaciones, han sido empleadas ambas técnicas que proporcionan información de la secuencia parcial, tales como la toma de huellas digitales y las comparaciones de secuencias, y técnicas que proporcionan la determinación de la secuencia completa. Véase, por ejemplo, Gibbs *et al.*, *Proc. Natl. Acad. Sci. USA* **86**: 1919-1923 (1989); Gyllensten *et al.*, *Proc. Natl. Acad. Sci. USA* **85**: 7652-7656 (1988); Carrano *et al.*, *Genomics* **4**: 129-136 (1989); Caetano-Annoles *et al.*, *Mol. Gen. Genet.* **235**: 157-165 (1992); Brenner y Livak, *Proc. Natl. Acad. Sci. USA* **86**: 8902-8906 (1989); Green *et al.*, *PCR Methods and Applications* **1**: 77-90 (1991); y Versalovic *et al.*, *Nucleic Acids Res.* **19**:6823-6831 (1991).

Los métodos de secuenciación de ADN más actualmente disponibles requieren la generación de un grupo de fragmentos de ADN que se ordenan por longitudes de acuerdo con la composición de nucleótidos. La generación de este grupo de fragmentos ordenados se hace en alguno de los dos modos siguientes: (1) degradación química en nucleótidos específicos usando el método de Maxam-Gilbert o (2) la incorporación de didesoxi-nucleótidos usando el método de Sanger. Véase Maxam y Gilbert, *Proc. Natl. Acad. Sci. USA* **74**:560-564 (1977); Sanger *et al.*, *Proc. Natl. Acad. Sci. USA* **74**: 5463-5467 (1977). El tipo y el número de etapas requeridas limitan intrínsecamente tanto el número de segmentos de ADN que pueden ser secuenciados en paralelo, como la cantidad de secuencia que puede ser determinada desde un sitio dado. Además, ambos métodos son propensos al error debido a la migración anómala de fragmentos de ADN en geles desnaturizantes. Las limitaciones de tiempo y espacio inherentes a estos métodos basados en geles han forzado a la búsqueda de métodos alternativos.

En un esfuerzo para satisfacer las demandas de secuenciación actuales a gran escala, se han hecho mejoras al método de Sanger. Por ejemplo, el uso de terminadores de cadena fluorescentes simplifica la detección de los nucleótidos. La síntesis de fragmentos de ADN más largos y la mejor resolución de los fragmentos produce más información de la secuencia en cada experimento. El análisis automatizado de fragmentos en geles o capilares ha reducido considerablemente el trabajo implicado en la recogida y el tratamiento de la información de la secuencia. Véase, por ejemplo, Prober *et al.*, *Science* **238**:336-341 (1987); Smith *et al.*, *Nature* **321**:674-679 (1986); Luckey *et al.*, *Nucleic Acids Res.* **18**:4417-4421 (1990); Dovichi, *Electrophoresis* **18**: 2393-2399 (1997).

Sin embargo, las tecnologías de secuenciación de ADN actuales todavía comportan tres limitaciones principales. Primero, requieren una gran cantidad de moléculas de ADN idénticas, que generalmente son obtenidas por clonación molecular o amplificación por reacción en cadena de la polimerasa (PCR) de secuencias de ADN. Los métodos actuales de detección son insensibles y por eso requieren un número crítico mínimo de oligonucleótidos marcados. También, muchas copias idénticas del oligonucleótido son necesarias para generar una escala de secuencia. Una segunda limitación es que las técnicas de secuenciación actuales dependen del cebado de oligodesoxinucleótidos específico de secuencias que deben ser sintetizadas antes de la iniciación del procedimiento de secuenciación. Sanger y Coulson, *J. Mol. Biol.* **94**:441-448 (1975). La necesidad de múltiples plantillas idénticas hace necesario el cebado sincrónico de cada copia a partir del mismo sitio predeterminado. Tercero, las técnicas de secuenciación actuales dependen de técnicas de electroforesis muy largas y que requieren de mucho trabajo, que son limitadas por la velocidad a la cual puedan ser separados los fragmentos y también están limitadas por el número de bases que puedan ser secuenciadas en un experimento dado por la resolución obtenible en el gel.

En un esfuerzo para prescindir de la necesidad de las técnicas de electroforesis, fue desarrollado un método de secuenciación que usa terminadores de cadena que pueden ser liberados, o desprotegidos, para una posterior extensión. Véase, la patente de EE.UU. N°. 5.302.509; Metzker *et al.*, *Nucleic Acids Res.* **22**:4259-4267 (1994). Este método implica ciclos repetitivos de incorporación de bases, detección de la incorporación y la reactivación del terminador de cadena para permitir el siguiente ciclo de síntesis de ADN. Así, detectando cada base añadida mientras la cadena de ADN crece, se elimina la necesidad del fraccionamiento por tamaños. Este método es, sin embargo, todavía altamente dependiente de grandes cantidades del ácido nucleico que se secuencia y del uso de secuencias conocidas para cebar la iniciación del crecimiento de la cadena. Además, esta técnica está perjudicada por cualquier ineficiencia de incorporación y desprotección. Como la incorporación y la regeneración de 3'-OH no son completamente eficientes,

un grupo de cadenas que se extienden inicialmente idénticas puede volverse rápidamente asincrónica y las secuencias no pueden ser resueltas más allá de unas adiciones iniciales limitadas.

El documento WO 99/05315 A2 (Medical Biosystems Ltd., *et al.*) describe un método que comprende hacer reaccionar un polinucleótido objetivo con una enzima polimerasa inmovilizada sobre un soporte sólido, y los nucleótidos diferentes, en condiciones suficientes para que se produzca la reacción de la polimerasa, y detectar un efecto consiguiente en la incorporación de un nucleótido específico complementario al polinucleótido objetivo.

De acuerdo con el documento WO 90/13666 A1 (Amersham Internacional PLC, *et al.*) un complejo inmovilizado de una plantilla que se secuenciar y un cebador es expuesto para fluir conteniendo sólo un dNTP a la vez. Cada acontecimiento de polimerización es detectado, preferiblemente directamente y por medios espectroscópicos.

El documento WO 93/21340 A1 (Medical Research Council, *et al.*) describe un método para determinar la secuencia de un ácido nucleico que comprende formar una plantilla monocatenaria que comprende el ácido nucleico que se secuenciar, hibridar un cebador a la plantilla para formar un complejo de plantilla/cebador, ampliar el cebador por la adición de un nucleótido marcado individual, determinar el tipo del nucleótido marcado añadido en el cebador, eliminar o neutralizar el marcador y luego repetir tales etapas secuencialmente y registrar el orden de incorporación de nucleótidos marcados. Los cuatro tipos diferentes de nucleótido son añadidos secuencialmente.

El documento WO 00/36152 A1 (Li-Cor, Inc., *et al.*) describe métodos para la secuenciación y el genotipado del ácido nucleico en una configuración de molécula individual mediante la detección de una molécula individual de restos de PPI fluorescentes marcados liberados a partir de NTP según se crea un producto de extensión de la polimerasa.

Por consiguiente, existe todavía una necesidad en la técnica de un método de alto rendimiento, rápido y rentable para secuenciar muestras de ácido nucleico desconocidas que elimine la necesidad de la amplificación; el conocimiento anterior de algunas de la secuencia de nucleótidos para generar cebadores de secuenciación; y técnicas de electroforesis en las que se emplea mucha mano de obra.

Sumario de la invención

De acuerdo con la presente invención se proporciona por lo tanto un método para secuenciar bases de nucleótidos que comprende las etapas secuenciales de:

- (a) introducir en una cámara de reacción que contiene una polimerasa inmovilizada sobre un soporte sólido una muestra de ácido nucleico, una pluralidad de cebadores de oligonucleótido diferentes, y una solución tamponada que contiene los cuatro nucleótidos diferentes, siendo cada nucleótido diferencialmente marcado con un grupo marcador desprendible y bloqueado en la porción 3' con un grupo bloqueante desprendible, en el que dicha muestra de ácido nucleico se hibrida a un cebador, y dicha polimerasa extiende dicho cebador incorporando un nucleótido diferencialmente marcado que es complementario al ácido nucleico de la muestra;
- (b) eliminar los nucleótidos que no han sido incorporados en el cebador;
- (c) determinar la identidad del nucleótido incorporado en el cebador de alargamiento detectando su grupo marcador, identificando así el complemento del nucleótido 3'-bloqueado marcado;
- (d) separar el grupo bloqueante 3' y el grupo marcador a partir del nucleótido incorporado;
- (e) eliminar el grupo bloqueante 3' separado y el grupo marcador separado de la etapa (d);
- (f) confirmar la separación y la eliminación del grupo bloqueante 3' a partir del nucleótido incorporado en el cebador monitorizando la presencia de nucleótidos marcados en la cámara de reacción mediante una segunda etapa de detección en la que la presencia del nucleótido marcado en la cámara de reacción indica que el grupo bloqueante 3' no ha sido eliminado;
- (g) introducir en la cámara de reacción solución tamponada recién preparada que contenga dichos cuatro diferentes nucleótidos para iniciar de nuevo la síntesis; y
- (h) repetir las etapas de (b) a (g) hasta que ningún nuevo nucleótido sea incorporado en la etapa (g) o el grupo bloqueante 3' persista en no ser separado y eliminado en las etapas (d) y (e);

por lo cual el orden en el cual los nucleótidos marcados en las etapas (b) son detectados corresponde al complemento de la secuencia de al menos una parte de dicha muestra de ácido nucleico.

La presente invención proporciona así métodos de rendimiento alto, rápidos y rentables para secuenciar muestras desconocidas de ácido nucleico que eliminan la necesidad de la amplificación; el conocimiento anterior de alguna de la secuencia de nucleótidos para generar cebadores de secuenciación; y técnicas de electroforesis que requieran

mucho trabajo. Los métodos de la presente invención permiten la Secuenciación de Ácido Nucleico Directa (SAND) de moléculas individuales de ácido nucleico.

De acuerdo con los métodos de la presente invención, una pluralidad de moléculas de polimerasa pueden ser inmovilizadas sobre un soporte sólido a través de una interacción covalente o no covalente. Una muestra de ácido nucleico y cebadores de oligonucleótido son introducidos en la cámara de reacción en una solución tamponada que contiene los cuatro terminadores de trifosfato de nucleósido marcados-atrapados. La elongación dirigida por la plantilla de un ácido nucleico es mediada por las polimerasas unidas usando los terminadores de trifosfato de nucleósido marcados-atrapados. Los centros de reacción pueden ser monitorizados por el sistema microscópico hasta que una mayoría de sitios contenga la polimerasa inmovilizada unida a una plantilla de ácido nucleico con un terminador de nucleótido marcado-atrapado individual incorporado. Entonces la cámara de reacción puede ser enjuagada con un tampón de lavado. La incorporación del nucleótido específico entonces se determina para cada centro de reacción activo. Después de la detección, la cámara de reacción puede ser irradiada para liberar el nucleótido incorporado y enjuagada con tampón de lavado otra vez. La presencia de nucleótidos marcados-atrapados se monitoriza otra vez antes de que sean añadidos reactivos recién preparados para iniciar de nuevo la síntesis, para verificar que los centros de reacción son liberados satisfactoriamente. Un fracaso persistente de la liberación o incorporación, sin embargo, indica el fracaso de un centro de reacción. Un fracaso persistente de la liberación o incorporación consiste en 2-20 ciclos, preferiblemente 3-10 ciclos, más preferiblemente 3-5 ciclos, en el que la presencia de un nucleótido marcado-atrapado es detectada durante la segunda etapa de detección, indicando que el centro de reacción no fue liberado satisfactoriamente. El ciclo de secuenciación referido anteriormente es repetido hasta que una proporción grande de centros de reacción falla.

Los nucleótidos diferencialmente marcados usados en los métodos de secuenciación de la presente invención tienen un grupo marcador desprendible y son bloqueados en la porción 3' con un grupo bloqueante desprendible. En una realización preferida, el grupo marcador es unido directamente al grupo bloqueante 3' desprendible. La liberación de los nucleótidos puede ser lograda enzimáticamente, químicamente o, preferiblemente, fotolíticamente dependiendo del enlazante desprendible usado para unir el grupo marcador y el grupo bloqueante 3' al nucleótido.

En otra realización preferida, el grupo marcador es unido a la base de cada nucleótido con un enlazante desprendible más bien que al grupo bloqueante 3' desprendible. El grupo marcador y el grupo bloqueante 3' pueden ser eliminados enzimáticamente, químicamente o fotolíticamente. La alternativa, el grupo marcador puede ser eliminado por un método diferente y el grupo bloqueante 3'. Por ejemplo, el grupo marcador puede ser eliminado enzimáticamente mientras que el grupo bloqueante 3' es eliminado químicamente, o por activación fotoquímica.

Muchas reacciones independientes ocurren simultáneamente dentro de la cámara de reacción, generando cada centro de reacción individual unos cien, o miles, de pares de bases. Este aparato tiene la capacidad de secuenciar en paralelo miles y, posiblemente, millones de plantillas separadas a partir de puntos de secuencia especificados o aleatorios. La secuencia combinada de cada ciclo está sobre el orden de varios millones de pares de bases de secuencia y no requiere la amplificación, el conocimiento anterior de una parte de la secuencia objetivo o la resolución de fragmentos en geles o capilares. Las preparaciones de ADN simples a partir de cualquier fuente pueden ser secuenciadas con los aparatos y métodos de la presente invención.

Breve descripción de los dibujos

La Figura 1 (Paneles A-C) es una representación esquemática de nucleótidos terminadores marcados-atrapados para su uso en la secuenciación de ácido nucleico directa. El Panel A representa un desoxiadenosin-trifosfato modificado por la unión de un conjugado de fluorocromo-enlazante fotolábil al carbono 3' de la ribosa. El Panel B representa una configuración alternativa, en la que el fluorocromo está unido a la base del nucleótido por vía de un enlazante fotolábil. El Panel C representa los cuatro nucleótidos diferentes cada uno marcado con un fluorocromo con propiedades espectrales distintas, lo que permite que los cuatro nucleótidos sean distinguidos durante la fase de detección de un ciclo de reacción de secuenciación de ácido nucleico directo.

La figura 2 es una representación esquemática de las etapas de un ciclo de secuenciación de ácido nucleico directo, en el que la etapa 1 ilustra la incorporación de un nucleótido marcado-atrapado, la etapa 2 ilustra la detección del marcador, y la etapa 3 ilustra el desbloqueo del compartimiento de 3'-OH.

La figura 3 es una representación esquemática de un centro de reacción que representa una polimerasa inmovilizada y una muestra de ácido nucleico que se secuencian.

La figura 4 es una representación esquemática del ensamblaje de la cámara de reacción que alberga la matriz de centros de reacción de SAND y media el intercambio de reactivos y tampón.

La Figura 5 es una representación esquemática de una matriz de centro de reacción. El panel del lado izquierdo (Campo de Microscopio) representa la vista de una matriz entera según se registra por cuatro acontecimientos de detección sucesivos (uno para cada uno de los fluorocromos separados). El panel del centro representa una vista ampliada de una parte del campo mostrando el espacio de centros de reacción individuales. El panel lejano derecho representa la vista de la cámara de un centro de reacción solo.

La figura 6 es una representación esquemática del principio de la onda evanescente.

La figura 7 es una representación esquemática de una secuenciación de ácido nucleico directa establecida utilizando microscopía de fluorescencia de reflexión total interna.

La figura 8 es una representación esquemática de un ejemplo de un algoritmo de adquisición de datos obtenido a partir de una matriz 3x3.

Descripción detallada de la invención

El método de la presente invención, denominado en este documento Secuenciación de Ácido Nucleico Directa (SAND), ofrece un método de rendimiento alto, rápido y rentable por el cual pueden ser secuenciadas fácilmente moléculas de ácido nucleico a partir de cualquier fuente sin necesidad de una amplificación anterior. La SAND puede ser usada para determinar la secuencia de nucleótido de numerosas moléculas de ácido nucleico individuales en paralelo.

1. Matriz de centro de reacción de SAND

Las polimerasas se unen al soporte sólido, espaciadas a intervalos regulares, en una matriz de centros de reacción, presentes a una periodicidad mayor que el poder de resolución óptico del sistema microscópico. Preferiblemente, sólo una molécula de polimerasa está presente en cada centro de reacción, y cada centro de reacción se localiza a una distancia ópticamente resoluble desde los otros centros de reacción.

Las reacciones de secuenciación ocurren preferiblemente en una cámara de reacción acuosa delgada que comprende una deslizadera de cubierta sellada y un soporte sólido ópticamente transparente.

La inmovilización de moléculas de polimerasa para su uso en la secuenciación de ácido nucleico ha sido descrita por Densham en la solicitud PCT WO 99/05315. Densham describe la unión de grupos amino seleccionados dentro de la polimerasa a un dextrano o una superficie activada con éster de N-hidroxisuccinimida. Los documentos WO 99/05315; EP-A-0589867; Löfas *et al.*, *Biosens. Bioelectron.* **10**: 813-822 (1995). Estas técnicas pueden ser modificadas en la presente invención para asegurar que el área activada sea lo bastante pequeña de modo que el impedimento estérico prevenga la unión de más de una polimerasa en cualquier punto dado en la matriz.

La matriz de centros de reacción que contienen una molécula de polimerasa individual es construida usando técnicas litográficas comúnmente usadas en la construcción de circuitos electrónicos integrados. Esta metodología ha sido usada en la técnica para construir matrices microscópicas de oligodesoxinucleótidos y matrices de motores de proteína individuales. Véase, por ejemplo, Chee *et al.*, *Science* **274**:610-614 (1996); Fodor *et al.*, *Nature* **364**:555-556 (1993); Fodor *et al.*, *Science* **251**:767-773 (1991); Gushin, *et al.*, *Anal. Biochem.* **250**: 203-211 (1997); Kinoshita *et al.*, *Cell* **93**: 21-24 (1998); Kato-Yamada *et al.*, *J. Biol. Chem.* **273**: 19375-19377 (1998); y Yasuda *et al.*, *Cell* **93**: 1117-1124 (1998). Usando técnicas tales como la fotolitografía y/o la litografía de haz de electrones [Rai-Choudhury, "Handbook of Microlithography, Micromachining, and Microfabrication", Volumen I: Microlithography, Volumen PM39, ediciones SPIE (1997); Service, *Science* **283**: 27-28 (1999), el sustrato se sensibiliza con un grupo de unión que permite la unión de una proteína modificada individual. De forma alternativa, una matriz de sitios sensibilizados puede ser generada usando la tecnología de película fina tal como el Langmuir-Blodgett. Véase, por ejemplo, Zasadzinski *et al.*, *Science* **263**: 1726-1733 (1994).

El espacio regular de las proteínas se logra por la unión de la proteína a estos sitios sensibilizados en el sustrato. Las polimerasas que contienen el marcador apropiado son incubadas con el sustrato sensibilizado de modo que una molécula de polimerasa individual se una a cada sitio sensibilizado. La unión de la polimerasa puede ser lograda vía una interacción covalente o no covalente. Los ejemplos de tales enlaces comunes en la técnica incluyen Ni²⁺/hexahistidina, estreptavidina/biotina, avidina/biotina, glutatióna S-transferasa (GST)/glutatióna, anticuerpo monoclonal/antígeno y proteína de unión de maltosa/maltosa.

Una representación esquemática de un centro de reacción es presentada en la figura 3. Una ADN polimerasa (por ejemplo, de *Thermus aquaticus*) se une a un portaobjetos de vidrio de microscopio. La unión es mediada por un marcador de hexahistidina en la polimerasa, unido por una interacción no covalente fuerte a un átomo de Ni²⁺, que, a su vez, es sostenido en el vidrio por ácido nitrilotriacético y una molécula enlazante. El ácido nitrilotriacético se une covalentemente al vidrio por un enlazante unido por la química del silano. La química del silano está limitada a puntos de pequeño diámetro atacados a intervalos uniformemente espaciados sobre el vidrio por litografía de haz de electrones o fotolitografía. Además de la polimerasa unida, el centro de reacción incluye la molécula de ADN plantilla y un cebador de oligonucleótido ambos unidos a la polimerasa. El portaobjetos constituye el portaobjetos inferior de la cámara de reacción de SAND.

El alojamiento de la matriz de centros de reacción de SAND y la mediación del intercambio de reactivos y el tampón constituyen el ensamblaje de la cámara de reacción. Un ejemplo de ensamblaje de cámara de reacción de SAND es ilustrado en la figura 4. La cámara de reacción es un compartimiento sellado con portaobjetos transparentes superiores e inferiores. Los portaobjetos son mantenidos en su lugar por un carcasa de metal o plástico, que pueden ser ensamblados y desmontados para permitir el reemplazo de los portaobjetos. Hay dos puertos que permiten el acceso a la cámara. Un puerto permite la entrada del tampón (y reactivos) y el otro puerto permite que el tampón (y

los productos de reacción) sean retirados de la cámara. El portaobjetos inferior lleva la matriz de centro de reacción. Además, se une un prisma al portaobjetos inferior para dirigir la luz láser en el portaobjetos inferior a tal ángulo como para producir la reflexión total interna de la luz láser dentro del portaobjetos inferior. Este ensamblaje permite que una onda evanescente sea generada sobre la matriz de centro de reacción. Una lente del objetivo de abertura numérica alta es usada para enfocar la imagen de la matriz de centro de reacción en el sistema de la cámara digital. La carcasa de la cámara de reacción puede ser ajustado con elementos de calentamiento y refrigeración, tal como un dispositivo Peltier, para regular la temperatura de las reacciones.

Fijando el sitio de incorporación del nucleótido dentro del sistema óptico, la información de la secuencia puede ser obtenida a partir de muchas moléculas de ácido nucleico distintas simultáneamente. Se proporciona un diagrama de la matriz de centro de reacción de SAND en la Figura 5. Como se describe anteriormente, cada centro de reacción está unido al portaobjetos inferior de la cámara de reacción. Se representa en el panel del lado izquierdo (Campo de Microscopio) la vista de una matriz entera cuando se registra mediante cuatro acontecimientos de detección sucesivos (uno para cada uno de los fluorocromos separados). El panel del centro es una vista ampliada de una parte del campo que muestra el espaciado de centros de reacción individuales. Finalmente, el panel derecho lejano representa la vista de la cámara de un centro de reacción individual. A cada centro de reacción se le asignan 100 píxeles para asegurar que está realmente aislado. Se muestra el área de visualización de un píxel individual en relación con el área $1\ \mu\text{m} \times 1\ \mu\text{m}$ asignada a cada centro de reacción. La densidad de centros de reacción está limitada por la resolución óptica del sistema microscópico. Prácticamente, esto significa que los centros de reacción deben estar separados por al menos $0,2\ \mu\text{m}$ para ser detectados como sitios distintos.

2. Selección enzimática

En general, cualquier macromolécula que catalice la formación de una secuencia de polinucleótidos puede ser usada como la polimerasa. En algunas realizaciones, la polimerasa puede ser un complejo enzimático que: 1) promueva la asociación (por ejemplo, por puentes de hidrógeno o apareamiento de bases) de un marcador (por ejemplo, un nucleótido normal o modificado, o cualquier compuesto capaz de una asociación específica con nucleótidos de plantilla complementarios) con el nucleótido de plantilla complementario en el sitio activo; 2) catalice la formación de un enlace covalente entre el marcador y la cadena sintética o cebador; y 3) traduzca el sitio activo al siguiente nucleótido de plantilla.

Aunque las polimerasas serán enzimas típicamente proteicas, es obvio para cualquier experto en la técnica que la actividad de la polimerasa no tiene que estar asociada con una enzima proteica. Por ejemplo, la polimerasa puede ser un ácido nucleico en sí mismo, como en el caso de ribozimas o enzimas basadas en ADN.

Una gran selección de enzimas proteicas está disponible para su uso en la presente invención. Por ejemplo, la polimerasa puede ser una enzima tal como una ADN polimerasa ADN-dirigida, una ADN polimerasa ARN-dirigida, una ARN polimerasa ADN-dirigida y/o la ARN polimerasa ARN-dirigida o estar constituidas por un núcleo enzimático y factores asociados que realcen la actividad del núcleo (por ejemplo, que aumenten la capacidad de procesamiento o la fidelidad de la subunidad del núcleo). La enzima debe ser modificada para unirla al soporte. La enzima puede ser clonada por técnicas conocidas en la técnica, para producir una proteína recombinante con un marcador de enlace adecuado. En una realización preferida, este enlace es un marcador de hexahistidina que permite la fuerte unión a iones de níquel sobre el soporte sólido. Las enzimas preferidas son altamente procesables, es decir, permanecen asociadas a la secuencia del nucleótido plantilla para una sucesión de adiciones de nucleótido, y son capaces de mantener un complejo de polimerasa-polinucleótido incluso cuando no se sintetizan activamente. Además, las polimerasas preferidas son capaces de incorporar nucleótidos 3'-modificados. Suficientes cantidades de una enzima son obtenidas usando técnicas recombinantes estándar conocidas en la técnica. Véase, por ejemplo, Dabrowski y Kur. *Protein Expr. Purif.* **14**: 131-138 (1998).

2.1 ADN polimerasa

En una realización preferida, la secuenciación se hace con una ADN polimerasa dependiente de ADN. Las ADN polimerasas dependientes de ADN catalizan la polimerización de desoxinucleótidos para formar la cadena complementaria de un ADN plantilla cebado. Los ejemplos de ADN polimerasas dependientes de ADN incluyen, pero no están limitados a, la ADN polimerasa de *Bacillus stearothermophilus* (Bst), el fragmento de Klenow de la ADN polimerasa I de *E. coli*, la holoenzima de la ADN polimerasa III de *E. coli*, las ADN polimerasas del bacteriófago T4 y T7, y las de *Thermus aquaticus* (Taq), *Pyrococcus furiosus* (Pfu), y *Thermococcus litoralis* (Vent). La polimerasa del gen 5 de T7 también puede ser usado cuando se compleja a tioredoxina. Tabor *et al.*, *J. Biol. Chem.*, 262: 1612-1623 (1987). La ADN polimerasa de Bst es preferida porque se ha demostrado que incorpora de manera eficiente 3'-O-(-2-Nitrobenzil)-dATP en una cadena de ADN creciente, es altamente procesable, muy estable y carece de actividad con 3'-5' exonucleasa. La secuencia de codificación de esta enzima ha sido determinada. Véanse las patentes de EE.UU. N^{os}. 5.830.714 y 5.814.506.

En una realización alternativa preferida en la que se usa ARN como plantilla, la ADN polimerasa dependiente de ADN seleccionada funciona como una ADN polimerasa dependiente de ARN, o transcriptasa inversa. Por ejemplo, la ADN polimerasa de *Thermus thermophilus* (Tth) se ha publicado que funciona como una ADN polimerasa dependiente de ARN, o transcriptasa inversa, bajo ciertas condiciones. Véase, Meyers y Gelfand, *Biochem.* **30**: 7661-7666 (1991). Así, la ADN polimerasa de Tth se une al sustrato y la reacción de secuenciación se lleva a cabo en condi-

ciones en las que esta enzima secuenciará una plantilla de ARN, produciéndose así una cadena de ADN complementario.

En algunas realizaciones, una subunidad de polimerasa o fragmento se une al soporte, y otras subunidades necesarias o fragmentos son añadidos como parte de un complejo con la muestra que se secuenciará. Este enfoque es útil para sistemas de polimerasa que implican un número de factores de replicación diferentes. Por ejemplo, para usar el sistema de replicación de bacteriófago T4 para la secuenciación de SAND, la polimerasa gp43 puede ser unida a un soporte. Otros factores de replicación, tales como el cargador de abrazadera (gp44/62) y la abrazadera deslizante (gp45), pueden ser añadidos con la plantilla de ácido nucleico para aumentar la procesabilidad del sistema de replicación. Un enfoque similar puede ser usado con el sistema de la polimerasa III de *E. coli*. en el que el núcleo de la polimerasa es inmovilizado en la matriz y la subunidad del dímero β (abrazadera deslizante) y el subensamblaje τ y γ (cargador de abrazadera) son añadidos a la muestra de ácido nucleico antes de la secuenciación de SAND. Además, este enfoque puede ser usado con ADN polimerasas de eucariotas (por ejemplo, α o β) y el correspondiente ANCP (antígeno nuclear de células proliferantes). En algunas realizaciones, la abrazadera deslizante es el factor de replicación que se une en la matriz y el resto de polimerasa es añadido junto con la muestra de ácido nucleico.

2.2 Transcriptasa inversa

Una transcriptasa inversa es una ADN polimerasa dependiente de ARN - una enzima que produce una cadena de ADN complementaria a una plantilla de ARN. En una realización alternativa preferida, una enzima de transcriptasa inversa se une al soporte para su uso en la secuenciación de moléculas de ARN. Esto permite la secuenciación de ARN tomados directamente a partir de tejidos, sin una transcripción inversa previa. Los ejemplos de transcriptasas inversas incluyen, pero no están limitados a, la transcriptasa inversa del virus de Avian Myeloblastosis (VAM), virus de la leucemia murino de Moloney, y el virus 1 de la inmunodeficiencia humano (VIH-1). La transcriptasa inversa del VIH-1 es particularmente preferida porque está muy bien caracterizado tanto estructuralmente como bioquímicamente. Véase, por ejemplo, Huang, *et al.*, *Science* **282**: 1669-1675 (1998).

En una realización alternativa preferida, la transcriptasa inversa inmovilizada funciona como una ADN polimerasa dependiente de ADN, produciéndose así una copia de ADN de la muestra o de la cadena de ADN plantilla objetivo.

2.3 ARN polimerasa

En otra realización más alternativa preferida, una ARN polimerasa dependiente de ADN es unida al soporte, y usa ribonucleótidos marcados-atrapados para generar una copia de ARN de la muestra o de la cadena de ADN objetivo que se secuenciará. Los ejemplos preferidos de estas enzimas incluyen, pero no están limitados a, ARN polimerasa de *E. coli* [Yin, *et al.*, *Science* **270**: 1653-1657 (1995)] y ARN polimerasas de los bacteriófagos T7, T3 y SP6. En una realización alternativa preferida, una ARN polimerasa de T7 modificada funciona como una ADN polimerasa dependiente de ADN. Esta ARN polimerasa se une al soporte y usa desoxiribonucleótidos marcados-atrapados para generar una copia de ADN de una ADN plantilla. Véase, por ejemplo, Izawa, *et al.*, *J. Biol. Chem.* **273**: 14242-14246 (1998).

2.4 ARN polimerasa dependiente de ARN

Muchos virus emplean ARN polimerasas dependientes de ARN en sus ciclos de vida. En una realización preferida, una ARN polimerasa dependiente de ARN se une al soporte, y usa ribonucleótidos marcados-atrapados para generar una copia de ARN de una cadena de ARN de la muestra que se secuenciará. Los ejemplos preferidos de estas enzimas incluyen, pero no están limitados a, ARN polimerasas dependientes de ARN de las familias virales: bromovirus, tobamovirus, tombusvirus, levivirus, virus del tipo del de la hepatitis C y picornavirus. Véase, por ejemplo, Huang *et al.*, *Science* **282**: 1668-1675 (1998); Lohmann *et al.*, *J. Virol.* **71**: 8416-8428 (1997); Lohmann *et al.*, *Virology* **249**: 108-118 (1998), y O'Reilly y Kao, *Virology* **252**: 287-303 (1998).

3. Preparación de la muestra

El ácido nucleico que se secuenciará puede ser obtenido a partir de cualquier fuente. Las muestras de ácido nucleico ejemplares que se secuencian incluyen ADN bicatenario, ADN monocatenario, ADN de plásmido, cADN de primera cadena, ADN genómico total, ARN, ADN cortado/fin-modificado (por ejemplo, con el promotor de ARN polimerasa), transposón marcado *in vitro* (por ejemplo, inserción aleatoria del promotor de la ARN polimerasa). El ácido nucleico objetivo o muestra que se secuenciará preferiblemente se secciona (o se corta) a un cierto tamaño, y se hibrida con cebadores de oligodesoxinucleótido usando técnicas bien conocidas en la técnica. Preferiblemente, el ácido nucleico de la muestra se desnaturaliza, neutraliza y precipita y luego se diluye a una concentración apropiada, se mezcla con cebadores de oligodesoxinucleótido, se calienta a 65°C y luego se enfría a temperatura ambiente en un tampón adecuado. El ácido nucleico se añade luego a la cámara de reacción después de que la polimerasa haya sido inmovilizada sobre el soporte o, de forma alternativa, se combina con la polimerasa antes de la etapa de inmovilización.

3.1 Marcación del transposón *in vitro* del ADN plantilla

En una realización alternativa preferida, las transposasas purificadas y los marcadores del elemento reemplazables serán usados para insertar al azar las secuencias específicas en el ADN plantilla bicatenario. En una configuración, el

elemento reemplazable contiene al promotor para la ARN polimerasa específica. De forma alternativa, las repeticiones invertidas de los elementos reemplazables pueden ser hibridadas con cebadores de oligodesoxinucleótido complementarios para SAND con ADN polimerasas. Los ejemplos preferidos de estas transposasas y elementos reemplazables incluyen, pero no están limitados a, TC1 y TC3A de *C. elegans* y el sistema Sleeping Beauty de teleosteos manipulado. Véase, por ejemplo, Ivics *et al.*, *Cell* **91**:501-510 (1997); Plasterk, *Curr. Top. Microbiol. Immunol.* **204**: 125-143 (1996); van Luenen *et al.*, *EMBO J.* **12**: 2513-2520 (1993), y Vos *et al.*, *Genes Dev.* **10**:775-761 (1996).

3.2 ADN plantilla bicatenario

En otra realización más, el ADN bicatenario se secuenciará por la ADN polimerasa de Bst sin necesidad de ninguna hibridación del cebador. Véase, por ejemplo, Lu *et al.* *Chin. J. Biotechnol.* **8**: 29-32 (1992).

3.3 Cebadores

Se conocen varios cebadores y promotores en la técnica y pueden ser adecuados para la extensión de la secuencia en SAND. Los ejemplos incluyen cebadores aleatorios, genotecas de cebador de punto de anclaje, genoteca de enmascaramiento/cebador de la proteína de unión monocatenario, y la primasa.

En una realización preferida, son usados cebadores anclados en vez de cebadores aleatorios. Los cebadores de ancla son cebadores de oligonucleótido para secuencias identificadas previamente. Los cebadores de ancla pueden ser usados para la determinación rápida de secuencias específicas del ADN genómico entero, a partir de los cDNA o RNA. Esto será de uso particular para el genotipado rápido, y/o para el rastreo clínico para detectar polimorfismos o mutaciones en genes relacionados con alguna enfermedad previamente identificados u otros genes de interés. Una vez que los proyectos de genoma, y otros estudios, han identificado las secuencias de interés particular entonces pueden ser diseñados oligonucleótidos correspondientes a varias posiciones en y alrededor de aquella secuencia para su uso en SAND. Esto maximizará la cantidad de datos útiles que pueden ser obtenidos a partir de un ciclo de secuenciación individual, particularmente útil cuando son usadas muestras de ADN complejas. Para la identificación de genes de enfermedades mutados o polimórficos esta técnica evitará la necesidad de realizar el genotipado por cualquier otro medio en uso en este momento, incluyendo el uso del polimorfismo de conformación monocatenario (PCM) [Orita *et al.*, *Genomics* **5**:874-879 (1989)], la secuenciación por PCR o la tecnología de hibridación de matrices de ADN [Hacia, *Nat. Genet.* **21**: 42-47 (1999)]. La secuenciación directa del gen de la enfermedad es superior a PCM y las tecnologías de hibridación porque son relativamente insensibles y con frecuencia pueden identificar positivamente o negativamente mutaciones. Muchos oligonucleótidos de ancla pueden ser mezclados juntos de modo que cientos o miles de genes o secuencias puedan ser identificadas simultáneamente. En esencia, cada gen conocido o potencial relacionado con la enfermedad puede ser secuenciado simultáneamente a partir de una muestra dada.

4. Nucleótidos de terminación marcados-atrapados

Para ser útil como un sustrato de terminación de cadena para los métodos de la presente invención, un nucleótido debe contener un marcador detectable que lo distinga de los otros tres nucleótidos. Además, los nucleótidos de terminación de cadena deben permitir la incorporación de la base, deben terminar la elongación en la incorporación y deben ser capaces de ser liberados para permitir además la elongación de la cadena, permitiendo así ciclos repetitivos de incorporación, monitorizando para identificar las bases incorporadas, y liberando para permitir el siguiente ciclo de elongación de la cadena. La liberación de los nucleótidos puede ser lograda enzimáticamente, químicamente o, preferiblemente, fotolíticamente.

La molécula básica es un NTP con modificación en el 3'-OH (R), el 2'-OH (R'), o la base (R''). En un didesoxi NTP estándar, R=H, R'=H y R''=H.

R=H, R'=OH y R''=H es un terminador de cadena para ARN polimerasas.

Un grupo de nucleótidos de terminación de cadena útiles para los métodos de la presente invención es R = atrapado/marcador, R' = (H o OH), y R'' = H. En una realización preferida, el nucleótido modificado es un marcador (por ejemplo, un fluoróforo) unido a un resto de azúcar por un grupo 3'-O-(-2-Nitrobencil). El 3'-O-(-2-Nitrobencil)-dNTP modificado es incorporado en la cadena de ADN creciente por la ADN polimerasa de Bst unida a un soporte. Para reasumir la elongación de la cadena, el nucleótido es liberado por la eliminación del grupo 2-Nitrobencil (con su marcador detectable correspondiente) por la exposición a una luz de frecuencia apropiada. El nucleótido 3'-O-(-2-Nitrobencil)-dATP modificado ha sido usado previamente en una ronda individual de incorporación de nucleótido y liberación. Metzker *et al.*, *Nucleic Acids Res.*, **22**: 4259-4267 (1994). Véase también Cheesman, patente de EE.UU. N°. 5.302.509.

Un grupo alternativo de nucleótidos de terminación de cadena útiles tiene la configuración R = jaula, R' = (H o OH), y R'' = jaula/marcador. En una realización preferida, el grupo marcador desprendible es un marcador (por ejemplo, un fluoróforo) unido a la base del nucleótido por un grupo 2-Nitrobencil, y el grupo bloqueante desprendible es un grupo 3'-O-(-2-Nitrobencil). El nucleótido modificado es incorporado en la cadena de ADN creciente por la ADN polimerasa de Bst unida a un soporte. Para reasumir la elongación de la cadena, el nucleótido es liberado por la eliminación tanto del grupo marcador como del grupo bloqueante por la exposición a una luz de la frecuencia apropiada.

En cualquiera de estas configuraciones puede demostrarse ventajoso colocar dos marcadores (por ejemplo, dos fluorocromos) sobre cada jaula, como ha sido descrito en el documento WO 98/33939.

Para secuenciar cuando la cadena sintética es ARN, los ribonucleótidos marcados-atrapados (es decir, $R' = OH$) son sintetizados como nucleótidos modificados diseñados para la incorporación por la ARN polimerasa unida al soporte.

4.1 Marcadores fluorescentes

El uso de marcadores fluorescentes para identificar nucleótidos en secuenciación de ácidos nucleicos es bien conocido en la técnica. Véase, por ejemplo, las patentes de EE.UU. N^{os}. 4.811.218; 5.405.747; 5.547.839 y 5.821.058. Metzker y Gibbs recientemente han descrito una familia de nucleótidos fluorescentemente marcados basados en los fluoróforos de Cy con mejores características espectrales. Patente de EE.UU. N^o. 5.728.529. Grupos alternativos de fluoróforos incluyen: los fluoróforos basados en rodamina, TARAM, ROX, JOE y FAM; los fluoróforos BigDye[®] (Applied Biosystems, Inc.); y los fluoróforos BODIPY[®] (patente de EE.UU. N^o. 5.728.529).

En una realización preferida de la presente invención, un marcador fluorescente es unido al grupo bloqueante 3' fotolábil (es decir, la jaula). Los ejemplos de nucleótidos modificados para SAND son ilustrados esquemáticamente en la Figura 1 (Paneles A-C). El Panel A representa un desoxiadenosin-trifosfato modificado por la unión de un conjugado de fluorocromo-enlazante fotolábil al carbono 3' de la ribosa. La fotólisis del enlazante por luz de <360 nm hace que el fluorocromo se disocie, dejando el grupo 3'-OH del nucleótido intacto. El Panel B representa una configuración alternativa en la cual el fluorocromo está unido a la base del nucleótido por vía de un enlazante fotolábil. Los 3'-OH se bloquea por un grupo fotolábil separado. Los nucleótidos modificados, tales como los representados en los Paneles A y B, son ejemplos de desoxiribonucleótidos marcados-atrapados para su uso en SAND. Una variedad de fluorocromos y grupos fotolábiles puede ser usada en la síntesis de desoxiribonucleótidos marcados-atrapados. Además, también pueden ser sintetizados ribonucleótidos para su uso con ARN polimerasas. Cuatro fluorocromos con propiedades espectrales distintas permiten a los cuatro nucleótidos ser distinguidos durante la fase de detección del ciclo de reacción de SAND. La Figura 1 (Panel C) proporciona una representación esquemática de cuatro nucleótidos terminadores marcados-atrapados diferentes para su uso en secuenciación de ácido nucleico directa.

Después de la incorporación de los nucleótidos terminadores marcados-atrapados por las moléculas de polimerasa inmovilizadas, los fluoróforos son iluminados para excitar la fluorescencia en cada una de las cuatro especies de fluoróforos. La emisión en cada punto en la matriz es detectada ópticamente y registrada. Una vez que la información de la secuencia ha sido obtenida, los enlazantes fotolábiles son eliminados por iluminación con una luz en la longitud de onda de liberación (<360 nm).

En la Figura 2 se representa una ronda individual del ciclo de reacción, es decir, (1) la incorporación de un nucleótido marcado-atrapado; (2) la detección del nucleótido marcado; y (3) el desbloqueo del nucleótido atrapado. La información de la secuencia primaria es deducida por medio de rondas sucesivas del ciclo de reacción de SAND. En el primer panel (Etapa 1) hay un ejemplo de ADN plantilla monocatenario (3'-AGCAGTCAG-5') en el lado izquierdo hay una secuencia de cebador corta (5'-TC-3') y un dGTP marcado-atrapado que sufre la incorporación. En el panel medio (Etapa 2) el fluorocromo, BODIPY^{564/570}, es excitado por la iluminación del láser YAG a 532 nm. El fluorocromo emite una luz centrada en una longitud de onda de 570 nm, que es detectada por el sistema microscópico. Finalmente, en la Etapa 3, la fotólisis del enlazante por una iluminación con luz de <360 nm disocia simultáneamente el marcador de fluorocromo y libera el bloque 3'. Como resultado, el cebador es ampliado en una base (5'-TCG-3') y el 3'-OH es restaurado de modo que otro nucleótido pueda ser incorporado en el siguiente ciclo.

4.2 Marcadores de punto cuántico

En una realización alternativa preferida de la presente invención, se marca cada uno de los terminadores atrapados con un tipo diferente de punto cuántico. Recientemente, puntos cuánticos (QD) de semiconductores altamente luminiscentes han sido acoplados covalentemente a biomoléculas. Chan y Nie, *Science* 281: 2016-2018 (1998). Estos marcadores luminiscentes exhiben mejores características espectrales en tintes orgánicos tradicionales, y se ha demostrado que permiten la detección sensible con un microscopio de fluorescencia confocal a nivel individual del punto. En esta realización, los terminadores cuánticos de punto atrapados son incorporados, detectados y liberados de una manera similar a la descrita anteriormente para los terminadores atrapados fluorescentes.

4.3 Partículas de resonancia de plasmón

En una realización preferida, se marca cada uno de los terminadores atrapados con una partícula resonante de plasmón de plata coloidal (PRP). Schultz *et al.*, *J. Clin. Ligand Assay* 22:214-216 (1999); Schultz *et al.*, *Proc. Natl. Acad. Sci.* 97:996-1001 (2000). Las PRP son nanopartículas metálicas, típicamente de 40-100 nm de diámetro que pueden ser manipuladas para dispersar la luz hacia todas partes de manera eficiente en el intervalo visible del espectro. Estas partículas son bastante brillantes para ser usadas en la detección de moléculas individuales. Las PRP mostraron que producían un flujo de dispersión equivalente al de 5 millones de moléculas de fluoresceína, y más de 105 veces mayor que el de los puntos cuánticos típicos. Schultz *et al.*, *Proc. Natl. Acad. Sci.* 97:996-1001 (2000). Además, cuando se visualiza en un CCD estándar, el pico espacial puede ser localizado con una precisión de 10 Å, una precisión similar a la observada con imaginería de fluoróforos individuales en nanopartículas de oro. Denk y Webb, *Appl. Opt.* 29: 2382-2391 (1990). Para facilitar la detección, en ciertas realizaciones, cada tipo diferente de nucleótido es

modificado con una PRP de un color diferente. Para resolver la señal de dos PRP incorporadas en una muestra en centros de reacción vecinos, los centros de reacción deben estar separados al menos por una longitud de coherencia (aproximadamente la longitud de onda de la luz de iluminación). Además, puede ser usada la dispersión Raman para detectar las PRP. Nie y Emory, *Science* **275**: 1102-1106 (1997).

5. Detección de nucleótidos incorporados

Los avances en técnicas microscópicas han permitido la detección espectroscópica de moléculas individuales. Véase, Nie y Zare, *Annu. Rev. Biophys. Biomol. Struct.* **26**:567-596 (1997), y Keller *et al.*, *Appl. Spectrosc.* **50**:12A-32A (1996). Por ejemplo, pueden ser visualizadas moléculas fluorescentes simples en una solución acuosa con microscopía de fluorescencia de reflexión total interna (TIRFM), microscopía confocal, transferencia de energía de resonancia de fluorescencia (TRASTE), o espectroscopia de resonancia de plasmón de superficie (SPR). Véase, Dickson *et al.*, *Nature* **388**:355-358 (1997); Dickson *et al.*, *Science* **274**:966-969 (1996); Ishijima *et al.*, *Cell* **92**:161-171 (1998); Iwane *et al.*, *FEBS Lett.* **407**: 235-238 (1997); Nie *et al.*, *Science* **266**: 1018-1021 (1994); Pierce *et al.*, *Nature* **388**:338 (1997); Ha *et al.*, *Proc. Natl. Acad. Sci. USA* **93**: 6264-6268 (1996), y Gordon *et al.*, *Biophys. J.* **74**: 2702-2713 (1998). Yokota *et al.*, *Phys. Rev. Letts.* **80**: 4606-4609 (1998). Ya que pueden ser detectadas espectroscópicamente moléculas individuales, no son necesarias más para la secuenciación muestras de ácido nucleico clonadas. Una copia individual de la plantilla, contenida dentro de un centro de reacción tiene un tamaño suficiente de muestra. Los aparatos y los métodos de la presente invención permiten la resolución de señales desde marcadores de nucleótidos individuales dentro de un plano óptico y su conversión subsecuente a información digital. Los fotones son recogidos a partir de un plano delgado aproximadamente equivalente al volumen dentro del cual la enzima y la base recién sintetizada residen.

5.1 TIRFM

Cuando la luz se dirige en un ángulo particular en un medio refractivo de ancho fijo, tal como un portaobjetos, se obtendrá la reflexión total interna (TIR). Por encima del plano del medio de refracción ocurre un fenómeno electromagnético conocido como onda evanescente. El principio de la onda evanescente es representado en la figura 6. La onda evanescente se extiende desde la superficie a una distancia del orden de la longitud de onda de la luz. Destacadamente, una onda evanescente puede ser usada para excitar fluorocromos dentro de esta distancia. Cuando se usa este fenómeno para la microscopía se le llama microscopía de fluorescencia de reflexión total interna (TIRFM). El ensamblaje de portaobjetos del microscopio, prisma y rayo láser representado en esta figura conducirá a una TIR dentro del portaobjetos inferior y, como consecuencia, una onda evanescente será generada a ~150 nm de la superficie superior del portaobjetos inferior. Las moléculas de fluorocromo, tales como aquellas dentro de centros de reacción de SAND, serán excitadas y pueden ser detectadas ópticamente usando lentes de objetivos, un microscopio y el sistema de cámara. Una alta relación de señal-a-ruido es alcanzada usando la excitación de la onda evanescente porque sólo son estimuladas aquellas moléculas de fluorocromo dentro de la onda evanescente.

En una realización preferida se usa TIRFM para la detección. En la figura 7 se representa el ensamblaje del equipo requerido para llevar a cabo SAND usando TIRFM. Un soporte de microscopio de laboratorio estándar aloja el ensamblaje de la cámara de reacción, la lente del objetivo, la rueda de filtros, el intensificador con placa microcanal, y la cámara CCD enfriada. La luz láser es dirigida en el prisma por espejos dicróicos y obturadores controlados por ordenador. La excitación de la onda evanescente es usada para estimular a la muestra. La excitación de la onda evanescente es alcanzada por la reflexión total interna en la interfase vidrio-líquido. En esta interfase, el campo electromagnético óptico no cae bruscamente al cero, pero decae exponencialmente en la fase líquida. La rápida descomposición del campo (onda evanescente) puede ser usada para excitar moléculas fluorescentes en una capa delgada de aproximadamente 150 nm inmediatamente al lado de esta interfase. Véase, la solicitud de patente PCT WO 98/33939. La sensibilidad que permite la detección de una molécula individual proviene del pequeño volumen de la muestra sonado. Una ventaja del TIRFM consiste en que la matriz del centro de reacción entera puede ser visualizada simultáneamente. Las imágenes de la matriz del centro de reacción son enfocadas sobre la cara del intensificador con placa microcanal mediante filtros de barrera llevados en la rueda de filtros. El intensificador con placa microcanal amplifica la imagen y la transfiere a la cara de la cámara CCD enfriada. Los datos de imagen son leídos a partir del chip de CCD y se procesan en un microordenador. Un láser de estimulación, o grupo de láseres de estimulación, es dirigido a la muestra mediante una tabla óptica. Otro láser libera al grupo protector de 3'-OH. Se pueden requerir láseres adicionales para la estimulación óptima del fluorocromo. Una rueda de filtros también es incluida en la invención para cambiar los filtros de barrera de modo que los cuatro fluorocromos diferentes (cada uno correspondiente a un tipo diferente de nucleótido marcado-atrapado) sean distinguidos inequívocamente.

Como se muestra en la figura 7, se construye un prisma en el portaobjetos del microscopio para dirigir el láser en el portaobjetos fuera del microscopio. Ishijima *et al.*, *Cell* **92**: 161-171 (1998). De forma alternativa, puede ser usado el TIRFM tipo objetivo para la detección de la fluorescencia. La luz láser es dirigida por una lente del objetivo fuera del centro tal que el ángulo crítico sea alcanzado usando la propia lente del objetivo. Véase, Tokunaga *et al.*, *Biochem. Biophys. Res. Comm.* **235**:47-53 (1997).

5.2 Microscopía confocal

En una realización alternativa preferida, se usa para la detección la microscopía confocal. En la microscopía confocal, un rayo láser es llevado a su foco limitado por difracción dentro de una muestra usando un aceite de inmersión, lente del objetivo con alta apertura numérica (NA). Moléculas individuales han sido detectadas en solución por fluo-

rescencia confocal multifotón. Mertz, *et al.*, *Opt. Lett.* 20: 2532-2534 (1995). En una realización de esta invención, los marcadores del nucleótido son detectados mediante microscopía confocal multifotón de exploración. Nie *et al.*, *Science* **266**: 1018-1021 (1994).

5.3 Transferencia de energía de resonancia de fluorescencia (TRASTE)

En una realización alternativa preferida, la tecnología TRASTE es usada para la detección. La transferencia de energía de resonancia de fluorescencia es una interacción dependiente de la distancia entre los estados electrónicos excitados de dos moléculas de tinte en las cuales la excitación es transferida desde una molécula donadora a una molécula aceptadora sin la emisión de un fotón. El TRASTE es dependiente de la inversa en sexta potencia de la separación intermolecular, haciéndolo útil en distancias comparables a las dimensiones de macromoléculas biológicas. Así, el TRASTE es una técnica importante para investigar una variedad de fenómenos biológicos que producen cambios en la proximidad molecular.

Esta técnica hace uso de algunas propiedades insólitas de moléculas de tintes. En los experimentos que usan tintes fluorescentes, la molécula del tinte típicamente se excita a una longitud de onda de luz y los datos son recogidos a una longitud de onda más larga. Sin embargo, cuando dos moléculas de tinte diferentes son colocadas muy cerca juntas, la luz puede ser absorbida por una molécula (el donador), y su emisión puede ser entonces inmediatamente capturada por la molécula adyacente (el aceptador). Una luz a una longitud de onda todavía más larga es entonces emitida desde el aceptador. En la mayor parte de las aplicaciones, los tintes del donador y del aceptador son diferentes, en el caso de que el TRASTE pueda ser detectado por el aspecto de la fluorescencia sensibilizada del aceptador o inactivando la fluorescencia del donador. Cuando el donador y el aceptador son los mismos, el TRASTE puede ser detectado por la despolarización de la fluorescencia resultante. Las moléculas donadoras y aceptadoras deben estar en una proximidad cercana (típicamente 10-100 Å). El espectro de absorción del aceptador debe sobrelapar el espectro de emisión de fluorescencia del donador, y las orientaciones del dipolo de transición del donador y el aceptador deben ser aproximadamente paralelas.

El TRASTE puede ser empleado para aumentar las relaciones de señal-a-ruido. Además, el TRASTE puede ser usado en SAND para evitar la necesidad de un enlazante fotolábil en los fluorocromos. El TRASTE es usado comúnmente para medir la distancia entre moléculas o sus partes, o para detectar interacciones moleculares transitorias. En la práctica, son modificadas moléculas candidato, o partes diferentes de la misma molécula, con dos grupos fluorescentes diferentes. La solución entonces se excita con una luz correspondiente a la longitud de onda de excitación más corta de los dos fluorocromos. Cuando el segundo fluorocromo está en una proximidad cercana al primero, se excitará por la energía emitida del primero y emitirá en su propia longitud de onda característica. La eficacia (rendimiento cuántico) de la conversión se refiere directamente a la distancia física entre los dos fluorocromos. Para la aplicación específica al SAND, son marcadas moléculas de polimerasa con un fluorocromo que se comporta como un donador de fotones para los nucleótidos modificados. Esto limitaría su excitación al sitio activo de la polimerasa o cualquier otra parte apropiada de la polimerasa. Tal ensamblaje aumentaría considerablemente la relación de señal-a-ruido de detección del nucleótido. Además, porque sólo los nucleótidos dentro de la polimerasa son excitables con TRASTE según se aplica a SAND, sería innecesaria la eliminación de los restos fluorescentes previamente incorporados. El TRASTE ha sido realizado a nivel de una molécula individual como se requiere para SAND [Ha *et al.*, *Proc. Natl. Acad. Sci. USA* **93**: 6264-6268 (1996)], y ha sido optimizado para la cuantificación en microscopía de fluorescencia. Gordon *et al.*, *Biophys. J.* **74**: 2702-2713 (1998). Óptimamente la polimerasa sería sintetizada como una proteína de fusión de una proteína fluorescente verde recombinante (GFP) ya que esto eliminaría la necesidad de derivatizar la polimerasa y, a diferencia de los fluorocromos más comúnmente usados, GFP es sustancialmente resistente al fotoblanqueo. Sin embargo, los inventores encuentran que el ensamblaje óptimo es una polimerasa químicamente modificada a la cual ha sido unido un fluorocromo sintético o punto cuántico.

5.4 Resonancia de plasmón de superficie

En una realización, la espectroscopia de resonancia de plasmón de superficie (SPR) es usada para detectar la incorporación del marcador en la muestra de ácido nucleico. La SPR se usa para medir las propiedades de una solución detectando las diferencias del índice de refracción entre la fase bruta de la solución y la región de la onda evanescente. La SPR ha sido recientemente usada para visualizar moléculas individuales de proteínas fluorescentemente marcadas en metal mediante plasmones de superficie en una solución acuosa. Yokota *et al.*, *Phys. Rev. Letts.* **80**: 4606-4609 (1998). Esta técnica implica el revestimiento de la superficie de la cámara de reacción con una capa delgada de metal para realzar la señal de los nucleótidos fluorescentemente marcados.

5.5 El detector de SAND

El detector es una cámara CCD enfriada ajustada con un intensificador con placa microcanal. Un diagrama de bloques de la estructura del instrumento se representa en la figura 7. Las cámaras CCD actualmente disponibles enfriadas-intensificadas tienen una resolución de al menos 1000x1000 píxeles. En una realización preferida de esta invención, una matriz consiste en 100x100 centros de reacción. Así, cuando la matriz es visualizada en la cara de la cámara, cada centro de reacción es asignado con aproximadamente 10x10 píxeles. El SAND usa una lente 63x 1,4 NA para visualizar una matriz (100x100 µm de rejilla) de centros de reacción espaciados con regularidad, representado en la figura 5. La información puede ser registrada simultáneamente a partir de 10.000 centros de reacción. Esta resolución esperada es comparable a la alcanzada en una publicación reciente, en la cual fue usado un TIRFM para visualizar

una muestra de fluoróforos rojo Nilo, y se produjeron imágenes de un gran número de moléculas individuales. Una molécula de rojo Nilo individual fue visualizada inequívocamente en un cuadrado de 8x8 píxeles. Dickson *et al.*, *Nature* **388**: 355-358 (1997).

5 6. El ciclo de secuenciación

La carcasa de la matriz de centros de reacción de SAND y la mediación del intercambio de reactivos y tampón comporta el ensamblaje de la cámara de reacción. La cámara de reacción es un compartimiento sellado con portaobjetos transparentes superiores e inferiores. Los portaobjetos son mantenidos en su lugar mediante una carcasa de metal o de plástico, que puede ser ensamblada y desmontada para permitir el reemplazo de los portaobjetos. Hay dos puertos que permiten el acceso a la cámara. Un puerto permite la entrada de tampón (y reactivos) y el otro puerto permite que el tampón (y los productos de reacción) sean retirados de la cámara. El portaobjetos inferior lleva la matriz del centro de reacción. Además, se une un prisma al portaobjetos inferior para dirigir la luz láser en el portaobjetos inferior a tal ángulo como para producir la reflexión total interna de la luz láser dentro del portaobjetos inferior. Este ensamblaje permite que se genere una onda evanescente en la matriz del centro de reacción. Una lente de objetivo de abertura numérica alta es usada para enfocar la imagen de la matriz del centro de reacción en el sistema de la cámara digital. La carcasa de la cámara de reacción puede ser ajustada con elementos de calentamiento y de refrigeración, tal como un dispositivo Peltier, para regular la temperatura de las reacciones. Una muestra de ácido nucleico es introducida a la cámara de reacción en la solución tamponada que contiene los cuatro terminadores de trifosfato de nucleósido marcados.

Una representación esquemática del ensamblaje de la cámara de reacción se representa en la figura 4. Los centros de reacción son monitorizados por el sistema microscópico hasta que una mayoría de centros de reacción contenga la polimerasa inmovilizada unida a la plantilla con un nucleótido terminador marcado-atrapado incorporado individual. Después la cámara de reacción se enjuaga con tampón de lavado. La incorporación del nucleótido específico entonces se determina para cada centro de reacción. Después de la detección, la cámara de reacción es irradiada para liberar al nucleótido incorporado y es aclarada con tampón de lavado una vez más. La presencia de nucleótidos marcados se monitoriza otra vez antes de que los reactivos recién preparados sean añadidos para iniciar de nuevo la síntesis. Esta segunda detección verifica que se libera un centro de reacción satisfactoriamente. La presencia de un nucleótido marcado en la cámara durante esta etapa indica que el centro de reacción no ha sido liberado. En consecuencia, la lectura subsecuente de este centro de reacción durante la siguiente etapa de detección del ciclo será obviada. Así, obviando las señales de los centros de reacción que no sean liberados satisfactoriamente, los métodos de la presente invención evitan los problemas causados por la liberación incompleta en los métodos de secuenciación de la técnica anterior. El ciclo de secuenciación descrito anteriormente es repetido hasta que una proporción grande de centros de reacción falle continuamente en incorporarse o en liberar a nucleótidos adicionales.

Métodos para regular el suministro (y la eliminación) de reactivos a los centros de reacción, así como el ambiente de la cámara de reacción (por ejemplo, la temperatura y el ambiente oxidativo) son incorporados en la cámara de reacción usando técnicas comunes en la técnica. Los ejemplos de esta tecnología son descritos en: Kricka, *Clinic Chem.* **44**: 2008-2014 (1998); véase también la patente de EE.UU. N°. 5.846.727.

7. Software de adquisición de secuencia

El software de adquisición de la secuencia adquiere y analiza los datos de las imágenes durante el ciclo de secuenciación. Al principio de un experimento de secuenciación, una imagen binaria de píxeles que contienen cada centro de reacción es determinada. Durante cada ciclo de secuenciación, cuatro imágenes de la matriz entera son producidas, y cada imagen corresponde a la excitación de una de las cuatro bases de nucleótido fluorescentemente marcadas A, C, G o T (U). Para cada imagen binaria del centro de reacción, todas las cuatro imágenes son analizadas para determinar qué especie de nucleótido ha sido incorporado en aquel centro de reacción durante aquel ciclo. Como se describe anteriormente, la imagen binaria del centro de reacción correspondiente a un cierto centro de reacción contiene una matriz de 10x10 píxeles. El número total de fotones producidos por el fluoróforo individual en aquel centro de reacción es determinado por la adición de cada valor de píxeles en la matriz. Típicamente, 500-1500 fotones son emitidos desde un fluoróforo individual cuando se excita durante 100 milisegundos con un láser que produce una intensidad de 5 kW/cm² en la superficie del portaobjetos del microscopio. Dickson *et al.*, *Science* **274**: 966-969 (1996). Las sumas de las imágenes binarias del centro de reacción de cada una de las cuatro imágenes se comparan, y la imagen que produce una suma significativa corresponde a la base recién incorporada en aquel centro de reacción. Las imágenes son tratadas para cada uno de los centros de reacción y una matriz de nucleótidos incorporados se registra. Un ejemplo de un algoritmo de adquisición de datos es proporcionado en la Figura 8. Tal tratamiento se hace en tiempo real a un precio bajo con ordenadores de procesamiento de imágenes modernos.

Pueden ser necesarias lecturas múltiples de la matriz del centro de reacción durante la etapa de detección para asegurar que los cuatro nucleótidos son distinguidos correctamente. Los tiempos de exposición pueden ser tan bajos como de 100 ms, y el tiempo de lectura del chip CCD puede ser de no menos de 250 ms. Así, el tiempo máximo necesario para cuatro lecturas completas de la matriz es de 1,5 segundos. El tiempo total para un ciclo dado, incluyendo la adición de reactivos, la eliminación, y el lavado, es seguramente menos de 10 segundos. En consecuencia, un aparato de secuenciación que consista en una matriz de 10.000 centros de reacción es capaz de detectar al menos 360 bases por sitio por hora, o 3,6 Megabases por hora de secuencia total, como una estimación conservadora. Esta velocidad es considerablemente más rápida que la de las metodologías de secuenciación tradicionales.

Además de acortar los tiempos de secuenciación, los métodos de la presente invención no requieren procesos que lleven mucho tiempo de amplificación de la muestra (clonación o PCR) y la electroforesis de gel. La carencia de bienes consumibles necesarios para la amplificación de la muestra y la electroforesis, acoplada con pequeños volúmenes de reactivo (el volumen de la cámara de reacción está sobre el orden de 10 microlitros) y los bajos requerimientos de trabajo manual reducen drásticamente el coste por nucleótido secuenciado en relación con las técnicas de secuenciación tradicionales.

8. Software de análisis de secuencia

Se representa en la figura 8 un ejemplo de adquisición de datos de SAND usando una matriz de 3x3 centros de reacción. En una configuración típica, sin embargo, los SAND utilizarían una matriz de 100x100 centros de reacción. En este ejemplo, son representados cuatro ciclos de SAND. Para cada ciclo, cuatro imágenes de la matriz son producidas. Cada imagen corresponde a una combinación de la longitud de onda de excitación específica y el filtro de barrera, y así corresponde a la incorporación de un nucleótido modificado específico. Considérese la matriz superior izquierda (Ciclo 1, A). En este caso, cuando se usa el grupo BODIPY de nucleótidos modificados "A" es 3'-O-(DMNPE-(BODIPY⁴⁹³/₅₀₃))-2'-desoxi ATP. Así, la matriz de centro de reacción es iluminada con una luz de 488 nm con un láser de Ar y la imagen es enfocada mediante un filtro de barrera de 503 nm. Cada uno de los nueve elementos en la matriz 3x3 corresponde a una área de 10x10 píxeles de la salida de la cámara CCD. Para cada una de las cuatro imágenes cada grupo de píxeles del centro de reacción es analizado para determinar si ha sido incorporado un nucleótido dado. Así, se observa en el ejemplo que en el Ciclo 1, A, los desoxiATP modificados fueron incorporados en los centros de reacción X1 y Z1. A partir de ahí, en la tabla los primeros nucleótidos registrados para los centros de reacción X1 y Z1 son "A". Si consideramos un centro de reacción dado, por ejemplo, el centro de reacción X1, sobre los cuatro ciclos de SAND, se observa que en el primer ciclo el centro de reacción ha incorporado una "A", en el segundo ciclo una "C", en el tercer ciclo una "C" y en el cuarto ciclo una "T". A partir de ahí, el fragmento de secuencia del ADN plantilla unido al centro de reacción Y3 es el complemento inverso de 5'-ACCT-3', que es 5'-TGGA-3'. La secuencia primaria existe como una matriz de secuencias, cada una derivada a partir de un centro de reacción individual. La longitud de cada secuencia del centro de reacción dependerá del número de ciclos que un centro dado permanece activo en un experimento. Basado en la capacidad de procesar de las polimerasas clonadas mencionadas en la técnica, se esperan longitudes de secuencia de varios cientos a varios miles de bases.

En una realización de la presente invención, una muestra de ácido nucleico es cortada antes de su inclusión en un centro de reacción. Una vez que estos fragmentos han sido secuenciados, el software de análisis de secuencias se usa para ensamblar sus secuencias en extensiones contiguas. Existen muchos algoritmos en la técnica con los que se pueden comparar secuencias y deducir su correcto solapamiento. Se han diseñado nuevos algoritmos recientemente para tratar grandes cantidades de datos de secuencias con enfoques de secuenciación al azar (aleatoria).

En una realización preferida, un algoritmo reduce inicialmente la cantidad de datos que se procesan usando sólo dos secuencias más pequeñas derivadas a partir de uno u otro extremo de la secuencia deducida a partir de un centro de reacción individual en un experimento dado. Este enfoque ha sido propuesto para su uso en la secuenciación al azar del genoma humano. Rawlinson, *et al.*, *J. Virol* **70**: 8833-8849 (1996); Venter *et al.*, *Science* **280**: 1540-1542 (1998). Este emplea algoritmos desarrollados en el Instituto para la Investigación del Genoma (TIGR). Sutton, *et al.*, *Genoma Sci. Technol.* **1**: 9 (1995).

En una realización alternativa preferida, los datos en bruto son comprimidos en una huella digital de palabras más pequeñas (por ejemplo, sitios de la enzima de restricción del hexanucleótido) y estas huellas digitales pueden compararse y ensamblarse en bloques continuos más grandes de secuencia (contigs). Esta técnica es similar a la usada para deducir secuencias solapantes después de la hibridación del oligonucleótido. Idury y Waterman, *J. Comput. Biol.* **2**:291-306 (1995). Otra realización más usa datos de secuencia existentes, a partir de mapas de ligamientos genéticos o físicos, para ayudar al ensamblaje de nuevos datos de secuencia de genomas enteros o grandes pedazos genómicos.

9. Utilidad del SAND

(a) Aplicaciones clínicas

La importancia de diagnósticos genéticos en medicina no puede ser subestimada. Lo más obvio es el uso de técnicas que puedan identificar a vehículos de rasgos genéticos dañinos para el diagnóstico prenatal y neonatal. Actualmente, los análisis bioquímicos y los análisis de cariotipo son las técnicas más comúnmente usadas, pero éstas tienen limitaciones claras. Los análisis bioquímicos son sólo útiles cuando hay un cambio de la actividad o de los niveles de una enzima o proteína que ha sido asociada con el estado de la enfermedad y para el cual ha sido determinada una prueba específica. Incluso cuando una proteína ha sido atribuida a un estado de enfermedad, el desarrollo de tales reactivos puede ser difícil, caro y largo. Los análisis cariotípicos son sólo útiles para identificar graves trastornos genéticos tales como ploidía, translocaciones y grandes deleciones. Aunque sea teóricamente posible determinar si los individuos poseen los alelos defectuosos de un gen dado por las técnicas de ADN actuales, los programas de rastreo eficaces son sólo practicablemente en casos en los cuales una mutación común está asociada con la enfermedad y su presencia puede ser determinada mediante técnicas no secuenciantes.

Los métodos de la presente invención permiten determinar grandes cantidades de datos de secuencia de ADN a partir de un paciente individual con poco esfuerzo técnico, y sin necesidad de clonar el ADN del paciente o amplificar las secuencias específicas por PCR. Pueden ser secuenciadas moléculas individuales directamente a partir de una preparación de ADN simple a partir de la sangre del paciente, muestras de tejido o a partir del fluido amniótico. En consecuencia, puede ser usado el SAND para el diagnóstico clínico de trastornos genéticos, rasgos u otras propiedades predecibles a partir de la información de la secuencia del ADN primario, tales como diagnósticos prenatales, neonatales y postnatales o la detección de trastornos congénitos; el análisis patológico de una enfermedad somática causada por recombinación genética y/o mutación; la identificación de pérdida de heterocigosidad, mutaciones puntuales, u otros cambios genéticos asociados al cáncer, o presentes en estados precancerosos.

Los métodos de la presente invención también pueden ser usados para identificar patógenos que causen enfermedades (por ejemplo, virales, bacterianos, micóticos) por secuenciación directa de los tejidos afectados.

15 (b) *Identificación génica funcional*

Los rastreos genéticos a gran escala para genes implicados en ciertos procesos, por ejemplo durante el desarrollo, son ahora comunes y son aplicados a vertebrados con grandes genomas tales como el pez cebra (*Danio rerio*) y el anfibio *Xenopus tropicalis*. Los intentos para clonar genes mutantes en ratón y ser humano han sido muy largos y difíciles y hasta en organismos genéticamente más manipulables como el pez cebra esto es todavía un proceso largo y difícil.

Ya que los métodos de la presente invención permiten la secuenciación de un genoma entero, el tamaño de un mamífero en un período corto de tiempo, la identificación de genes mutantes puede ser lograda por el rastreo de la secuencia en bruto, es decir, la secuenciación de genomas enteros o grandes segmentos genómicos de un vehículo, y comparando con la secuencia de genomas enteros o grandes segmentos genómicos de los diferentes miembros de una especie dada.

Asimismo, los métodos de la presente invención permiten la fácil secuenciación de genomas bacterianos enteros. La información de la secuencia generada de esta manera puede ser usada para la identificación rápida de genes que codifican nuevas enzimas a partir de una amplia variedad de organismos, incluyendo bacterias extremófilas.

Además, los métodos de la presente invención también pueden ser usados para la evaluación de velocidades de mutación en respuesta a mutágenos y radiación en cualquier tejido o tipo de célula. Esta técnica es útil para la optimización de protocolos para futuros rastreos de mutaciones.

35 (c) *Análisis de alteraciones genéticas en tumores*

Muchos cánceres, posiblemente todos los cánceres, comienzan con alteraciones específicas en el genoma de una célula o unas células, que entonces se vuelven descontroladas por los controles del crecimiento normal. La mayor parte del tratamiento de cánceres es dependiente de la respuesta fisiológica específica de estas células anormales frente a agentes particulares.

El método de la presente invención permitirá la generación rápida de un perfil genético de tumores individuales, permitiendo a los investigadores seguir con precisión qué cambios genéticos acompañan a las diversas etapas de progresión del tumor. Esta información también permitirá el diseño de agentes específicos para actuar sobre células cancerígenas para intervenciones hechas a medida sobre tumores individuales.

50 (d) *Análisis de la variación genética*

Muchos rasgos fisiológicos importantes, tales como el control de la tensión arterial, son controlados por una multiplicidad de *loci* genéticos. Actualmente, estos rasgos son analizados por el estudio de *locus* de rasgos cuantitativos (QTL). Generalmente, en el análisis QTL, un grupo de marcadores de ligamiento genético polimórficos es utilizado sobre un grupo de sujetos con un rasgo particular, tal como la hipertensión crónica familiar. A través de un análisis del ligamiento de los marcadores con el rasgo, se perfila una correlación entre un grupo de *loci* particulares y el rasgo. Por lo general, un puñado de *loci* contribuyen a la mayoría del rasgo y un grupo más grande de *loci* tendrá efectos menores sobre el rasgo.

Los métodos de la presente invención permiten la secuenciación del genoma entero rápidamente. Así, usando los métodos de la presente invención, el análisis QTL es realizado a una escala muy fina y, con un grupo grande de sujetos, todos los *loci* principales que contribuyen a un rasgo dado y la mayor parte de los *loci* menores son fácilmente identificados.

Además, el método de la presente invención puede ser usado para construir árboles filogenéticos y/o relaciones de parentesco por valoración de recombinaciones genómicas anteriores (por ejemplo, inversión, translocación, delección, mutación de punto), o por acontecimientos de recombinación meiótica previos que afectan a la distribución de marcadores polimórficos. El método de la presente invención puede ser usado para identificar mutaciones o polimorfismos, con el objetivo de asociar el genotipo con el fenotipo. El método de la presente invención también puede ser

usado para identificar la secuencia de aquellos genes mutantes o polimórficos que causan un fenotipo específico, o que contribuyen a un rasgo poligénico.

(e) *Aplicaciones agrícolas*

La eficacia agrícola y la productividad son aumentadas generando clases de plantas y animales con características genéticas óptimas. Los métodos de la presente invención pueden ser usados, por ejemplo, para revelar la variación genética que es la base tanto de rasgos deseables como indeseables en plantas agrícola y en animales. Además, los métodos de la presente invención pueden ser usados para identificar patógenos de plantas y animales y para diseñar métodos para combatir a éstos.

(f) *Aplicaciones forenses*

Los métodos de la presente invención pueden ser usados en investigaciones criminales y forenses, o con el objetivo de determinar la paternidad/maternidad identificando genéticamente muestras de sangre, pelo, piel y otros tejidos para establecer inequívocamente un eslabón entre un individuo problema y muestras forensicamente relevantes. Los resultados obtenidos serán análogos a los resultados obtenidos con técnicas de huella digital genética, pero proporcionarán una información mucho más detallada y, con menor probabilidad, proporcionarán una identificación falsa positiva. Además, puede ser determinada la identidad de individuos de una muestra mixta.

(g) *Aplicaciones en investigación*

Los métodos de la presente invención pueden ser usados para varias aplicaciones en investigación, tales como la secuenciación de constructos de ADN artificiales para confirmar/obtener su secuencia primaria, y/o para aislar a clones mutantes específicos a partir de rastreos de mutagénesis aleatorios; la secuenciación del cADN a partir de células individuales, tejidos enteros u organismos en cualquier etapa del desarrollo o circunstancia ambiental para determinar el perfil de expresión génica de aquel espécimen; la secuenciación de productos por PCR y/o los fragmentos de ADN clonados de cualquier tamaño aislados a partir de cualquier fuente.

Los métodos de la presente invención también pueden ser usados para la secuenciación de fragmentos de ADN generados por técnicas analíticas que sondan la estructura de ADN de orden más alto por su sensibilidad diferencial frente a enzimas, radiación o tratamiento químico (por ejemplo, tratamiento parcial de DNasa de cromatina), o para la determinación del estado de metilación de ADN comparando la secuencia generada de un tejido dado con o sin tratamiento previo con sustancias químicas que convierten la metil-citosina en timina (u otro nucleótido) como la base eficaz reconocida por la polimerasa. Además, los métodos de la presente invención pueden ser usados para analizar cambios fisiológicos celulares que ocurren durante el desarrollo o la senectud a nivel de la secuencia primaria.

Los métodos de la presente invención también pueden ser usados para la secuenciación de genomas enteros o grandes segmentos genómicos de células transformadas para seleccionar individuos con un estado de integración deseado. Por ejemplo, SAND puede ser usado para el rastreo de líneas de células madre embrionarias transfectadas para la integración correcta de constructos específicos, o para el rastreo de organismos tales como drosófila, pez cebra, ratón o tejidos humanos para acontecimientos de integración específicos.

Además, el método de la presente invención puede ser usado para identificar genes nuevos mediante la identificación de bloques conservados de secuencia o motivos de organismos evolucionariamente divergentes. El método de la presente invención también puede ser usado para la identificación de otros elementos genéticos (por ejemplo, secuencias reguladoras y sitios de unión a proteína) por conservación de la secuencia y posición genética relativa.

Los detalles de una o varias realizaciones de la invención han sido expuestos en la descripción que acompaña anterior. Aunque cualquier método y materiales similares o equivalentes a los descritos en este documento puedan ser usados en la práctica o pruebas de la presente invención, los métodos preferidos y los materiales son descritos a continuación. Otras propiedades, objetos y ventajas de la invención serán evidentes a partir de la descripción y de las reivindicaciones. En la memoria descriptiva y las reivindicaciones añadidas, las formas singulares incluyen referentes plurales a no ser que el contexto dicte claramente otra cosa. A no ser que se defina de otra manera, todos los términos y expresiones técnicas y científicas usadas en este documento tienen el mismo significado que comúnmente se entiende por cualquier experto ordinario en la técnica a la cual esta invención pertenece.

Los Ejemplos siguientes son presentados para ilustrar más generalmente las realizaciones preferidas de la invención. Estos Ejemplos de ninguna manera deberían ser interpretados como una limitación del alcance de la invención, como se define por las reivindicaciones añadidas.

Ejemplo 1

Preparación del sustrato de la cámara de reacción, conjugado Niquel/quelante

La unidad fundamental de la metodología SAND es el centro de reacción (figura 3). El centro de reacción comprende una molécula de polimerasa unida a una molécula de ácido nucleico plantilla, y unida a una posición fija sobre

un sustrato transparente vía una interacción de alta afinidad entre grupos unidos a la polimerasa y el sustrato, respectivamente. En una configuración, las reacciones de SAND ocurren en una cámara de reacción cuya base, el sustrato, está hecho de vidrio (SiO_2) modificado de modo que las moléculas de polimerasa puedan ser unidas en una matriz regular. Usando litografía de haz de electrones se genera una matriz cuadrada de dimensiones $100\ \mu\text{m} \times 100\ \mu\text{m}$. Rai-Choudhury, "Handbook of Microlithography, Micromachining, and Microfabrication", Volumen I: Microlithography, Volumen PM39, ediciones SPIE (1997). Un pequeño punto, $< 50\ \text{nm}$ de diámetro, es atacado en cada intervalo de $1\ \mu\text{m}$ en el material resistente que reviste el portaobjetos. Este ataque químico expone al vidrio a una derivatización subsecuente en la cual un grupo de ácido nitrilotriacético es unido covalentemente mediante la química del silano. Schmid, *et al.*, *Anal. Chem.* 69: 1979-1985 (1997). Cada grupo de ácido nitrilotriacético sirve como un quelante para un ión de Ni^{2+} . El ión de Ni^{2+} coordinado puede ser unido entonces por restos de hexahistidina manipulados en una variedad de moléculas de polimerasa. Así, una matriz de 10.000 moléculas de polimerasa es generada en una matriz de $100\ \mu\text{m} \times 100\ \mu\text{m}$, que serán observadas en un sistema de microscópico óptico. En una configuración alternativa, se une covalentemente biotina a cada punto mediante la química del silano. La biotina entonces se une a través de restos de estreptavidina covalentemente unidos, o manipulados, en las moléculas de polimerasa.

Ejemplo 2

La cámara de reacción microfluídica permite el intercambio rápido de reactantes, tampón y productos

La cámara de reacción es un dispositivo que aloja la matriz de centros de reacción y regula el ambiente. Como se describe en el Ejemplo 1, el sustrato es un portaobjetos de microscopio de vidrio preparado con una matriz microscópica regular de restos covalentemente unidos. Un prisma se une al portaobjetos sobre la superficie opuesta a la matriz. El prisma dirige la luz láser en el portaobjetos con tal ángulo que la reflexión total interna de la luz láser es alcanzada dentro del portaobjetos. En estas condiciones, una onda evanescente es generada sobre la matriz durante el ciclo de reacción de secuenciación. El portaobjetos y el prisma están fijados en una ensamblaje, que generará una cámara sellada con un volumen de $1\text{-}10\ \mu\text{l}$ (figura 4). Los reactivos y el tampón son bombeados hacia y desde la cámara a través de puertos microfluídicos por todos los lados de la cámara. Los intercambios completos de volumen ocurren en 1 segundo y son mediados por válvulas electrónicamente controladas y bombas.

Ejemplo 3

Preparación de nucleótidos de terminación de cadena marcados-atrapados

Preparación del conjugado enlazante de fluorocromo-fotolábil

Los derivados de 2-nitrobencilo unidos por fluorocromo son generados primero como se describe en Anasawa, *et al.*, documento WO 98/33939. De forma alternativa, un enlazante fotolábil sensibilizado (por ejemplo, usando el kit de atrapamiento de DMNPE, número de catálogo D-2516, Molecular Probes, Inc.) puede ser unido primero al grupo 3' del dNTP como se detalla más abajo y luego unirse a un fluorocromo usando la química de succinimida o de otra forma. Puede demostrarse óptimo usar un enlazante de longitud variable entre el fluorocromo y el grupo atrapante para reducir el impedimento estérico posible causado por grandes grupos químicos. Brandis, *et al.*, *Biochemistry* 35: 2189-2200 (1996).

Preparación de análogos de 3'-O-modificado-2'-desoxinucleótido

Los 3'-O-modificado-2'-desoxinucleótidos son sintetizados por esterificación del grupo 3'-OH de dATP, dCTP, dGTP y dTTP. Esto se logra con varios métodos generales. Metzker, *et al.*, *Nucleic Acids Res.* 22:4259-4267 (1994).

Método 1

Primero se hacen reaccionar los 2'-desoxi-5'-hidroxi-dNTP con *tert*-butildifenilsililo (TBDPS) en presencia de imidazol y dimetilformamida (DMF) produciendo desoxinucleótidos 5'-protegidos. Después el 2'-desoxi-5'-*tert*-butildifenilsililo dNTP resultante se disuelve en benceno y se mezcla con el derivado de haluro del conjugado enlazante del fluorocromo-fotolábil en presencia de hidróxido de tetrabutilamonio (TBAH) (y además NaOH en algunos casos) y se agita a 25°C durante 16 horas. La capa orgánica se extrae con acetato de etilo y se lava con agua desionizada, NaCl saturado, se seca con Na_2SO_4 y se purifica con cromatografía flash usando un gradiente gradual (metanol/acetato de etilo del 10% hasta metanol/acetato de etilo del 5% en intervalos del 2%).

Método 2

Los 2'-desoxi-5'-*tert*-butildifenilsililo dNTP preparados como se detalla anteriormente se hacen reaccionar directamente con el anhídrido de ácido del conjugado de enlazante de fluorocromo-fotolábil en piridina seca en presencia de 4-dimetilaminopiridina (DMAP) a 25°C durante 6 horas. La piridina entonces es eliminada bajo el vacío, el residuo se disuelve en agua desionizada, se extrae en cloroformo, se lava con agua desionizada, con HCl del 10%, NaHCO_3 saturado, NaCl saturado, se seca con Na_2SO_4 , y se purifica por cromatografía flash.

ES 2 310 514 T3

Método 3

Los 2'-desoxi-5'-*terc*-butildifenilsililo dNTP son secados por co-evaporación repetida con piridina, disueltos en DMF caliente y enfriados a 0°C en un baño de hielo. El NaOH se disuelve en DMF después del lavado con benceno seco, luego se añade al 2'-desoxi-5'-*terc*-butildifenilsililo disuelto y se agita durante 45 minutos. Un derivado halogenado del conjugado enlazante de fluorocromo-fotolábil en DMF es añadido y la reacción es agitada durante unas horas. La reacción entonces es inactivada con agua fría desionizada y se agita de la noche a la mañana. El sólido obtenido se filtra, se seca, y se recrystaliza en etanol.

Método 4

Pueden ser preparados 3'-atrapados NTP directamente a partir del trifosfato de acuerdo con Hiratsuka *et al.*, *Biochim Biophys Acta* **742**: 496-508 (1983).

En el caso de los métodos 1-3, los compuestos resultantes posteriormente son desilados por la adición de 1,0 equivalentes de fluoruro de tetrabutilamonio (Bu_4NF). Las reacciones son monitorizadas por cromatografía de capa fina y después de la terminación (aproximadamente 15 minutos), las reacciones son inactivadas con 1 equivalente de ácido acético glacial. El disolvente se elimina, y los residuos se purifican por cromatografía de columna de sílice. Los derivados de 5'-trifosfato del compuesto generado por los métodos 1-3 son sintetizados según el protocolo siguiente. El nucleósido 3'-modificado (1,0 equivalentes) se disuelve en trimetilfosfato bajo una atmósfera de nitrógeno. Se añade oxiclورو de fósforo (POCl_3) (3,0 equivalentes) y la reacción se agita a -10°C durante 4 horas. La reacción se inactiva con una solución de trifosfato de tributilamonio (5,0 equivalentes) en DMF y tributilamina. Después de agitar enérgicamente durante 10 minutos, la reacción se inactiva con TEAB a pH 7,5. La solución se concentra, y el derivado de trifosfato se aísla por gradiente lineal (TEAB de 0,01 M a 0,5 M) usando una columna de celulosa DEAE (HCO_3^- -forma).

Los productos sintéticos finales son purificados por HPLC, y además pueden ser purificados mediante limpieza enzimática si fuera necesario [Metzker, *et al.*, *Biotechniques* **25**:814-817 (1998)], una técnica que utiliza la preferencia enzimática extrema de muchas polimerasas para desoxinucleótidos frente a sus parejas 3'-bloqueadas. Esto probablemente es el resultado de la baja eficacia de la formación catalítica del enlace de fosfodiéster cuando los nucleótidos 3'-modificados están presentes en el sitio activo enzimático de modo que la enzima tiende a agotar rápidamente los desoxinucleótidos contaminantes normales primero. Brandis, *et al.*, *Biochemistry* **35**: 2189-2200 (1996).

En una configuración alternativa, un grupo fotolábil se une a 3'-OH usando succinimida u otra química y un conjugado enlazante del fluorocromo-fotolábil es unido directamente a la base del nucleótido como se describe en Anasawa *et al.*, documento WO 98/33939. El grupo fotolábil unido de 3' servirá como un terminador de cadena reversible [Metzker, *et al.*, *Nucleic Acids Res.* **22**: 4259-4267 (1994)] y el enlazante de fluorocromo-fotolábil unido por base servirá como un marcador desprendible. En esta configuración con cada ciclo ambos grupos fotolábiles serán eliminados por fotólisis antes de que se permita una incorporación más. Tal configuración puede ser preferida si se encuentra que el impedimento estérico de grandes grupos fluorocromos unidos al 3'-OH del nucleótido previenen al nucleótido de la entrada de la polimerasa.

Ejemplo 4

SAND usando una ADN polimerasa marcada con hexahistidina clonada, ADN plantilla monocatenario cebado aleatorio y microscopia de fluorescencia de reflexión total interna

Hay dos fases para el procedimiento.

Fase 1

La primera fase es la fase de estructuración. La ADN polimerasa marcada con hexahistidina se lava en la cámara de reacción y se permite que se una a la matriz de Ni^{2+} -nitrilotriacético. Como ejemplo, podría ser usada la ADN polimerasa marcada con hexahistidina de *Thermus aquaticus*. Dabrowski, *et al.*, *Acta Biochim Pol* **45**: 661-667 (1998). El ADN plantilla se prepara cortando o por digestión de restricción, seguido de la desnaturación a 95°C e hibridando con una mezcla de cebadores de oligodesoxinucleótido aleatorios. El ADN plantilla monocatenario cebado entonces se bombea en la cámara de reacción.

Fase 2

La segunda fase del procedimiento es el ciclo de secuenciación principal. El ciclo es como sigue:

1. El tampón de reacción que contiene desoxinucleósido trifosfatos (dNTP) de terminación de cadena marcados-atrapados se bombea en la cámara de reacción. El tampón de reacción consiste en: Tris HCl 10 mM, pH 8,3; KCl 50 mM; y MgCl_2 2,5 mM. Los dNTP están cada uno a una concentración de 0,02-0,2 mM.
2. El tampón de reacción sin los dNTP se aclara a través de la cámara de reacción.

3. Para cada uno de los 10.000 centros de reacción, la identidad del nucleótido recién incorporado se determina mediante microscopía de fluorescencia de reflexión total interna (TIRFM). Se hacen múltiples registros de la matriz de centro de reacción de modo que cada uno de los cuatro nucleótidos sea distinguido. Los fluorocromos usados tienen altos coeficientes de extinción y/o altos rendimientos cuánticos para la fluorescencia. Además, los fluorocromos tienen bien resueltos los máximos de emisión y/o excitación. Hay varias familias de fluorocromos que serán usadas, por ejemplo, la familia BODIPY de fluorocromos (Molecular Probes, Inc.). Utilizando los fluorocromos BODIPY y el enlazante fotolábil 1-(4,5-dimetoxi-2-nitrofenil) etilo (DMNPE) puede emplearse el siguiente grupo de análogos de nucleótidos para SAND:

3'-O-(DMNPE-(BODIPY⁴⁹³/₅₀₃))-2' desoxi ATP

3'-O-(DMNPE-(BODIPY⁵³⁰/₅₅₀))-2' desoxi CTP

3'-O-(DMNPE-(BODIPY⁵⁶⁴/₅₇₀))-2' desoxi GTP

3'-O-(DMNPE-(BODIPY⁵⁸¹/₅₉₁))-2' desoxi TTP

Así las 'A' incorporadas se detectan con una iluminación láser con ión argón de 488 nm y un filtro de barrera centrado a 503 nm. Las 'C', 'G' y 'T' incorporadas son detectadas con una iluminación con láser YAG a 532 nm y filtros de barrera centrados en 550 nm, 570 nm y 591 nm, respectivamente.

Para cada uno de los acontecimientos de iluminación separados es generada una onda evanescente en la matriz del centro de reacción y la imagen de la matriz es enfocada a través del sistema microscópico en la cara de una cámara enfriada-CCD intensificada de placa de microcanal.

4. Los nucleótidos recién incorporados son liberados ópticamente por una iluminación con la luz de otro láser YAG a < 360 nm. Esto causa la disociación del DMNPE-BODIPY a partir de la cadena de ácido nucleico naciente dejándolo intacto y listo para incorporar al siguiente nucleótido.

5. La eliminación del resto fluorescente es verificada por TIRFM y el ciclo de reacción es repetido hasta que ya no sean incorporados más nucleótidos.

Típicamente el tiempo de exposición para cada fluorocromo es de 100 ms. El tiempo de lectura del chip CCD es $\sim 0,25$ s. A partir de esto, la etapa de detección para cada ciclo tarda $< 1,5$ s. El volumen total de la cámara de reacción es $1-10 \mu\text{l}$. Menos que un segundo lleva rellenar completamente la cámara de reacción. Por consiguiente, el tiempo total para un ciclo dado es menos de 10 segundos. Por lo tanto, en 10 segundos/ciclo cada uno de los 10.000 centros de reacción de la máquina SAND es capaz de deducir al menos 360 bases de secuencia por hora, correspondiente a 3,6 Mbase/hora de la secuencia deducida por la máquina SAND en total.

Los obturadores que controlan la iluminación láser, las ruedas de filtro que llevan los filtros de barrera y la cámara CCD son controlados todos por un micro-ordenador. La recolección de las imágenes y el análisis de datos son ejecutados todos por el mismo micro-ordenador. Los datos de secuencia extraídos y las imágenes de la matriz son almacenados permanentemente en una unidad CD ROM según son recogidos.

Aunque hayan sido descritas realizaciones particulares en este documento detalladamente, esto se ha hecho para ejemplificar con objetivos de ilustración sólo, y no se requieren para limitar, en lo que concierna, el alcance de las reivindicaciones añadidas que siguen. La opción de la polimerasa particular, el ligamiento particular de la polimerasa al soporte sólido, o los terminadores de nucleótido particulares, por ejemplo, según se piensa, son materiales de rutina para cualquier experto ordinario en la técnica con el conocimiento de las realizaciones descritas en este documento.

REIVINDICACIONES

1. Un método para la secuenciación de base de nucleótido que comprende las etapas secuenciales de:

- (a) introducir en una cámara de reacción que contiene una polimerasa inmovilizada sobre un soporte sólido una muestra de ácido nucleico, una pluralidad de cebadores de oligonucleótido diferentes, y una solución tamponada que contiene los cuatro nucleótidos diferentes, siendo cada nucleótido diferencialmente marcado con un grupo marcador desprendible y bloqueado en la porción 3' con un grupo bloqueante desprendible, en el que dicha muestra de ácido nucleico se hibrida a un cebador, y dicha polimerasa extiende dicho cebador incorporando un nucleótido diferencialmente marcado que es complementario al ácido nucleico de la muestra;
- (b) eliminar los nucleótidos que no han sido incorporados en el cebador;
- (c) determinar la identidad del nucleótido incorporado en el cebador de elongación detectando su grupo marcador, identificando así el complemento del nucleótido 3'-bloqueado marcado;
- (d) separar el grupo bloqueante 3' y el grupo marcador a partir del nucleótido incorporado;
- (e) eliminar el grupo bloqueante 3' separado y el grupo marcador separado de la etapa (d);
- (f) confirmar la separación y la eliminación del grupo bloqueante 3' a partir del nucleótido incorporado en el cebador monitorizando la presencia de nucleótidos marcados en la cámara de reacción mediante una segunda etapa de detección, en la que la presencia del nucleótido marcado en la cámara de reacción indica que el grupo bloqueante 3' no ha sido eliminado;
- (g) introducir en la cámara de reacción solución tamponada recién preparada que contiene dichos cuatro nucleótidos diferentes para iniciar de nuevo la síntesis; y
- (h) repetir las etapas de (b) a (g) hasta que ningún nuevo nucleótido sea incorporado en la etapa (g) o el grupo bloqueante 3' persista en no ser separado y eliminado en las etapas (d) y (e);

por el cual el orden, en el cual los nucleótidos marcados en la etapa (c) son detectados, corresponde al complemento de la secuencia de al menos una parte de la muestra de ácido nucleico.

2. El método de la reivindicación 1, en el que el grupo bloqueante 3' y el grupo marcador son separados en la etapa (d) a partir del nucleótido incorporado por activación fotoquímica.

3. El método de la reivindicación 1, en el que el grupo bloqueante 3' y el grupo marcador son separados en la etapa (d) a partir del nucleótido incorporado por activación química o enzimática.

4. El método de la reivindicación 1, en el que dicho grupo marcador diferencialmente marcado es un marcador fluorescente, una partícula resonante de plasmón o un marcador cuántico de punto.

5. El método de la reivindicación 1, en el que dicho grupo marcador se une directamente al grupo bloqueante 3' desprendible.

6. El método de la reivindicación 5, en el que dicho grupo bloqueante 3' desprendible es un grupo 2-nitrobencilo.

7. El método de la reivindicación 1, en el que dicho grupo marcador es unido a la base de cada nucleótido con un enlazante desprendible.

8. El método de la reivindicación 7, en el que dicho enlazante desprendible es un grupo 2-nitrobencilo.

9. El método de la reivindicación 1, en el que dicha polimerasa es seleccionada a partir de la ADN polimerasa, la ARN polimerasa y la transcriptasa inversa.

10. El método de la reivindicación 9, en el que dicha ADN polimerasa es seleccionada a partir de la ADN polimerasa de *Bacillus stearothermophilus*, la ADN polimerasa de *Thermus aquaticus*, la ADN polimerasa de *Pyrococcus furiosus*, la ADN polimerasa de *Thermococcus litoralis*, la ADN polimerasa de *Thermus thermophilus*, la ADN polimerasa del bacteriófago T4, la ADN polimerasa del bacteriófago T7, el fragmento de Klenow de la ADN polimerasa I de *E. coli*, y la ADN polimerasa III de *E. coli*.

11. El método de la reivindicación 9, en el que dicha ARN polimerasa es seleccionada a partir de la ARN polimerasa de *E. coli*, la ARN polimerasa del bacteriófago T3, la ARN polimerasa del bacteriófago T7, la ARN polimerasa del bacteriófago SP6, y las ARN polimerasas de familias virales de bromovirus, tobamovirus, tombusvirus, levivirus, virus tipo C de la hepatitis y picornavirus.

ES 2 310 514 T3

12. El método de la reivindicación 9, en el que dicha transcriptasa inversa es seleccionada a partir de la transcriptasa inversa del virus de *Avian Myeloblastosis*, la transcriptasa inversa del virus de la leucemia de Moloney murino, la transcriptasa inversa del virus 1 de la inmunodeficiencia humana, y la polimerasa T7 modificada.

13. El método de la reivindicación 1, en el que dicho nucleótido marcado es detectado por microscopía de fluorescencia de reflexión total interna, microscopía confocal de fotón, resonancia de plasmón de superficie y transferencia de energía de resonancia de fluorescencia.

14. El método de la reivindicación 1, en el que dicho soporte sólido comprende una matriz de centros de reacción, comprendiendo cada centro de reacción una molécula de polimerasa individual inmovilizada sobre dicho soporte sólido, y en el que dichos centros de reacción son monitorizados por un sistema microscópico y están presentes en una periodicidad mayor que la potencia de resolución de dicho sistema microscópico.

15. El método de la reivindicación 14, en el que la separación física de dichos centros de reacción es de al menos $0,2\ \mu\text{m}$.

16. El método de la reivindicación 14, en el que la separación física de dichos centros de reacción es de al menos $1\ \mu\text{m}$.

17. El método de la reivindicación 14, en el que la separación física de dichos centros de reacción es de al menos $2\ \mu\text{m}$.

18. El método de la reivindicación 14, en el que la separación física de dichos centros de reacción es de al menos $10\ \mu\text{m}$.

19. El método de la reivindicación 14, en el que en la introducción de dichos nucleótidos diferencialmente marcados, dichos centros de reacción son monitorizados usando dicho sistema microscópico hasta que una mayoría de sitios contenga la polimerasa inmovilizada unida a una plantilla de ácido nucleico con un nucleótido marcado diferencialmente incorporado individual.

20. El método de la reivindicación 19, en el que dichos sitios de reacción son monitorizados para detectar los fallos al incorporar un nucleótido o liberar un bloqueo del grupo bloqueante 3' en más de 2-20 ciclos.

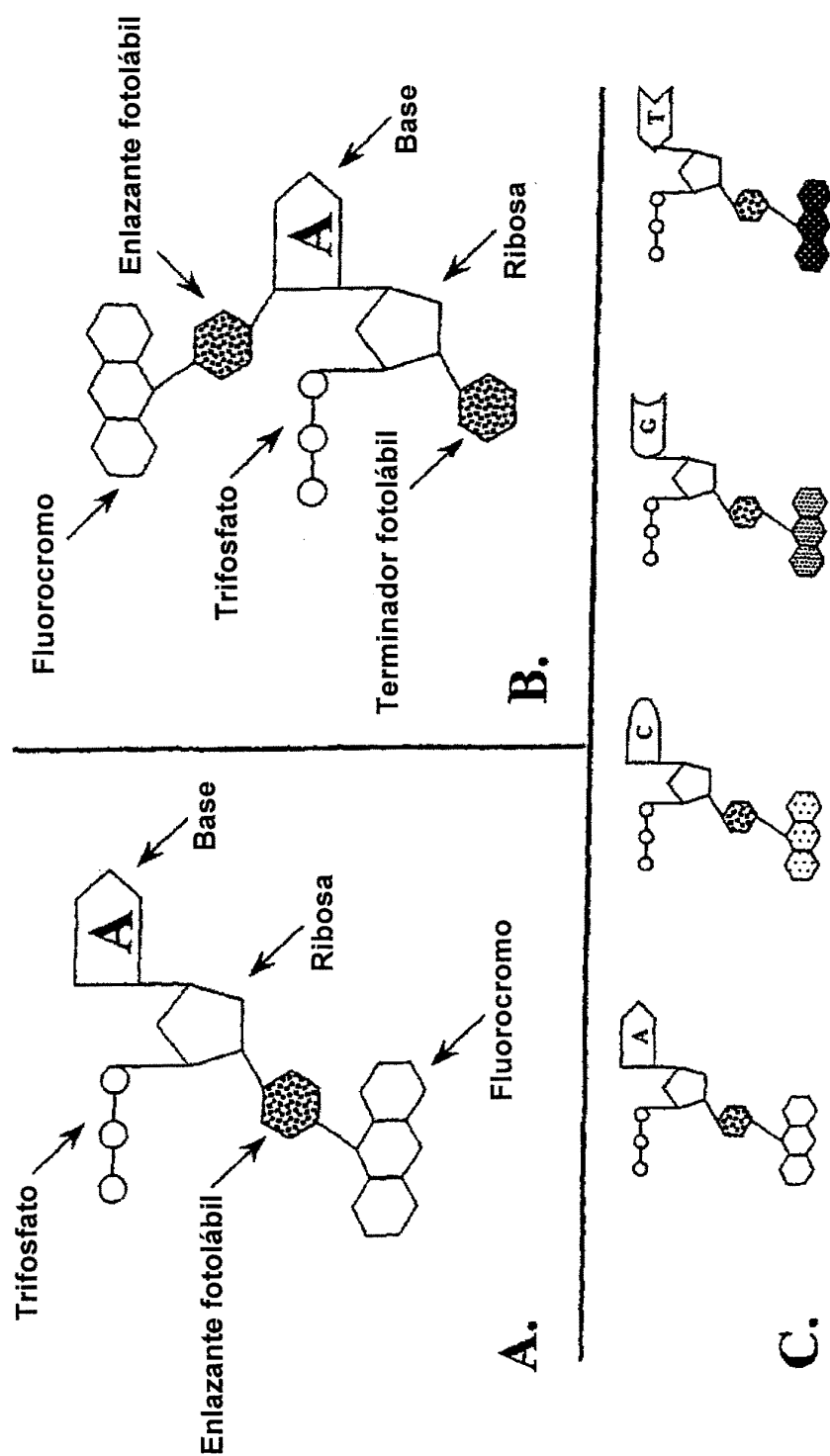


FIG. 1 Ejemplo de nucleótidos modificados para SAND

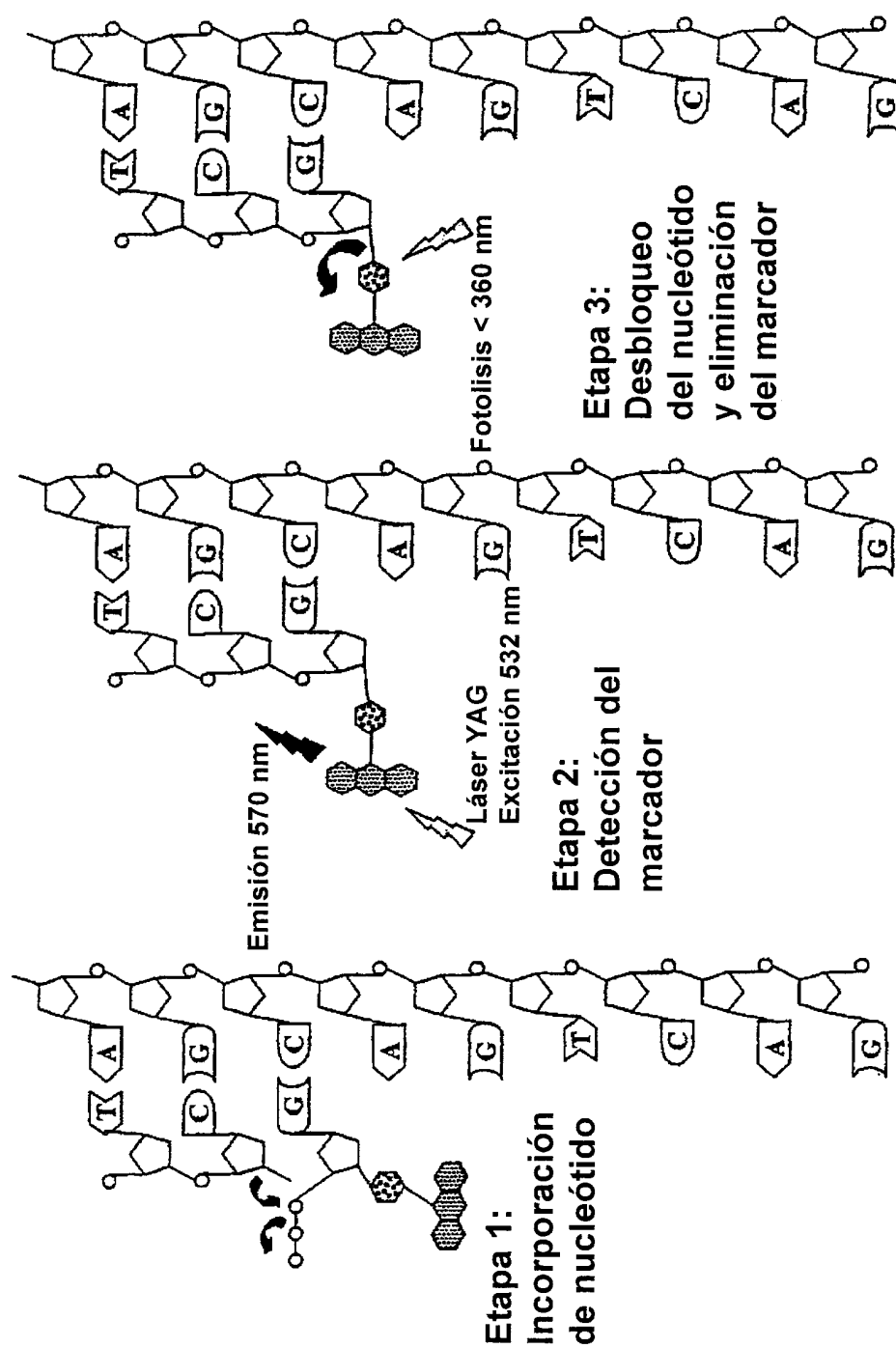


FIG. 2 El ciclo de reacción de SAND

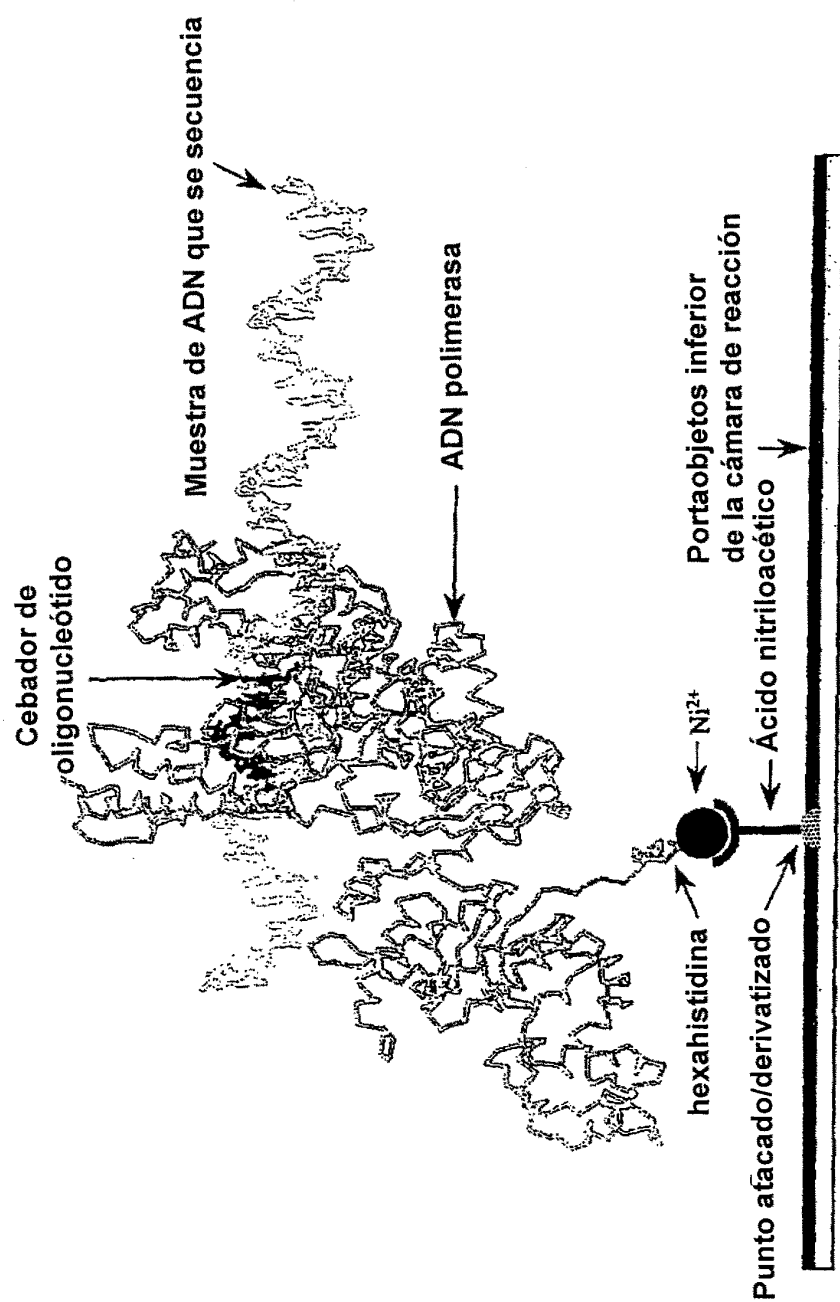


FIG. 3 Ejemplo de un centro de reacción de SAND

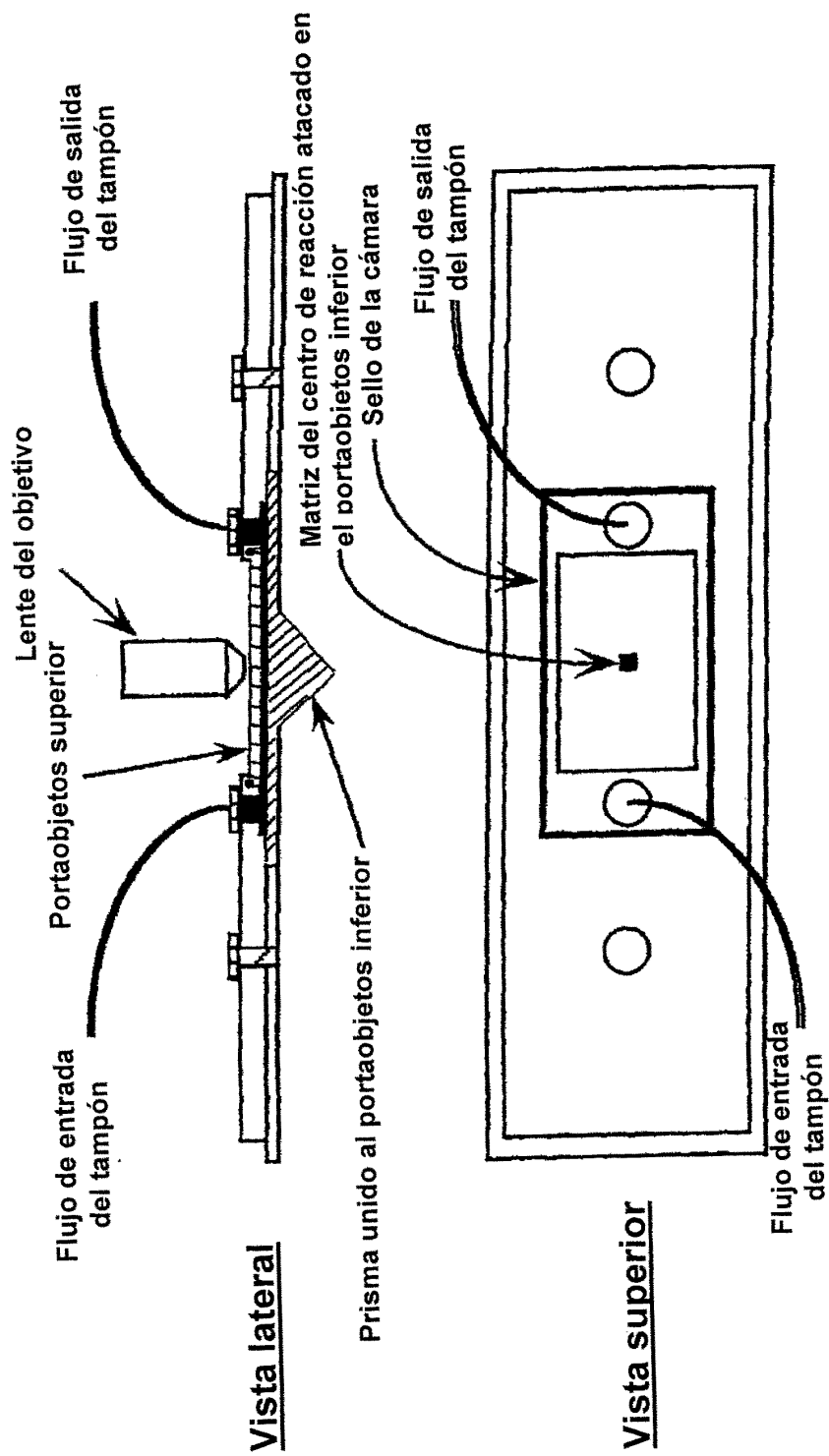


FIG. 4 Ensamblaje de la cámara de reacción de SAND

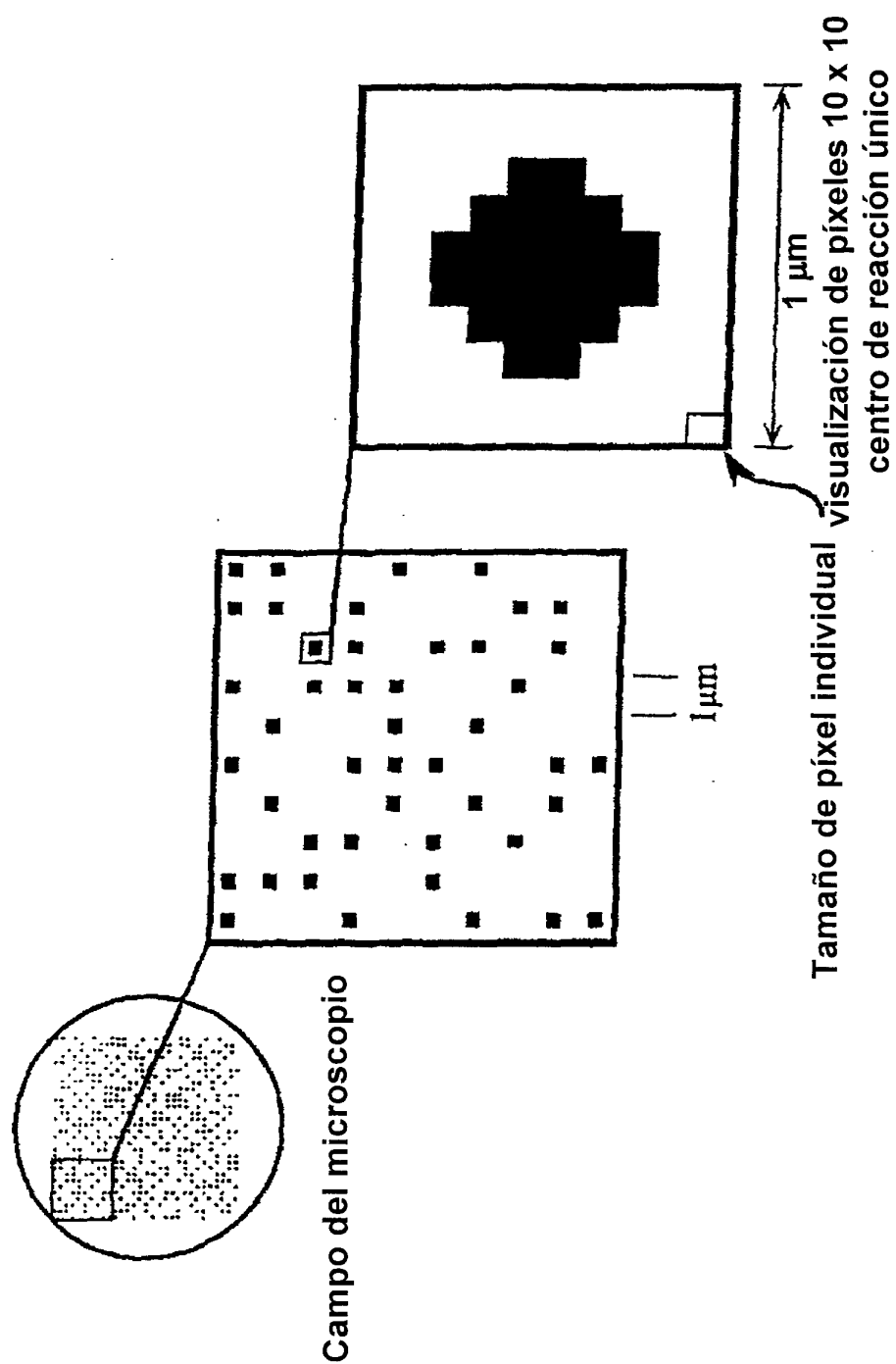


FIG. 5 Diagrama de la serie del centro de reacción de SAND

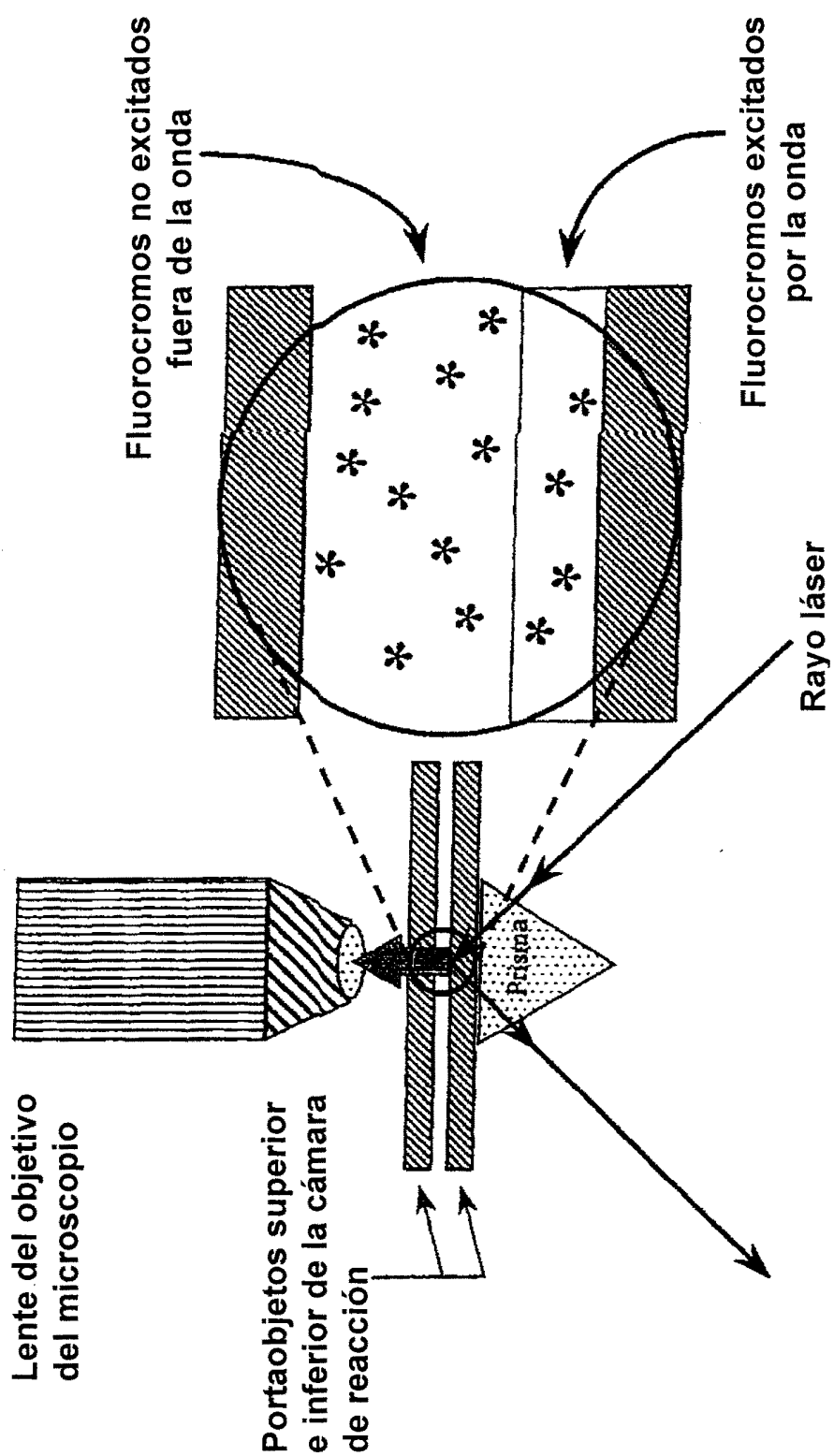


FIG. 6 Principio de la onda evanescente

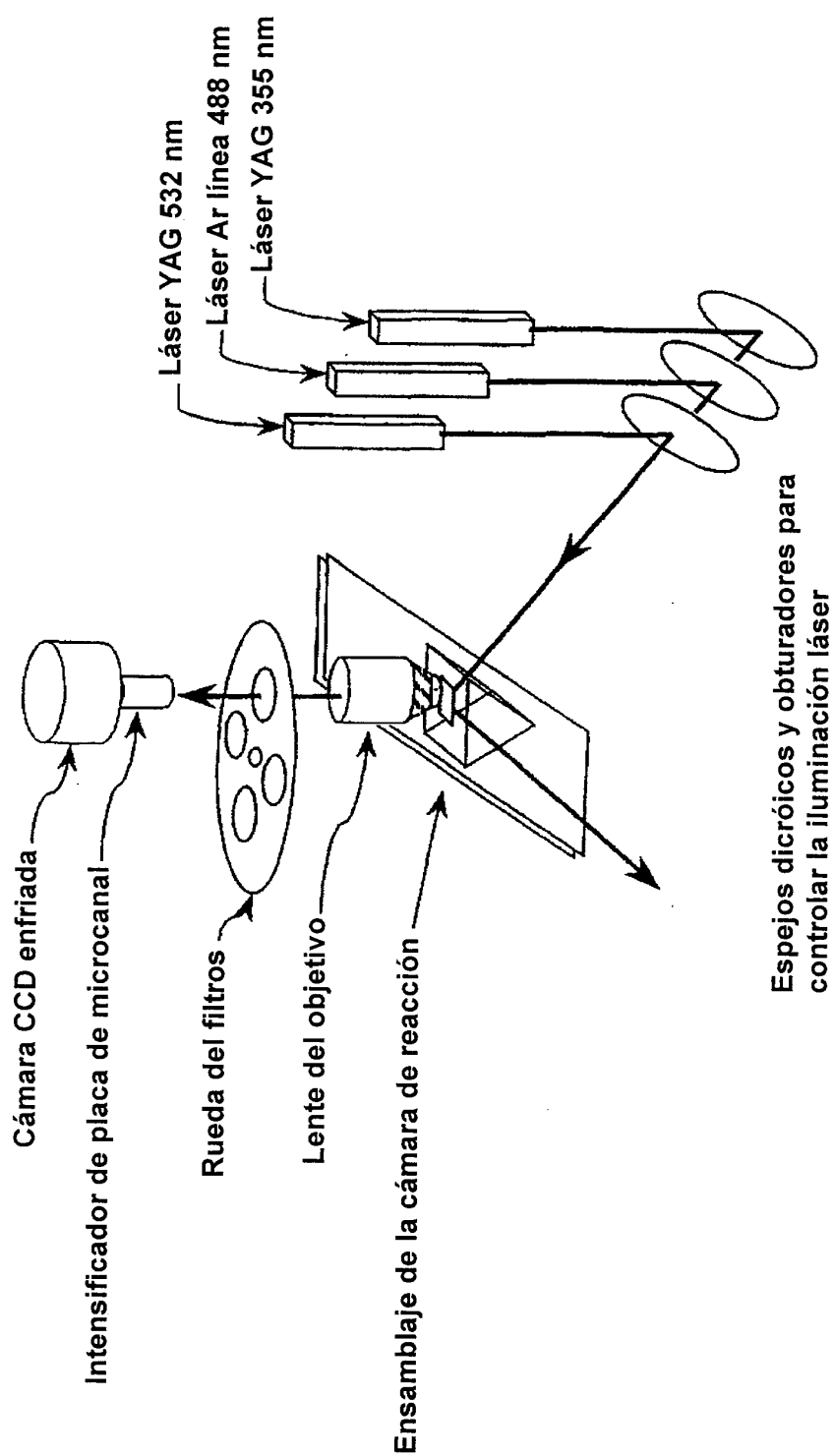


FIG. 7 Montaje para SAND usando TIRFM

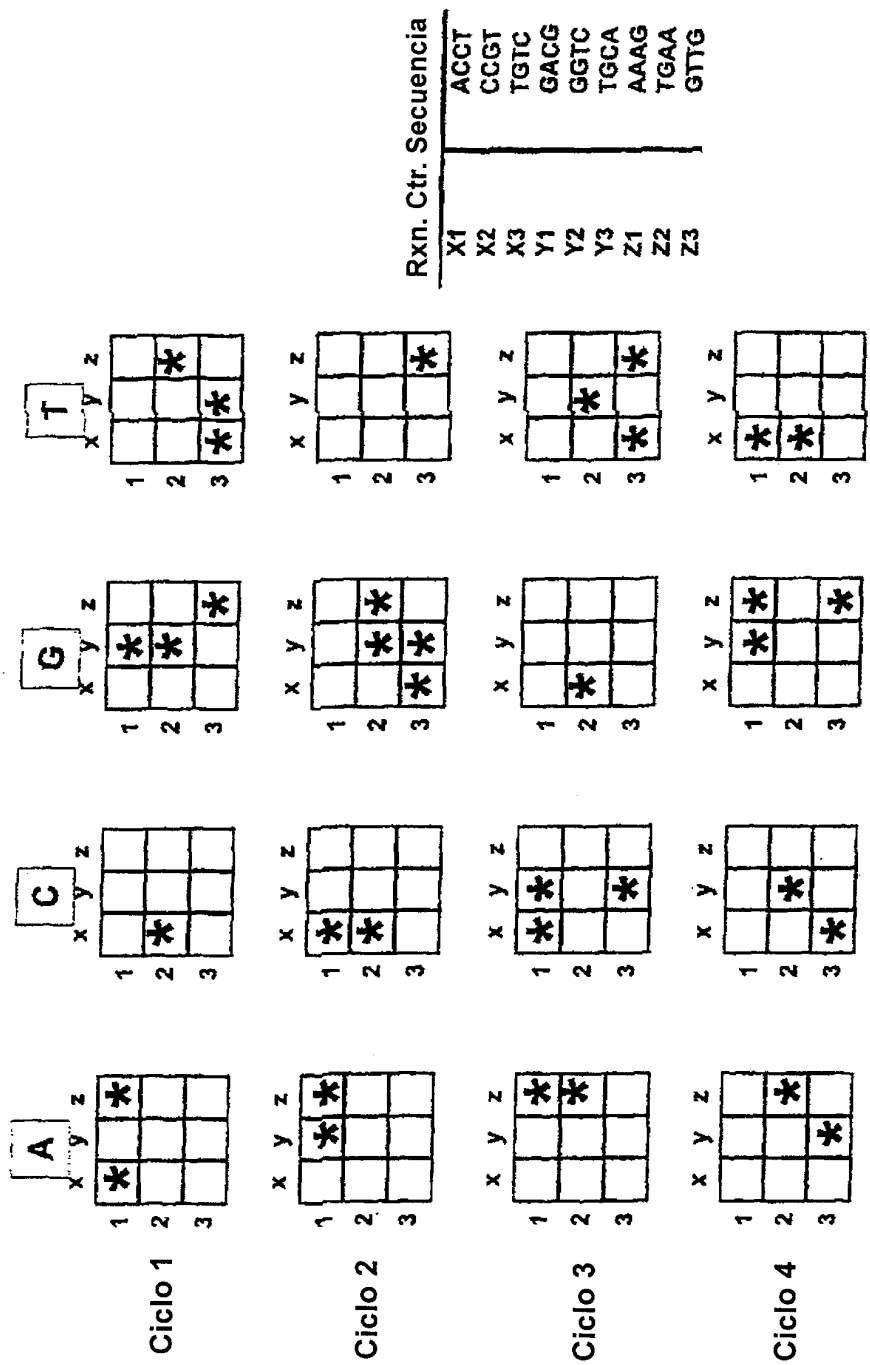


FIG. 8 Adquisición de datos del SAND: Un ejemplo con una matriz de 3x3