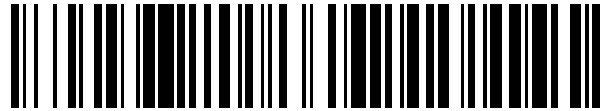


19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 563 643**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.04.1998 E 10011729 (0)**

97 Fecha y número de publicación de la concesión europea: **25.11.2015 EP 2327797**

54 Título: **Método de secuenciación de ácido nucleico**

30 Prioridad:

01.04.1997 GB 9706528

01.04.1997 GB 9706529

23.06.1997 GB 9713236

23.06.1997 GB 9713238

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

15.03.2016

73 Titular/es:

ILLUMINA CAMBRIDGE LIMITED (50.0%)

Chesterford Research Park

Little Chesterford Saffron Walden Essex CB10

1XL, GB y

ILLUMINA, INC. (50.0%)

72 Inventor/es:

KAWASHIMA, ERIC;

FARINELLI, LAURENT y

MAYER, PASCAL

74 Agente/Representante:

ZEA CHECA, Bernabé

ES 2 563 643 T3

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

DESCRIPCIÓN

Método de secuenciación de ácido nucleico

5 **[0001]** La presente invención se refiere, entre otras cosas, a la amplificación de ácidos nucleicos.

[0002] El desarrollo de la biología molecular y de la farmacología utiliza ahora exhaustivamente el análisis de ácidos nucleicos (Friedrich, G.A. Moving beyond the genome projects, *Nature Biotechnology* 14, 1234 (1996)). Las áreas más desafiantes son la secuenciación de todo el genoma, la detección de polimorfismos mononucleotídicos y la exploración y monitorización de la expresión génica. Actualmente, solo en proyectos de secuenciación de ADN, se han manejado hasta cientos de miles de muestras (Venter, J.C., H.O Smith, L. Hood, A new strategy for genome sequencing, *Nature* 381, 364 (1996)). Esta capacidad está limitada por la tecnología disponible. Proyectos tales como el "proyecto del genoma humano" (mapeo genético y secuenciación de ADN) e identificación de todos los polimorfismos en genes expresados que intervienen en enfermedades comunes, implican la secuenciación de millones de muestras de ADN.

[0003] Con la mayoría de las tecnologías de secuenciación de ADN actuales, simplemente no es posible disminuir indefinidamente el tiempo que se necesita para procesar una sola muestra. Una forma de aumentar el rendimiento es realizar muchos procesos en paralelo. La introducción de preparaciones y el suministro de muestras robóticas, de placas de 96 y 384 pocillos, de aparatos con rejilla de alta densidad (Maier, E., S. Meierewer, A.R. Ahmadi, J. Curtis, H. Lehrach, Application of robotic technology to automated sequence fingerprint analysis by oligonucleotide hybridization, *Journal Of Biotechnology* 35,191 (1994)) y recientemente el desarrollo de matrices oligonucleotídicas de alta densidad (Chee, M., R.Yang, E. Hubbell, A. Berno, X.C. Huang, D. Stern, J. Winkler, D.J. Lockhart, M.S. Morris, y S.P.A. Fodor, Accessing genetic information with high-density DNA arrays, *Science* 274(5287): 610-614, (1996)) están comenzado a dar respuestas a exigencias cada vez con un mayor rendimiento. Dichas tecnologías permiten procesar hasta 50.000-100.000 muestras a la vez, en el transcurso de días e incluso horas (Maier, E., Robotic technology in library screening, *Laboratory Robotics and Automation* 7, 123 (1995)).

[0004] En la mayoría de los métodos conocidos para realizar análisis de ácido nucleico, es necesario extraer primero los ácidos nucleicos de interés (por ejemplo, ADN genómico o mitocondrial o ARN mensajero (ARNm)) de un organismo. Después es necesario aislar los ácidos nucleicos de interés de la mezcla de todos los ácidos nucleicos y habitualmente, amplificar estos ácidos nucleicos para obtener cantidades idóneas para su caracterización y/o detección. El aislamiento de fragmentos de ácidos nucleicos se ha considerado necesario incluso cuando se está interesado en un conjunto representativo, aunque aleatorio, de todos los diferentes ácidos nucleicos, por ejemplo, un conjunto representativo de todos los ARNm presentes en una célula o de todos los fragmentos obtenidos después de haber cortado al azar el ADN genómico en trozos pequeños.

[0005] Para amplificar el ADN con medios biológicos se pueden utilizar diversos métodos que son muy conocidos por los expertos en la técnica. Generalmente, los fragmentos de ADN se insertan primero en vectores utilizando enzimas de restricción y ADN ligasas. Un vector que contiene un fragmento de interés puede introducirse después en un hospedador biológico y amplificarse mediante protocolos bien establecidos. Normalmente los hospedadores están dispersos al azar sobre un medio de cultivo (por ejemplo, placas de agar). Después pueden copiarse para proporcionar colonias que se originan de células hospedadoras individuales.

[0006] En dichos hospedadores se pueden realizar, de manera simultánea, hasta millones de amplificaciones simultáneas de fragmentos de ADN clonados. La densidad de colonias es del orden de 1 colonia/mm². Para obtener ADN de dichas colonias una opción es transferir las colonias a una membrana, y después inmovilizar el ADN desde dentro de los hospedadores biológicos directamente a la membrana (Grunstein, M. y D.S. Hogness, Colony Hybridization: A method for the isolation of cloned DNAs that contain a specific gene, *Proceedings of the National Academy of Science, USA*, 72: 3961 (1975)). Sin embargo, con estas opciones, la cantidad de ADN transferida está limitada y a menudo es insuficiente para la detección no radioactiva.

[0007] Otra opción es transferir, mediante técnicas de esterilización, individualmente cada colonia a un envase (por ejemplo, a placas de 96 pocillos) donde además se pueda producir la replicación de células las hospedadoras, de tal manera que pueda obtenerse más ADN de las colonias. Después los ácidos nucleicos amplificados pueden recuperarse de las células hospedadoras con un proceso de purificación apropiado. Sin embargo dicho procedimiento es generalmente laborioso y requiere mucho tiempo, y es difícil de automatizar.

[0008] La técnica revolucionaria de la amplificación de ADN utilizando la reacción en cadena de la polimerasa (PCR) se propuso en 1985 por Mullis et al. (Saiki, R., S. Scharf, F. Faloona, K. Mullis, G. Horn, H. Erlich y N. Arnheim, *Science* 230, 1350-1354 (1985)) y ahora es muy conocida por los expertos en la técnica. En este proceso de amplificación, un fragmento de ADN de interés puede amplificarse utilizando dos oligonucleótidos cortos (generalmente de aproximadamente 20 bases de longitud) que flanquean una región a amplificar, y que habitualmente reciben el nombre de "cebadores". La amplificación se produce durante los tiempos de ciclado de la PCR, que incluye una etapa durante la cual las moléculas de ADN bicatenarias se desnaturalizan (generalmente

calentando una mezcla de reacción, por ejemplo, a 95 °C para separar las moléculas de ADN bicatenario en dos fragmentos monocatenarios), una etapa de apareamiento (en la que la mezcla de reacción se lleva a una temperatura de, por ejemplo, 45 °C para permitir que los cebadores se apareen con los moldes monocatenarios) y una etapa de alargamiento (el ADN complementario al fragmento monocatenario se sintetiza mediante incorporación de nucleótidos secuencial en los extremos de los cebadores con la enzima ADN polimerasa).

[0009] El procedimiento anterior se realiza normalmente en solución, por lo cual no los cebadores ni el molde se ligan a ninguna matriz sólida.

10 **[0010]** Sin embargo, más recientemente, se ha propuesto el uso de un cebador injertado a una superficie junto con cebadores libres en solución para amplificar e injertar de manera simultánea un producto PCR sobre la superficie (Oroskar, A.A., S.E. Rasmussen, H.N. Rasmussen, S.R. Rasmussen, B.M. Sullivan, y A. Johansson, Detection of immobilised amplicons by ELISA-like techniques, *Clinical Chemistry* 42: 1547 (1996)). (El término "injerto" se usa en este documento para indicar que un resto comienza a unirse a una superficie y permanece ahí, a menos que se retire o se desee retirarlo). Generalmente la amplificación se realiza en envases (por ejemplo, en placas de formato 15 de 96 pocillos) de tal manera que cada envase contiene el producto (o los productos) de la PCR de una reacción. Con dichos métodos, algunos de los productos de la PCR comienzan a injertarse sobre una superficie del envase que tiene cebadores en su interior que se han puesto en contacto con el reactante durante el ciclado de la PCR. El injerto en la superficie simplifica ensayos posteriores y permite realizar una automatización eficiente.

20 **[0011]** La disposición de muestras de ADN se realiza más clásicamente en membranas (por ejemplo, membranas de nylon o nitrocelulosa). Con el uso de robótica adecuada (por ejemplo, Q-bot™, Genetix Ltd, Dorset BH23 3TG UK) es posible alcanzar una densidad de hasta 10 muestras/mm². En este caso, el ADN se liga por enlace covalente a una membrana por medios fisicoquímicos (por ejemplo, radiación UV). Estas tecnologías permiten la disposición de 25 moléculas de ADN grandes (por ejemplo, moléculas de más de 100 nucleótidos de longitud) así como moléculas de ADN más pequeñas. Por tanto, pueden disponerse tanto moldes como sondas.

[0012] Nuevas estrategias basadas en portaobjetos de vidrio previamente dispuestos (disposiciones de áreas reactivas obtenidas por tecnología de inyección (Blanchard, A.P. y L. Hood, Oligonucleotide array synthesis using ink jets, *Microbial and Comparative Genomics*, 1: 225 (1996)) o matrices de geles de poliacrilamida reactivos (Yershov, G. et al., DNA analysis and diagnostics on oligonucleotide micromicroplacas, *Proceedings of the National Academy of Science, USA*, 93: 4913-4918 (1996)) permiten la disposición de hasta 100 muestras/mm². Con estas tecnologías, solo se ha descrito el injerto de sondas (oligonucleótidos). El número indicado de muestras/mm² sigue siendo muy bajo (de 25 a 64).

35 **[0013]** Pueden alcanzarse densidades de muestra más altas utilizando microplacas de ADN, que pueden ser matrices de oligonucleótidos que se unen, mediante enlace covalente, a una superficie y que se pueden obtener utilizando técnicas de microlitografía (Fodor, S.P.A. et al., Light directed, spatially addressable parallel chemical synthesis, *Science* 251: 767(1991)). Actualmente, en aplicaciones de biología molecular, se utilizan microplacas con 40 625 sondas/mm² (Lockhart, D.J. et al., Expression monitoring by hybridisation to high-density oligonucleotide arrays, *Nature Biotechnology* 14: 1675 (1996)). Se reivindica que se pueden alcanzar densidades de sonda de hasta 250 000 muestras/cm² (Chee, M. et al., Accessing genetic information with high-density DNA arrays, *Science* 274: 610 (1996)). Actualmente, en una sola microplaca de aproximadamente 2,5 cm² se pueden disponer hasta 132 000 oligonucleótidos diferentes. En la actualidad, estas microplacas se fabrican mediante síntesis sólida de 45 oligonucleótidos en fase sólida con el extremo 3'OH del oligonucleótido unido a la superficie. Por tanto estas microplacas se han utilizado para proporcionar sondas oligonucleotídicas que no pueden actuar como cebadores en una etapa de alargamiento mediada por la ADN polimerasa.

[0014] Cuando los productos de la PCR se ligan en el recipiente en el que se realiza la amplificación por PCR, esto 50 puede considerarse como un proceso de disposición de matrices directo. La densidad de los productos resultantes de la PCR en la matriz, está por tanto limitada al recipiente que se dispone. Actualmente los recipientes de los que se dispone están solo en un formato de placa de microtitulación de 96 pocillos. Estos permiten obtener solo aproximadamente 0,02 muestras de productos PCR /mm² de superficie.

55 **[0015]** Utilizando el sistema Nucleolink™ disponible en el comercio, que puede obtenerse en Nunc A/S (Roskilde, Dinamarca), es posible conseguir la amplificación y disposición simultáneas de muestras en envases en la superficie sobre la cual se han injertado los cebadores oligonucleotídicos. Sin embargo, en este caso, la densidad de la matriz de las muestras la fija el tamaño del recipiente. En la actualidad se puede alcanzar una densidad de 0,02 60 muestras/mm² para el formato de placa de 96 pocillos. Aumentar esta densidad es difícil. Esto es apreciable ya que, por ejemplo, la disponibilidad de las placas de 384 pocillos (0,08 muestras/mm²), adecuadas para la PCR, se ha demorado debido a problemas técnicos (por ejemplo, transferencia de calor y efectos capilares durante el llenado). Por lo tanto, en un futuro previsible, con esta estrategia es poco probable poder conseguir mejoras en cuanto a los órdenes de magnitud en la densidad de las muestras dispuestas en una matriz.

65 **[0016]** La presente invención tiene por objeto superar, o al menos aliviar, algunas de las desventajas de los

métodos de secuenciación de ácido nucleico de la técnica anterior. La presente invención se ocupa de un método de secuenciación de una molécula de ácido nucleico como se define en las reivindicaciones.

[0017] De acuerdo con la presente descripción se desvela un método de amplificación de ácido nucleico que comprende las etapas de:

A. proporcionar una pluralidad de cebadores que están inmovilizados pero que tienen un extremo expuesto para permitir la extensión del cebador;

10 B. permitir que una molécula de ácido nucleico diana monocatenario se aparee con uno de dicha pluralidad de cebadores sobre parte de la longitud de dicha molécula de ácido nucleico monocatenario y después extender ese cebador utilizando, como molde, la molécula de ácido nucleico monocatenario apareada, para proporcionar una cadena de ácido nucleico inmovilizada extendida;

C. separar la molécula de ácido nucleico diana de la cadena de ácido nucleico inmovilizada extendida;

15 D. permitir que la cadena de ácido nucleico inmovilizada extendida se aparee con uno de dicha pluralidad de cebadores indicados en la etapa A) y después extender ese cebador utilizando, como molde, la cadena de ácido nucleico inmovilizada extendida, para proporcionar otra cadena de ácido nucleico inmovilizada extendida; y opcionalmente,

E. separar las cadenas de ácido nucleico inmovilizadas extendidas apareadas entre sí.

20 **[0018]** Preferentemente el método también comprende la etapa de:

F. utilizar al menos una cadena de ácido nucleico inmovilizada extendida para repetir las etapas D) y E), para proporcionar cadenas de ácido nucleico inmovilizadas extendidas adicionales y, opcionalmente,

25 G. repetir la etapa F) una vez o más.

[0019] A ser posible la secuencia de ácido nucleico diana monocatenario se proporciona mediante un método en el cual dicho ácido nucleico diana monocatenario se produce proporcionando una secuencia de ácido nucleico determinada a amplificar (cuya secuencia puede ser conocida o no) y añadir a la misma una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico; en el que dicha primera secuencia de ácido nucleico se hibrida con uno de dicha pluralidad de cebadores y dicha segunda secuencia de ácido nucleico es complementaria a una secuencia que se hibrida a una de dicha pluralidad de cebadores.

[0020] La segunda secuencia de ácido nucleico puede ser una secuencia que es igual que la secuencia de uno de la pluralidad de cebadores. Por tanto, la secuencia de ácido nucleico diana monocatenario puede proporcionarse mediante un método en el que dicho ácido nucleico diana monocatenario se produzca proporcionando una secuencia de ácido nucleico determinada a amplificar (cuya secuencia puede conocerse o no) y añadir a la misma una primera secuencia de ácido nucleico y una segunda secuencia de ácido nucleico; en el que dicha primera secuencia de ácido nucleico se hibrida con uno de dicha pluralidad de cebadores y dicha segunda secuencia de ácido nucleico es igual que la secuencia de uno de dicha pluralidad de cebadores.

[0021] La primera y segunda secuencias de ácido nucleico pueden proporcionarse en el primer y segundo extremos de dicho ácido nucleico diana monocatenario, aunque esto no es esencial.

[0022] Si se desea puede proporcionarse un marcador para permitir identificar los productos de amplificación de una secuencia de ácido nucleico determinada.

Colonias

[0023] El método de la presente descripción permite proporcionar una o más áreas distintas, comprendiendo cada área distinta una pluralidad de cadenas de ácido nucleico inmovilizadas (de ahora en adelante en el presente documento denominadas "colonias"). Estas áreas pueden contener grandes cantidades de moléculas de ácido nucleico amplificado. Estas moléculas pueden ser moléculas de ADN y/o ARN y pueden proporcionarse en forma mono o bicatenaria. Tanto una cadena determinada como su cadena complementaria pueden proporcionarse en forma amplificada en una sola colonia.

[0024] Pueden proporcionarse colonias de cualquier tamaño particular.

[0025] Sin embargo, en toda su dimensión, el tamaño de las colonias preferidas es de 10 nm a 100 μm , más preferentemente de 100 nm a 10 μm . A ser posible, gran parte de las colonias presentes en una superficie (es decir, al menos el 50 % de la misma) presentan tamaños con los intervalos indicados anteriormente.

[0026] Las colonias pueden disponerse de una manera predeterminada o al azar. Son posibles configuraciones de colonias bi o tridimensionales. Las configuraciones pueden ser regulares (por ejemplo, tener un diseño poligonal o generalmente circular) o pueden ser irregulares.

65

[0027] Las colonias pueden proporcionarse a altas densidades. Pueden alcanzarse densidades de más de una colonia/mm² de superficie. De hecho, utilizando la presente invención, pueden alcanzarse densidades de más de 10², más de 10³ o incluso más de 10⁴ colonias/mm². En realizaciones preferidas, la presente invención proporciona densidades de colonia de 10⁴⁻⁵ colonias/mm², más preferentemente densidades de 10⁶⁻⁷ colonias/mm², por tanto se ofrece una mejora de 3 a 4 órdenes de magnitud con respecto a las densidades que pueden alcanzarse utilizando muchos de los métodos de la técnica anterior. Esta propiedad de la invención es lo que permite una gran ventaja sobre la técnica anterior, dado que la alta densidad de colonias de ADN permite disponer al azar y amplificar una gran cantidad de moldes de ADN (hasta 10⁶⁻⁷ colonias/mm²).

10 **Cebadores**

[0028] Los cebadores inmovilizados para su uso en la presente invención se pueden proporcionar mediante cualquier medio adecuado, siempre que su extremo 3'-OH libre esté disponible para la extensión del cebador. Cuando se van a amplificar muchas moléculas de ácido nucleico diferentes, se pueden proporcionar muchos cebadores diferentes. Como alternativa, se pueden utilizar cebadores "universales", por lo cual solo se puede utilizar uno o dos tipos de cebadores diferentes (dependiendo de la realización de la invención) para amplificar las moléculas de ácido nucleico diferentes. Se puede utilizar cebadores universales cuando las moléculas que se van a amplificar comprenden una primera y segunda secuencias, como se describe anteriormente. La adquisición de cebadores universales es ventajosa sobre métodos tales como los desvelados en el documento WO96/04404 (Mosaic Technologies, Inc.) en los que deben prepararse cebadores específicos para cada secuencia particular que se va a amplificar.

[0029] Los cebadores oligodesoxinucleotídicos sintéticos están disponibles en el comercio de muchos proveedores (por ejemplo, Microsynth, Suiza, Eurogentech, Bélgica).

[0030] Se ha descrito el injerto de cebadores sobre vidrio o cuarzo silanizado y sobre obleas de silicio o superficies de oro (Maskos, U. y E.M Southern, Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in situ, *Nucleic Acids Research* 20(7): 1679-84, 1992; Lamture, J.B., et al. Direct-detection of nucleic-acid hybridization on the surface of a charge-coupled-device, *Nucleic Acids Research* 22(11): 2121-2125, 1994; Chrisey, L.A., G.U. Lee, y C.E. Oferrall, Covalent attachment of synthetic DNA to self-assembled monolayer films, *Nucleic Acids Research* 24(15): 3031-3039, 1996).

[0031] El injerto de cebadores marcados con biotina en soportes cubiertos con estreptavidina es otra alternativa. Ese método de injerto se usa normalmente para bio macromoléculas en general.

[0032] En la presente invención también es posible el injerto no covalente de cebadores en la interfaz entre una fase acuosa y una fase hidrófoba a través de un anclaje hidrófobo. Dicho anclaje se usa normalmente para bio macromoléculas en general (S. Terrettaz et al.: Protein binding to supported lipid membranes, *Langmuir* 9,1361 (1993)). Formas preferidas de dichas interfases serían liposomas, vesículas lipídicas, emulsiones, biocapas con diseño, películas de Langmuir o de Langmuir-Blodgett. Los diseños se pueden obtener diseñando directamente en moldes, por ejemplo, en microplacas de silicio diseñadas a través de métodos microlitográficos (Goves, J.T. et al., Micropatterning Fluid Bilayers on Solid Supports, in *Science* 275,651 (1997)). Los diseños también se pueden obtener debido a las propiedades de auto ensamblaje de los "coloides", por ejemplo, emulsiones o partículas de látex (Larsen, A.E. y D.G. Grier, Like charge attractions in metastable colloidal crystallites, *Nature* 385,230 (1997)).

[0033] En los métodos anteriores, en una superficie se pueden injertar dos o más cebadores diferentes. Los cebadores se pueden injertar de manera homogénea y simultánea sobre la superficie.

[0034] Utilizando métodos microlitográficos es posible proporcionar cebadores inmovilizados de una manera controlada. Si se desea la síntesis directa de oligonucleótidos sobre un soporte sólido con un extremo 3'-OH libre, entonces se pueden utilizar métodos microlitográficos para sintetizar de manera simultánea muchos cebadores oligonucleotídicos diferentes (Pirung, M.C. y Bradley, J.C. Comparison of methods for photochemical phosphoramidite-based DNA- synthesis. *Journal of Organic Chemistry* 60(20): 6270-6276, 1995). Estos se pueden proporcionar en distintas áreas cuya configuración puede corresponder con la de las colonias que se van a formar (por ejemplo, puede ser a través de varios nanómetros o micrómetros). Dentro de cada área, es necesario proporcionar solo un único tipo de cebador oligonucleotídico. Como alternativa se puede proporcionar una mezcla que comprenda una pluralidad de cebadores diferentes. En cualquier caso, dentro de cada área, los cebadores se deben distribuir de manera homogénea. Estos se pueden proporcionar en forma de una matriz regular.

[0035] Cuando las áreas comprenden inicialmente un solo tipo de cebador inmovilizado, estas pueden modificarse, si se desea, para llevar dos o más tipos de cebadores diferentes. Una manera de conseguir esto es utilizar moléculas como moldes para la extensión del cebador que tenga extremos 3' que se hibriden con un solo tipo de cebador inicialmente presente y que tengan extremos 5' que se extiendan más allá de los extremos 3' de dichos cebadores. Proporcionando una mezcla de moldes con diferentes secuencias entre sí, la extensión del cebador de un tipo de cebador utilizando la mezcla de dichos moldes seguido por la separación de cadenas dará como resultado

cebadores modificados diferentes. (en el presente documento los cebadores modificados se denominan cebadores “extendidos” para diferenciarlos de los cebadores “primarios” inicialmente presentes en una superficie).

5 **[0036]** De esta manera, se pueden proporcionar uno, dos o más tipos de cebadores diferentes extendidos en cualquier área donde se localicen inicialmente cebadores primarios. Si se desea, se pueden utilizar partes sustancialmente iguales de moldes diferentes para dar proporciones sustancialmente iguales de diferentes tipos de cebadores extendidos inmovilizados sobre un área determinada. Por consiguiente entonces, si se desean diferentes proporciones de diferentes cebadores extendidos inmovilizados, esto puede conseguirse ajustando las proporciones de las diferentes moléculas molde inicialmente utilizadas.

10

[0037] Dentro del cebador puede haber un sitio de escisión de endonucleasas de restricción. También se puede proporcionar un cebador con un sitio de reconocimiento de endonucleasas de restricción que dirijan la escisión del ADN con una distancia de varias bases (endonucleasas de restricción de Tipo II). (Para evitar dudas, se considera que dichos sitios están presentes incluso si se requiere que, tanto el cebador como su complemento, estén presentes en una molécula bicatenaria para que se produzca el reconocimiento y/o la escisión). Como alternativa, cuando se extiende un cebador, puede producirse un sitio de escisión y/o de reconocimiento. En cualquier caso, las endonucleasas de restricción pueden ser útiles permitiendo que una molécula de ácido nucleico, inmovilizada en una colonia, se escinda para liberar al menos una parte de la misma. Como una alternativa al uso de otras endonucleasas de restricción, se pueden utilizar ribozimas para liberar al menos partes de las moléculas de ácido nucleico de una superficie (cuando dichas moléculas son moléculas de ARN). Son posibles otros métodos. Por ejemplo, si se utiliza un enlace covalente para ligar un cebador a una superficie, este enlace se puede degradar (por ejemplo, por medios químicos, físicos o enzimáticos).

15 **[0038]** Los cebadores para su uso en la presente invención tienen preferentemente una longitud de al menos cinco bases. Normalmente, tendrán una longitud menor de 100 o menor de 50 bases. Sin embargo esto no es esencial. En los cebadores puede haber bases de origen natural y/o no natural.

Moléculas de ácido nucleico diana

20 **[0039]** Volviendo ahora a las moléculas de ácido nucleico diana (denominadas también en el presente documento “moldes”) para su uso en el método de la presente invención, estas pueden proporcionarse mediante cualquier medio apropiado. Una molécula diana (cuando está en forma monocatenaria) comprende una primera parte que tiene una secuencia que se puede aparear con un primer cebador y una segunda parte que tiene una secuencia complementaria a una secuencia que se puede aparear con un segundo cebador. En una realización preferida la segunda parte tiene la misma secuencia que el segundo cebador.

25 **[0040]** El segundo cebador puede tener una secuencia que sea igual que, o diferente de, la secuencia del primer cebador.

30 **[0041]** La primera y segunda partes de las moléculas de ácido nucleico diana se localizan preferentemente en sus extremos 3' y 5' respectivamente. Sin embargo esto no es esencial. La molécula diana también comprenderá normalmente una tercera parte localizada entre la primera y segunda partes. Esta parte de la molécula comprende una secuencia particular que se va a copiar. Esta puede ser, si se desea, de cualquier fuente y puede tener una secuencia conocida o desconocida (algunas veces denominada “anónima”). Esta puede proceder, por ejemplo, de fraccionamiento al azar por medios mecánicos o por digestión con enzimas de restricción limitada de una muestra de ácido nucleico.

35 **[0042]** Si se desea se pueden proporcionar otras partes de las moléculas diana. Por ejemplo, se pueden proporcionar partes diseñadas que actúan como etiquetas. Una “etiqueta” se define por su función de permitir identificar una molécula de ácido nucleico (o su complemento) particular.

40 **[0043]** Independientemente de que estén presentes las partes, las moléculas de ácido nucleico diana pueden proporcionarse mediante técnicas conocidas por los expertos en la materia de manipulación de ácidos nucleicos. Por ejemplo, dos o más partes pueden unirse entre sí por ligamiento. Si es necesario, antes del ligamiento se pueden realizar modificaciones apropiadas para proporcionar moléculas en una forma lista para el ligamiento. Por ejemplo, si se desea un ligamiento de extremos romos entonces se puede utilizar una exonucleasa específica monocatenaria, tal como la nucleasa S1, para retirar las partes monocatenarias de las moléculas antes del ligamiento. También se pueden utilizar enlazadores y/o adaptadores en la manipulación de ácido nucleico. (Se desvelan técnicas útiles de manipulación de ácido nucleico, por ejemplo, en Sambrook et al, Molecular Cloning, 2ª edición, Cold Spring Harbor Laboratory Press (1989)).

45 **[0044]** Una vez que se ha sintetizado una molécula molde, esta puede clonarse en un vector y puede amplificarse en un hospedador adecuado antes de utilizarse en la presente invención. Como alternativa, esta puede amplificarse por PCR. Como una alternativa adicional, pueden sintetizarse químicamente lotes de moléculas molde utilizando sintetizadores de ADN automáticos (por ejemplo, de Perkin-Elmer/Applied Biosystems, Foster City, CA).

[0045] Sin embargo es importante observar que la presente invención permite proporcionar grandes cantidades de moléculas de ácido nucleico de idéntica secuencia en una colonia que se origina de una sola molécula molde. Adicionalmente, el molde se puede volver a utilizar para generar otras colonias. Por tanto, para su uso en la formación de colonias, no es esencial proporcionar grandes cantidades de moléculas molde.

5

[0046] El molde puede tener cualquier longitud que se desee, siempre que pueda participar en el método de la presente invención. Preferentemente tiene una longitud de al menos 10, más preferentemente de al menos 20 bases. Más preferentemente tiene una longitud de al menos 100 o de al menos 1000 bases. Como es el caso de los cebadores para su uso en la presente invención, los moldes pueden comprender bases de origen natural y/o no

10 natural.

Condiciones de reacción

[0047] Volviendo ahora a las condiciones de reacción adecuadas para el método de la presente invención, se apreciará que la presente invención utiliza etapas repetitivas de apareamiento de cebadores con moldes, extensión de cebadores y separación de cebadores extendidos de moldes. Estas etapas se pueden realizar generalmente utilizando reactivos y condiciones que conocen los expertos en técnicas de PCR (o transcriptasa inversa más PCR). Se desvelan técnicas de PCR, por ejemplo, en "PCR: Clinical Diagnostics and Research", publicado en 1992 por Springer-Verlag.

20

[0048] Por tanto puede utilizarse una polimerasa de ácido nucleico junto con un aporte de moléculas de nucleósido trifosfato (u otras moléculas que actúan como precursores de nucleótidos presentes en ADN/ARN, tales como nucleósido trifosfatos modificados) para extender cebadores en presencia de un molde adecuado.

[0049] De manera deseable, se proporciona un exceso de desoxirribonucleósido trifosfatos. Los desoxirribonucleósido trifosfatos preferidos se abrevian de la siguiente manera: dTTP (desoxitimidina nucleósido trifosfato), dATP (desoxiadenosina nucleósido trifosfato), dCTP (desoxicitosina nucleósido trifosfato) y dGTP (desoxiguanosina nucleósido trifosfato). Los ribonucleótido trifosfatos preferidos son UTP, ATP, CTP y GTP. Sin embargo son posibles alternativas. Pueden ser de origen natural o no natural. También puede proporcionarse un tampón del tipo generalmente usado en las reacciones de la PCR.

30

[0050] Para incorporar nucleótidos durante la extensión de los cebadores se usa una polimerasa de ácido nucleico que es preferentemente estable en las condiciones de reacción pertinentes para que pueda utilizarse varias veces. (Esto es particularmente útil en procedimientos de amplificación automáticos). Por tanto, cuando se usa calentamiento para separar de su molde una cadena de ácido nucleico recién sintetizada, es preferible que la polimerasa de ácido nucleico sea termoestable a la temperatura utilizada. Los expertos en la técnica conocen dichas polimerasas termoestables. Estas polimerasas pueden obtenerse de microorganismo termófilos y se incluye la ADN polimerasa dependiente de ADN conocida como *Taq* polimerasa y también sus derivados termoestables. (Sin embargo, no es necesario que la polimerasa de ácido nucleico sea dependiente de ADN y puede ser dependiente de

40 ARN. Por tanto puede ser una transcriptasa inversa, es decir, una ADN polimerasa dependiente de ARN).

[0051] Generalmente, el apareamiento de un cebador con su molde tiene lugar a una temperatura de 25 a 90 °C. Dicho intervalo de temperatura normalmente se mantendrá durante la extensión del cebador. Una vez transcurrido un tiempo suficiente como para permitir el apareamiento y también para permitir un grado de extensión del cebador deseado, la temperatura puede aumentarse, si se desea, para permitir la separación de la cadena. En esta fase la temperatura se aumentará generalmente a una temperatura de 60 a 100°C. (También se pueden utilizar temperaturas elevadas para reducir problemas de cebado no específicos antes del apareamiento). Estas se pueden utilizar para controlar el tiempo de inicio de la colonia, por ejemplo, para sincronizar el inicio de la colonia para diversas muestras. Como alternativa, las cadenas pueden separarse por tratamiento con una solución de baja salinidad y alto pH (>12) o utilizando una sal caotrópica (por ejemplo clorhidrato de guanidinio) o mediante un disolvente orgánico (por ejemplo formamida).

50

[0052] Después de la separación de la cadena (por ejemplo, por calentamiento), se realizará preferentemente una etapa de lavado. La etapa de lavado puede omitirse entre rondas de apareamiento iniciales, extensión del cebador y separación de la cadena, y si se desea, para mantener los mismos moldes cerca de los cebadores inmovilizados. Esto permite utilizar los moldes varias veces para iniciar la formación de la colonia. (Es preferible proporcionar una alta concentración de moléculas molde inicialmente para que se inicien muchas colonias en una fase).

55

[0053] El tamaño de las colonias puede controlarse, por ejemplo, controlando el número de ciclos de apareamiento, extensión de cebador y separación de cadenas que se producen. También pueden controlarse otros factores que afectan al tamaño de las colonias. Estos incluyen el número y la disposición sobre una superficie de cebadores inmovilizados, la conformación de un soporte sobre el cual se inmovilizan los cebadores, la longitud y rigidez del molde y/o de las moléculas cebadoras, la temperatura y la fuerza iónica y la viscosidad de un fluido en el que se pueden realizar los ciclos mencionados anteriormente.

65

Usos de las colonias

[0054] Una vez que se han formado las colonias estas se pueden utilizar para cualquier propósito deseado.

5 **[0055]** Por ejemplo, se pueden utilizar en la secuenciación de ácido nucleico (ya sea parcial o completa), en diagnóstico, en exploración, como soportes para otros componentes y/o con fines de investigación (los usos preferidos se describirán con detalle más adelante). Si se desea, las colonias pueden modificarse para proporcionar diferentes colonias (denominadas en el presente documento "colonias secundarias" para diferenciarlas de las "colonias primarias" inicialmente formadas).

10

Superficies que comprenden cadenas de ácido nucleico inmovilizadas

[0056] Una superficie que comprende cadenas de ácido nucleico inmovilizadas en forma de colonias de moléculas de ácido nucleico monocatenario también se incluye en el ámbito de la presente invención.

15

[0057] Normalmente cada cadena de ácido nucleico inmovilizada dentro de una colonia se localizará en la superficie de tal manera que una cadena de ácido nucleico inmovilizada y complementaria a la misma se localiza en la superficie a una distancia de la longitud de dicha cadena de ácido nucleico inmovilizada (es decir, en la longitud de la molécula). Esto permite proporcionar cadenas de ácido nucleico y sus complementarias a muy alta densidad en forma inmovilizada. Preferentemente serán proporciones sustancialmente iguales a las de una cadena de ácido nucleico determinada y su complementaria dentro de una colonia. Preferentemente, dentro de la colonia, una cadena de ácido nucleico y su complementaria se distribuirán sustancialmente de manera homogénea.

20 **[0058]** También es posible proporcionar una superficie que comprenda cadenas de ácido nucleico monocatenario en forma de colonias, en las que, en cada colonia, las cadenas sencillas codificantes y no codificantes se proporcionen en una forma tal que las dos cadenas ya no son en absoluto complementarias, o sean simplemente en parte complementarias. Dichas superficies también se incluyen en el ámbito de la presente invención. Normalmente, dichas superficies se obtienen después de tratar colonias primarias, por ejemplo, por digestión parcial con enzimas de restricción o por digestión parcial después de la separación de las cadenas (por ejemplo, después de calentamiento) mediante una enzima que digiere ADN monocatenario, o por medios químicos o físicos (por ejemplo, irradiando con luz colonias que se han teñido con un colorante intercalante, por ejemplo, bromuro de etidio).

25 **[0059]** Una vez proporcionadas las colonias monocatenarias, estas se pueden utilizar para proporcionar moléculas bicatenarias. Esto se puede realizar, por ejemplo, proporcionando un cebador adecuado (preferentemente en solución) que se hibrida a los extremos 3' de moléculas monocatenarias inmovilizadas y después extendiendo ese cebador utilizando una polimerasa de ácido nucleico y un aporte de nucleósido trifosfatos (u otros precursores nucleotídicos).

30 **[0060]** Por tanto, las superficies que comprenden colonias de moléculas de ácido nucleico monocatenario que no tienen puente también se encuentran dentro del ámbito de la presente invención. (La expresión "que no tienen puente" se utiliza en el presente documento para indicar que las moléculas no tienen forma de estructuras de tipo puente como las mostradas, por ejemplo, en la figura 1h).

35 **[0061]** Utilizando la presente invención, se pueden proporcionar pequeñas colonias que contienen grandes cantidades de moléculas de ácido nucleico (bien mono o bicatenario). Por lo tanto, muchas colonias pueden localizarse sobre una superficie que tenga un área pequeña. Por tanto, las densidades de colonias que se pueden obtener pueden ser muy altas, como se ha indicado anteriormente.

40 **[0062]** Generalmente, diferentes colonias comprenderán diferentes cadenas de ácido nucleico amplificadas y cadenas complementarias amplificadas de las mismas. Por tanto, la presente invención permite colocar muchas poblaciones de moléculas de ácido nucleico amplificado diferentes y sus complementos en una sola superficie que tenga un área de superficie relativamente pequeña. La superficie será normalmente plana, aunque esto no es esencial.

55 Aparatos

[0063] La presente invención también proporciona un aparato para proporcionar una superficie que comprende colonias de las moléculas de ácido nucleico inmovilizadas indicadas anteriormente.

60 **[0064]** Dicho aparato puede incluir uno o más de lo siguiente:

a) medios para inmovilizar cebadores en una superficie (aunque esto no es necesario si ya se han proporcionado cebadores inmovilizados);

65 b) un aporte de una polimerasa de ácido nucleico;

c) un aporte de precursores de los nucleótidos a incorporar en un ácido nucleico (por ejemplo, un aporte de nucleósido trifosfatos);

5 d) medios para separar ácidos nucleicos apareados (por ejemplo, medios de calentamiento);
y

e) medios de control para coordinar las diferentes etapas requeridas para el método de la presente invención.

10 **[0065]** Otros aparatos se encuentran dentro del ámbito de la presente descripción. Estos permiten analizar los ácidos nucleicos inmovilizados producidos mediante el método de la presente invención. Esto puede incluir una fuente de reactantes y medios de detección para detectar una señal que pueda generarse una vez que uno o más reactantes se han aplicado a las moléculas de ácido nucleico inmovilizadas. Esto también puede proporcionarse con una superficie que comprenda moléculas de ácido nucleico inmovilizadas en forma de colonias, como se describe anteriormente.

15 **[0066]** A ser posible el medio para detectar una señal tiene suficiente resolución para permitir que diferencie entre señales generadas de diferentes colonias.

20 **[0067]** Los aparatos de la presente descripción (de cualquiera naturaleza) se proporcionan preferentemente en forma automatizada de tal manera que una vez que se han activado, puedan repetirse etapas de procesos individuales automáticamente.

25 **[0068]** La presente invención se describirá ahora sin limitación de la misma en las siguientes secciones A a I con referencia a los dibujos acompañantes.

30 **[0069]** Debe apreciarse que los procedimientos que utilizan las moléculas de ADN referidas en estas secciones son aplicables, con los cambios debidos, a moléculas de ARN, a menos que el contexto indique otra cosa.

[0070] También debe apreciarse que cuando se proporcionan secuencias en la siguiente descripción, estas se escriben en dirección 5' a 3' (de izquierda a derecha), a menos que el contexto indique otra cosa.

[0071] Las figuras proporcionadas se resumen a continuación:

35 La FIGURA 1 ilustra un método de amplificación e inmovilización simultánea de moléculas de ácido nucleico utilizando un solo tipo de cebador.

La FIGURA 2 ilustra cómo puede producirse el crecimiento de colonias utilizando un método de la presente invención.

40 La FIGURA 3 ilustra el principio del método utilizado para producir colonias de ADN utilizando la presente invención. También ilustra las etapas de apareamiento, alargamiento y desnaturalización que se utilizan para proporcionar dichas colonias.

45 La FIGURA 4 es un ejemplo de colonias de ADN formadas por amplificación de un molde específico con cebadores sencillos injertados en una superficie.

La FIGURA 5 ilustra un método de amplificación e inmovilización simultánea de moléculas de ácido nucleico utilizando dos tipos de cebadores.

50 La FIGURA 6 muestra colonias de ADN reales producidas mediante la presente invención.

La FIGURA 7 ilustra un método de amplificación e inmovilización simultánea de moléculas de ácido nucleico cuando se utiliza, como molde, una molécula diana que tiene secuencias internas que se aparean con cebadores.

55 La FIGURA 8 ilustra un método para sintetizar copias adicionales de las cadenas de ácido nucleico originales utilizando cadenas de ácido nucleico presentes en colonias. Las cadenas recién sintetizadas se muestran en solución pero si se desea pueden proporcionarse en forma inmovilizada.

60 La FIGURA 9 muestra la amplificación por PCR de ADN de ADN encontrado en colonias de ADN previamente formadas.

La FIGURA 10 ilustra cómo pueden generarse cebadores secundarios a partir de colonias de ADN.

65 La FIGURA 11 ilustra cómo pueden generarse colonias de ADN secundarias a partir de cebadores secundarios.

La FIGURA 12 ilustra cómo pueden generarse cebadores con diferentes secuencias a partir de una superficie funcionalizada con cebadores existentes.

5 La FIGURA 13 representa métodos de preparación de fragmentos de ADN adecuados para la generación de colonias de ADN.

La FIGURA 14 ilustra un método para sintetizar ARNc utilizando la colonia de ADN como un sustrato para la ARN polimerasa.

10 La FIGURA 15 ilustra un método preferible para determinar la secuencia de ADN de ADN presente en colonias individuales.

La FIGURA 16 ilustra un método de determinación de la secuencia de una colonia de ADN, *de novo*.

15 La FIGURA 17 ilustra la utilidad de colonias de ADN secundarias en el ensayo de niveles de expresión de ARNm.

Las FIGURAS 18 y 19 ilustran el uso de colonias de ADN secundarias en el aislamiento e identificación de genes expresados nuevos y poco frecuentes.

20 **A) Esquema que muestra la amplificación e inmovilización simultánea de moléculas de ácido nucleico utilizando un solo tipo de cebador**

[0072] En relación ahora con la Figura 1a), se proporciona una superficie que tiene unida a la misma una pluralidad de cebadores (para simplificar solo se muestra un cebador). Cada cebador (1) está unido a la superficie mediante un enlace indicado por un bloque de color oscuro. Este enlace puede ser covalente o no covalente pero debe ser suficientemente fuerte como para mantener un cebador en su sitio en la superficie. Se muestran cebadores que tienen una secuencia de nucleótidos corta (5'-ATT). Sin embargo, en la práctica podrían proporcionarse generalmente secuencias más largas.

30 [0073] La Figura 1b) muestra una molécula diana (II) que se ha apareado con un cebador. La molécula diana comprende en su extremo 3' una secuencia (5'-ATT) que es complementaria a la secuencia cebadora (5'-ATT). En su extremo 5' la molécula diana comprende una secuencia (5'-ATT) que es igual que la secuencia cebadora (aunque no se requiere una identidad exacta).

35 [0074] Entre los dos extremos puede proporcionarse cualquier secuencia a amplificar (o la complementaria de cualquier secuencia a amplificar). Como ejemplo, parte de la secuencia a amplificar se muestra como 5'-CCG.

40 [0075] En la Figura 1c) se muestra una extensión de cebador. En este caso se utiliza una ADN polimerasa junto con dATP, dTTP, dGTP y dCTP para extender el cebador (5'-ATT) desde su extremo 3', utilizando como molde la molécula diana.

45 [0076] Cuando la extensión del cebador es completa, como se muestra en la Figura 1d), puede observarse que se proporciona una cadena extendida inmovilizada (III) que es complementaria a la molécula diana. Después, la molécula diana puede separarse de la cadena inmovilizada extendida (por ejemplo, por calentamiento, como se muestra en la Figura 1e)). Esta etapa de separación libera la cadena extendida, inmovilizada, de tal manera que después se pueda utilizar para iniciar una ronda posterior de extensión de cebador, como se muestra en las figuras 1f) y 1g). En este caso, la cadena extendida inmovilizada se dobla de tal manera que un extremo de esta cadena (que tiene la secuencia terminal 5'-AAT) se aparea con otro cebador (2, 5'-ATT), como se muestra en la Figura 1f). Este cebador proporciona un extremo 3' a partir del cual puede producirse la extensión del cebador, esta vez utilizando como molde la cadena inmovilizada, extendida. En la Figura 1g) se muestra que se está produciendo la extensión del cebador y en la Figura 1h) que esta está finalizando.

55 [0077] La Figura 1i) muestra las dos cadenas inmovilizadas extendidas que se mostraron en la Figura 1h) después de su separación (por ejemplo, por calentamiento). Después, cada una de estas cadenas puede en sí misma utilizarse como moldes en rondas posteriores de extensión de cebador iniciada a partir de nuevos cebadores (3 y 4), como se muestra en las Figuras 1j) y 1k). Después de dos rondas de amplificación se pueden proporcionar cuatro cadenas inmovilizadas monocatenarias seguido de una etapa de separación de cadenas (por ejemplo, por calentamiento), como muestra en la Figura 11). Dos de estas tienen secuencias correspondientes a la secuencia de la molécula diana originalmente utilizada como un molde. Las otras dos tienen secuencias complementarias a la secuencia de la molécula diana originalmente utilizada como un molde. (En la práctica una cadena inmovilizada determinada y su complementaria inmovilizada pueden aparearse a la vez).

60 [0078] Se apreciará por tanto que una secuencia determinada y su complementaria pueden proporcionarse en las mismas cantidades en forma inmovilizada y que pueden distribuirse sustancialmente de manera homogénea dentro

de una colonia.

[0079] Por supuesto, se pueden realizar rondas de amplificación posteriores además de las mostradas en la Figura 1, de tal manera que puedan proporcionarse colonias que comprendan grandes cantidades de una molécula de ácido nucleico monocatenario determinada y una cadena complementaria a la misma. Únicamente es necesario utilizar un solo molde para iniciar cada colonia, aunque, si se desea, se puede volver a utilizar un molde para iniciar varias colonias.

[0080] Se apreciará que la presente invención permite proporcionar densidades muy altas de moléculas de ácido nucleico extendidas, inmovilizadas. Dentro de una colonia cada molécula extendida, inmovilizada se localizará en una superficie dentro de una longitud de molécula de otra molécula inmovilizada extendida. Por tanto, la posición 3 mostrada en la Figura 11) está dentro de una longitud de molécula de posición 1; la posición 1 está dentro de una longitud de molécula de posición 2; y la posición 2 está dentro de una longitud de molécula de posición 4.

[0081] La Figura 2 se proporciona para ilustrar cómo puede producirse el crecimiento de colonias (utilizando el método descrito con referencia a la figura 1 y a la figura 6 o a cualquier método de la presente invención para proporcionar moléculas de ácido nucleico inmovilizadas).

[0082] Se muestra una vista esquemática en planta de una placa plana que tiene cebadores inmovilizados sobre la misma en un diseño de rejilla cuadrada (los cebadores se indican mediante puntos pequeños). Se utiliza una rejilla regular exclusivamente por una cuestión de sencillez: en muchos casos reales, las posiciones de los cebadores deberían de hecho estar menos ordenadas o al azar.

[0083] En la posición indicada con una flecha X una molécula molde se ha apareado con un cebador y se ha producido un trozo inicial de extensión de cebador para proporcionar una cadena de ácido nucleico extendida, inmovilizada. Después de la separación de las cadenas, un extremo de esta cadena se queda libre para aparearse con cebadores adicionales de tal manera que pueden producirse otras cadenas de ácido nucleico extendidas, inmovilizadas. Se muestra que esto se ha producido secuencialmente en las posiciones indicadas por la letra Y. Por una cuestión de sencillez, el cebador seleccionado para el apareamiento está ubicado cerca del cebador que lleva la cadena de ácido nucleico: en casos reales, la cadena de ácido nucleico se aparearía con un cebador que no es el vecino más cercano. Sin embargo, es obvio que este cebador estará a una distancia igual a la longitud de la cadena de ácido nucleico.

[0084] Se apreciará que, para que se produzca el crecimiento celular de la colonia, solo se requiere el apareamiento en una de estas posiciones (en lugar de en todas).

[0085] Después de haberse proporcionado otras moléculas de ácido nucleico monocatenario, extendidas, inmovilizadas, en las posiciones indicadas con la letra Y, las moléculas resultantes pueden de por sí aparearse con otros cebadores y el proceso puede continuar para proporcionar una colonia que comprenda una gran cantidad de moléculas de ácido nucleico inmovilizadas en un área relativamente pequeña.

[0086] La Figura 3 muestra una versión simplificada del ciclo de apareamiento, alargamiento y desnaturalización. También representa las observaciones típicas que se pueden hacer, como se puede observar en los ejemplos mostrados en las figuras 4 y 6. La amplificación e inmovilización simultánea de los ácidos nucleicos utilizando cebadores en fase sólida se ha realizado satisfactoriamente utilizando el procedimiento descrito en los Ejemplos 1, 2 y 3 más adelante:

EJEMPLO 1

[0087] Se injertaron oligonucleótidos, fosforilados en su extremo 5' (Microsynth GmbH, Suiza), en pocillos de microtitulación de plástico Nucleolink (Nunc, Roskilde, Dinamarca). La secuencia del oligonucleótido p57 corresponde a la secuencia 5'-TTTTTTCACCAACCCAAACCAACCCAAACC y la de p58 corresponde a la secuencia 5'-TTTTTTAGAAGGAGAAGGAAAGGGAAAGGG. Los pocillos de microtitulación con p57 o p58 se prepararon de la siguiente manera. A cada pocillo Nucleolink, se añadieron 30 µl de una solución del oligonucleótido 160nM en 1-metil-imidazol 10 mM (pH 7,0) (Sigma Chemicals, St.Louis, MO). A cada pocillo, se añadieron 10 µl de 1-etil-3-(3-dimetilaminopropil)-carbodiimida 40 mM (pH 7,0) (Sigma Chemicals) en 1-metil-imidazol 10 mM a la solución de oligonucleótidos. Después, los pocillos se cerraron y se incubaron a 50 °C durante una noche. Después de la incubación; los pocillos se lavaron dos veces con 200 µl de RS (NaOH 0,4 N, Tween 20 al 0,25 % (Fluka Chemicals, Suiza)), se incubaron durante 15 minutos con RS 200 µl RS, se lavaron dos veces con RS 200 µl y dos veces con TNT (TrisHCl 100 mM, pH 7,5, NaCl 150 mM, Tween 20 al 0,1 %) 200 µl. Los tubos se secaron a 50 °C y se guardaron en una bolsa de plástico cerrada a una temperatura de 4 °C.

[0088] La generación de las colonias se inició en cada pocillo con 15 µl de mezcla de cebado; 1 nanogramo de ADN molde (en el que el ADN molde comienza con la secuencia 5'-AGAAGGAGAAGGAAAGGGAAAGGG y finaliza

en el extremo 3' con la secuencia CCCTTCCCTTTCCTTCTCCTTCT-3'), los cuatro dNTP (0,2 mM), BSA (albúmina de suero bovino, Boehringer-Mannheim, Alemania) al 0,1 %, Tween 20 al 0,1 %, DMSO (dimetilsulfóxido, Fluka Chemicals, Suiza) al 8 %, tampón para PCR 1X Amplitaq y 0,025 unidades/μl de ADN polimerasa AmpliAq (Perkin Elmer, Foster City, CA). La reacción de cebado fue una sola ronda de PCR en las siguientes condiciones: 94 °C durante 4 minutos, 60 °C durante 30 segundos y 72 °C durante 45 segundos en un termociclador (PTC 200, MJ Research, Watertown, MA). Después se utilizó tampón TE (trisHCl 10 mM, pH 7,5, EDTA 1 mM) 100 μl en tres lavados sucesivos de un minuto de duración a 94 °C. Después, las colonias de ADN se formaron añadiendo a cada pocillo, 20 μl de mezcla de polimerización, que era idéntica a la mezcla de cebado pero que carecía de ADN molde. Después, los pocillos se colocaron en el termociclador PTC 200 y el crecimiento de las colonias se realizó mediante incubación de los pocillos cerrados durante 4 minutos a 94 °C y ciclado durante 50 repeticiones a las siguientes condiciones: 94 °C durante 45 segundos, 65 °C durante 2 minutos, 72 °C durante 45 segundos. Después de finalizar este programa, los pocillos se mantuvieron a 8° C hasta su uso posterior.

[0089] Se amplificó por PCR un fragmento de 640 pares de bases correspondiente a la secuencia central del molde (pero que no incluía la secuencia 5'-AGAAGGAGAAGGAAAGGGAAAGGG). PCR. El fragmento aislado se marcó con biotina-N⁴-dCTP (NEN Life Sciences, Boston, MA) y una traza de [α -³²P]dCTP (Amersham, UK) utilizando el kit de marcaje Primet-it II (Stratagene, San Diego, CA) para generar una sonda marcada con biotina.

[0090] La sonda marcada con biotina se diluyó en una concentración de 2,5 nM en EasyHyb (Boehringer-Mannheim, Alemania) y 15 μl se hibridaron en cada muestra con el siguiente esquema de temperatura (termociclador PTC 200): 94 °C durante 5 minutos, seguido de 500 etapas de disminución de 0,1 °C de temperatura cada 12 segundos (en otras palabras, la temperatura e redujo a 45 °C en 100 minutos). Después, las muestras se lavaron de la siguiente manera; 1 vez con 2X SSC/SDS al 0,1 % (2X SSC; Na-Cl 0,3 M/citrato de sodio 0,03 M pH 7,0/dodecil sulfato sódico 0,001 mg/ml) a temperatura ambiente, una vez con 2X SSC/SDS al 0,1 % a 37 °C y una vez con 0,2X SSC/SDS al 0,1 % a 50 °C. Después, los pocillos se incubaron durante 30 minutos con 50 μl de fluorescente rojo, se cubrieron con Neutravidina, FluoSpheres^(R) 40 nm (excitación a 580 nm y emisión a 605 nm, Molecular Probes Inc., Eugene, OR) en TNT/BSA al 0,1 %. (La solución de microesferas se preparó a partir de una dilución de 2 μl de la solución de reserva de microesferas en 1 ml de TNT/BSA al 0,1 %, que después se sometió a ultrasonidos durante 5 minutos en un baño de agua ultrasonido a 50 W (Elgasonic, Suiza), seguido por filtración a través de un filtro de 0,22 μm (Millex GV4). Después, los pocillos se contaron (Cherenkov) en un contador de centelleo de placa Microbeta (WALLAC, Turku, Finlandia).

[0091] Las FluoSpheres^(R) sobrantes se retiraron mediante lavado durante 30 min en TNT/BSA al 0,1 % a temperatura ambiente. Se observaron imágenes de las muestras teñidas utilizando un objetivo de 20X en un microscopio invertido (Axiovert S100TV, Carl Zeiss AG, Oberkochen, Alemania) equipado con una cámara CCD Micromax512x768 (Princeton instruments, Trenton, NJ) a través de un conjunto de filtros XF43 (PB546/FT580/LP590, Omega Optical, Brattleboro, VT) con una exposición de 5 segundos.

[0092] La Figura 4 muestra los resultados de hibridación de la generación de colonias en tubos funcionalizados con cualquiera de: (a) el oligonucleótido p57 o (b) el oligonucleótido p58. La reacción de control muestra manchas muy poco fluorescentes, ya que la secuencia de las regiones flanqueantes en el molde no corresponden a las secuencias cebadoras injertadas sobre el pocillo. En cambio, la figura 4b muestra el número de manchas fluorescentes detectado cuando los cebadores injertados en los pocillos se corresponden con las secuencias flanqueantes en el molde iniciador de ADN. Calculando el número de manchas fluorescentes detectado y teniendo en cuenta el aumento, se puede estimar que hay entre 3 y 5 x 10⁷ colonias/cm². Las fotografías se generaron con el programa Winview 1.6.2 (Princeton Instruments, Trenton, NJ) con fondos e intensidades normalizados a los mismos valores.

B) Esquema que muestra la amplificación e inmovilización simultánea de moléculas de ácido nucleico utilizando dos tipos diferentes de cebadores

[0093] Referente ahora a la Figura 5, se ilustra otra realización de la presente invención. En esta realización, se utilizaron dos cebadores diferentes inmovilizados para proporcionar la extensión del cebador.

[0094] En esta realización la molécula diana mostrada se proporciona con una secuencia de nucleótidos en su extremo 3' (AAT-3') que es complementaria a la secuencia de un primer cebador, (5'-ATT, I), que se injerta en la superficie, de manera que puede producirse el apareamiento con ese cebador. La secuencia (5'-GGT) en el extremo 5' de la molécula diana, III, corresponde a la secuencia (5'-GGT) de un segundo cebador, II, que también se injerta en la superficie, de tal manera que la secuencia que es complementaria a la secuencia en el extremo 5' puede aparearse con este dicho segundo cebador. Generalmente dicha secuencia complementaria (5'-ACC) se selecciona de tal manera que no se aparee con el primer cebador (5'-ATT). A pesar de la situación descrita en el apartado A, una vez que el extremo 3' de una cadena recién sintetizada se aparee con un cebador en la superficie, tendrá que encontrar un cebador cuya secuencia sea diferente de la de la secuencia que lleva en su extremo 5' (véase la diferente entre las figuras 1f y 5f).

[0095] La realización mostrada en la figura 5 tiene una ventaja sobre la realización ilustrada en la figura 1 ya que la posibilidad de que un extremo de una molécula diana monocatenaria se aparee con otro extremo de la misma molécula en solución puede evitarse y por lo tanto la amplificación puede seguir adelante. La posibilidad de que se produzca el apareamiento entre ambos extremos de un complemento inmovilizado a una molécula diana también puede evitarse.

EJEMPLO 2

[0096] Una mezcla de dos oligonucleótidos que estaban fosforilados en el extremo 5' ((Microsynth GmbH, Balgach, Suiza) se injertó en placas Nucleolink de 96 pocillos (Nunc, Dinamarca) según recomendaciones del fabricante. Las placas resultantes se conservaron en seco a una temperatura de 4 °C. La secuencia del cebador, P1, era 5'-GCGCGTAATACGACTCACTA, la secuencia del otro cebador, P2, era 5'-CGCAATTAACCCTCACTAAA. Estas placas se formularon especialmente en Nunc, y permiten el injerto covalente de fragmentos de ADN fosforilados en el extremo 5' a través de un procedimiento estándar.

[0097] Se clonó un molde en un vector (pBlueScript Skminus, Stratagene Inc, San Diego, CA) con la secuencia de ADN apropiada en el sitio de clonación (es decir, correspondiente a P1 y P2 en la posición 621 y 794 respectivamente) y se obtuvo un molde de ADN bicatenario lineal de 174 pb de longitud mediante amplificación por PCR, utilizando P1 y P2. El producto PCR molde se purificó en columnas Qiagen Qia-quick (Qiagen GmbH, Hilden, Alemania) para retirar los nucleótidos y los cebadores utilizados durante la amplificación PCR.

[0098] El molde purificado (en una solución de 50 µl que contenía tampón 1X PCR (Perkin Elmer, Foster City, CA) con los cuatro desoxirribonucleósido trifosfatos (dNTP) a 0,2 mM (Pharmacia, Uppsala, Suecia) y 2,5 unidades de ADN polimerasa AmpliTaq Gold (Perkin Elmer, Foster City, CA)) se dispersó en el soporte, es decir, sobre las placas Nucleolink injertadas con P1 y P2 (las placas se habían lavado con una solución que contenía TRIS-HCl 100 mM (pH 7,5), NaCl 150 mM y Tween 20 al 0,1 % (Fluka, Suiza) a temperatura ambiente durante 15 min). Esta solución se incubó a 93 °C durante 9 minutos para activar la ADN polimerasa y después se realizaron 60 ciclos (94 °C/30 s; 48 °C/ 30 s; 72 °C/30 s) en un termociclador PTC 200. Se analizaron diversas concentraciones diferentes de molde PCR (aproximadamente 1, 0,5, 0,25, 0,125, 0,0625 ng/µl) y para cada muestra se realizó una reacción de control efectuada sin Taq polimerasa (las mismas condiciones que las anteriores pero sin ADN polimerasa).

[0099] Cada muestra se tiñó con YO-PRO (Molecular Probes, Portland OR), un tinte altamente sensible para el ADN bicatenario. Los productos resultantes se observaron con un microscopio confocal utilizando un objetivo de 40X (LSM 410, Carl Zeiss AG, Oberkochen, Alemania) con excitación apropiada (un láser de argón 488) y filtros de detección (un filtro de paso bajo 510) (nota: el fondo de cada pocillo es plano y permite la observación con un microscopio de fluorescencia invertido).

[0100] En la Figura 6A, el pocillo de control (sin molde de ADN añadido, panel a) muestra únicamente objetos extraños que pueden observarse en una superficie de color blanco (estos objetos eran útiles en esta fase para indicar que el foco era correcto). Estos objetos tenían una forma irregular, un tamaño de 20 a 100 micrómetros y tenían un grosor mucho más grande que la profundidad del campo de observación. En un pocillo donde había ADN polimerasa (figura 6A, panel ii), además de los objetos de forma irregular observados en el pocillo control, se observó una gran cantidad de manchas fluorescentes. Estas presentaban una forma circular, tenían un tamaño de 1 a 5 micrómetros y no ocupaban el campo visual. El número de manchas dependía de la concentración del molde utilizado para iniciar la formación de colonias. A partir del tamaño observado de las colonias, se puede estimar que pueden disponerse más de 10.000 colonias distintas dentro de 1 mm² de soporte.

EJEMPLO 3

[0101] Se injertaron oligonucleótidos (Microsynth GmbH, Suiza) en pocillos Nucleolink (Nunc, Dinamarca). El oligonucleótido P1 corresponde a la secuencia 5'-TTTTTCTCACTATAGGGCGAATTGG y el oligonucleótido P2 corresponde a 5'-TTTTTCTCACTAAAGGGAACAAAAGCTGG. A cada pocillo Nucleolink, se añadió una solución de 45 µl de 1-metil-imidazol 10 mM (pH 7,0) (Sigma Chemicals, St.Louis, MO) que contenía 360 fmol de P1 y 360 fmol de P2. A cada pocillo, a la solución de oligonucleótidos, se añadieron 15 µl de 1-etil-3-(3-dimetilaminopropil)-carbodiimida 40 nM (pH 7,0) (Sigma Chemicals) en 1-metil-imidazol. Después, los pocillos se cerraron y se incubaron a 50 °C durante 16 horas.

[0102] Después de la incubación, los pocillos se aclararon dos veces con RS 200 µl (NaOH 0,4 N, Tween 20 al 0,25 %), se incubaron durante 15 minutos con RS 200 µl, se lavaron dos veces con RS 200 µl, y dos veces con TNT (Tris/HCl 100 mM, pH 7,5, NaCl 150 mM, Tween 20 al 0,1 %) 200 µl, antes de ponerlos a secar a 50 °C en un horno. Los tubos secos se guardaron en una bolsa de plástico cerrada a una temperatura de 4 °C.

[0103] El crecimiento de las colonias se inició en cada pocillo con 15 µl de mezcla de inicio (tampón PCR 1X, los dNTP 0,2 mM y 0,75 unidades de ADN polimerasa AmpliTaq Gold, 20 nanogramos de ADN molde, en el que el

molde de ADN era ADN S1 o ADN S2 o una mezcla de diferentes proporciones de ADN S1 y ADN S2, como se indica en el comentario de la figura 6B. S1 y S2 son fragmentos de 704 y 659 pares de bases, respectivamente, que se han clonado en los plásmidos pBlue-Script Skminus y posteriormente amplificados a través de una PCR utilizando P1 y P2 como cebadores. Los fragmentos se purificaron en columnas Qia-quick de Qiagen (QIAGEN GmbH, Alemania) para retirar los nucleótidos y los cebadores.

10 [0104] Cada pocillo se cerró con Cycleseal™ (Robbins Scientific Corp., Sunnyvale, CA), y se incubó a 93 °C durante 9 minutos, a 65 °C durante 5 minutos y a 72 °C durante 2 minutos y de nuevo a 93 °C. Después, se utilizó una solución de TNT 200 µl en tres lavados sucesivos un minuto de duración a 93 °C. La mezcla de inicio se reemplazó después por mezcla de crecimiento 15 µl (la misma que la mezcla de inicio, pero sin ADN molde), y el crecimiento se realizó incubando los pocillos sellados durante 9 minutos a 93 °C y repitiendo 40 veces las siguientes condiciones: 93 °C durante 45 segundos, 65 °C durante 3 minutos, 72 °C durante 2 minutos. Después de finalizar este programa, los pocillos se conservaron a 6 °C hasta su utilización. El control de la temperatura se realizó en un termociclador PTC 200, utilizando una almohadilla de silicio proporcionada en el kit Nucleolink y la placa se calentó 15 (104 °C) en el termociclador PTC 200.

20 [0105] Por PCR se amplificó un fragmento de 640 pares de bases correspondiente a la secuencia central del fragmento S1, pero que no incluía la secuencia P1 o P2, como se ha descrito anteriormente. La sonda se marcó con biotina-16-dUTP (Boehringer-Mannheim, Alemania) utilizando el kit de marcaje de cebadores al azar Prime-it II (Stratagene, San Diego, CA) de acuerdo con las instrucciones del fabricante.

25 [0106] Las sondas marcadas con biotina se hibridaron con las muestras en tampón EasyHyb (Boehringer-Mannheim, Alemania), utilizando el siguiente esquema de temperatura (en el termociclador PTC 20): 94 °C durante 5 minutos, seguido de 68 etapas de disminución de 0,5 °C de temperatura cada 30 segundos (en otras palabras, la temperatura disminuyó a 60 °C en 34 minutos), utilizando pocillos cerrados. Después, las muestras se lavaron 3 veces con 200 µl de TNT a temperatura ambiente. A continuación, los pocillos se incubaron durante 30 minutos con TNT 50 µl que contenía BSA 0,1 mg/ml. Después, los pocillos se incubaron durante 5 minutos con 15 µl de solución de fluorescente rojo, se cubrieron con Neutravidina, FluoSpheres^(R) 40 nm (excitación a 580nm y emisión a 605 nm, Molecular Probes, Portland, OR). La solución de microesferas se preparó a partir de una dilución de 2 µl de la solución de reserva de microesferas, que se había sometido a ultrasonido durante 5 minutos en un baño de agua ultrasonido a 50 W (Elgasonic, Suiza), diluido en 1 ml de solución TNT que contenía BSA 0,1 mg/ml y se filtró con un filtro de tamaño de poro de 0,22 µm Millex GV4 (Millipore, Bedford, MA).

35 [0107] Las muestras teñidas se observaron utilizando un microscopio Axiovert 10 invertido utilizando un objetivo de 20X (Carl Zeiss AG, Oberkochen, Alemania) equipado con una cámara Micromax 512x768 CCD (Princeton Instruments, Trenton, NJ), utilizando un conjunto de filtros XF43 (PB546/FT580/LP590, Omega Optical, Brattleboro, VT), y 10 segundos de recogida de luz. Los archivos se convirtieron a formato TIFF y se procesaron en el programa informático adecuado (PhotoPaint, Corel Corp., Ottawa, Canadá). El procesamiento consistía en una potenciación de inversión y contraste lineal, para proporcionar una imagen adecuada para una impresión en blanco y negro en 40 una impresora láser.

[0108] La figura 6B muestra los resultados de 3 proporciones diferentes de los moldes S1/S2 utilizados en la reacción de inicio: i) la proporción de S1/S2 es de 1/0, pueden observarse muchas manchas, ii) la proporción de S1/S2 es de 1/10, y el número de manchas es de aproximadamente 1/10 del número de manchas que se puede observar en la imagen i), como se esperaba y iii) la proporción S1/S2 es de 0/1, y solo pueden observarse algunas manchas poco frecuentes.

50 **Esquema que muestra la amplificación e inmovilización simultánea de moléculas de ácido nucleico cuando la molécula diana contiene secuencias internas complementarias a los cebadores inmovilizados**

[0109] La Figura 7 se proporciona para mostrar que no es necesario que las secuencias mostradas en los extremos 5' y 3' de la molécula diana ilustrada en las Figuras 1 y 5 se localicen en los extremos de una molécula diana.

55 [0110] Una molécula de ácido nucleico diana (II) puede tener una secuencia en cada (o en cualquier) extremo que no esté implicada en el apareamiento con un cebador ni actúe como molde para proporcionar una secuencia complementaria que se aparee con un cebador (secuencia 5'-AAA y secuencia 5'-CCC). Una de las secuencias internas (5'-AAT) se utiliza como molde para sintetizar una secuencia complementaria, III, a la misma (5'-TTT) como se desprende de las figuras 7(a) a 7(e).

60 [0111] Sin embargo, la propia secuencia 5'-TTT no se utiliza para proporcionar una secuencia complementaria a la misma, como se desprende de las figuras 7(f) a 7(k). En la Figura 7(l) se puede observar que solo una de las cuatro cadenas inmovilizadas se muestran después de dos rondas de extensión con cebadores y que una etapa de separación de cadenas comprende la secuencia adicional 5'-TTT y que no hay ninguna cadena que comprenda una

secuencia complementaria (5'-AAA) a esta secuencia (es decir, solo hay una cadena que es significativamente más larga que las otras). Después de varias rondas de amplificación la cadena que comprende la secuencia 5'-TTT representará una proporción insignificativa del número total de moléculas de ácido nucleico extendidas, inmovilizadas presentes.

5

D) El uso de cadenas de ácido nucleico presentes en colonias para sintetizar copias adicionales de cadenas de ácido nucleico

10 **[0112]** Las moléculas de ácido nucleico monocatenario, amplificadas, presentes en colonias proporcionadas por la presente invención se pueden en sí mismas utilizarse como moldes para sintetizar cadenas de ácido nucleico adicionales.

15 **[0113]** La Figura 8 ilustra un método para sintetizar ácidos nucleicos adicionales utilizando ácidos nucleicos inmovilizados como un punto de partida.

15 **[0114]** Las colonias normalmente comprenderán tanto una cadena de ácido nucleico determinada como su complementaria en forma inmovilizada (figura 8a). Por tanto estas se pueden utilizar para proporcionar copias adicionales no solo en una cadena de ácido nucleico determinada sino también de su complementaria.

20 **[0115]** Una manera de hacer esto es proporcionar uno o más cebadores (cebadores TTA y TGG) en solución que se aparean con cadenas de ácido nucleico inmovilizadas, amplificadas, presentes en colonias (figura 8c) proporcionadas por la presente invención. (Estos cebadores pueden ser los mismos que los cebadores inicialmente utilizados para proporcionar las colonias inmovilizadas, proporcionándose separados, en forma libre, en lugar de en forma inmovilizada). La colonia de ADN original está desnaturalizada por calentamiento en su forma monocatenaria (figura 8b), lo que permite que los cebadores TTA y TGG se apareen en el extremo 3' disponible de cada cadena de ADN. Después, la extensión con cebadores, utilizando ADN polimerasa AmpliTaq y los cuatro desoxirribonucleósido trifosfatos (marcados o no) se puede utilizar para sintetizar cadenas complementarias con cadenas de ácido nucleico inmovilizadas o al menos con partes de las mismas (etapa (iii)).

25 **[0116]** Una vez sintetizadas las cadenas recién formadas (figura 8d) mediante el proceso descrito anteriormente, estas pueden separarse de las cadenas inmovilizadas con las que están hibridadas (por ejemplo, calentamiento). El proceso puede repetirse si se desea utilizando la reacción PCR, para proporcionar un gran número de dichas cadenas en solución (figura 8e).

30 **[0117]** Si se desea, las cadenas sintetizadas de esta manera, después de la separación de las cadenas inmovilizadas, pueden aparearse entre sí (es decir una cadena determinada y su complementaria pueden aparearse hibridarse) para proporcionar moléculas de ácido nucleico bicatenarias en solución. Como alternativa, pueden separarse entre sí para proporcionar poblaciones homogéneas de moléculas de ácido nucleico monocatenario en solución.

40 **[0118]** También ha de observarse que, una vez proporcionadas las moléculas monocatenarias en solución, estas se pueden utilizar como moldes para la PCR (o PCR inversa). Por lo tanto no es esencial continuar utilizando las cadenas de ácido nucleico inmovilizadas para obtener una amplificación adicional de cadenas determinadas o cadenas complementarias a las mismas.

45 **[0119]** Debe observarse que, cuando se proporciona una pluralidad de colonias y que las cadenas de ácido nucleico en diferentes colonias tienen diferentes secuencias, es posible seleccionar solo determinadas colonias para su uso como moldes en la síntesis de moléculas de ácido nucleico adicionales. Esto se puede realizar utilizando cebadores para la extensión de cebadores que son específicos para moléculas presentes en colonias seleccionadas.

50

[0120] Como alternativa pueden proporcionarse cebadores para permitir utilizar varias o todas las colonias como moldes. Dichos cebadores pueden ser una mezcla de muchos cebadores diferentes (por ejemplo, una mezcla de todos los cebadores originalmente utilizados para proporcionar todas las colonias, pero proporcionándose los cebadores en solución en lugar de en forma inmovilizada).

55

EJEMPLO 4

60 **[0121]** Se injertaron oligonucleótidos Microsynth GmbH Balgach, Suiza) en pocillos Nucleolink (Nunc, Dinamarca). El oligonucleótido P1 corresponde a la secuencia 5'-TTTTTTTTTACCAACCCAAACCAACCCAAACC y el oligonucleótido P2 corresponde a la 5'-TTTTTTTTTAGAAGGAGAAGGAAAGGGAAAGGG. En cada pocillo Nucleolink, se añadió una solución de 45 µl de 1-metil-imidazol 10 mM (pH 7,0) (Sigma Chemicals) que contenía 360 fmol de P1 y 360 fmol de P2. A cada pocillo, se añadieron 15 µl de 1-etil-3-(3-dimetilaminopropil)-carbodiimida 40 mM (pH 7,0) (Sigma Chemicals) en 1-metil-imidazol 10 mM a la solución de oligonucleótidos. Después los pocillos se cerraron y se incubaron a 50 °C durante 16 horas. Después de la incubación los pocillos se aclararon dos veces con

65 200 µl de RS (NaOH 0,4 N, Tween 20 al 0,25 %), se incubaron durante 15 minutos con RS 200 µl, se lavaron dos

veces con RS 200 μ l, y dos veces con TNT (Tris 100 mM/HCl, pH 7,5, NaCl 150 mM, Tween 20 al 0,1 %) 200 μ l, antes de ponerlos a secar en un horno a 50 °C. Los tubos secos se guardaron en una bolsa de plástico cerrada a una temperatura de 4 °C.

5 **[0122]** El crecimiento de las colonias se inició en cada pocillo con 15 μ l de mezcla de inicio (tampón PCR 1X, los dNTP 0,2 mM y 0,75 unidades de ADN polimerasa AmpliTaq, 20 nanogramos de ADN molde, donde el ADN molde era ADN S1 o ADN S2 o una mezcla 1/1 de ADN S1 y ADN S2, como se indica en el comentario del Ejemplo 3. S1 y S2 son fragmentos de 658 y 704 pares de bases, respectivamente, que se han preparado como se describe en el Ejemplo 3.

10

[0123] Cada pocillo se selló con Cycleseal™ (Robbins Scientific Corp., Sunnyvale, CA), y se incubó a 93 °C durante 9 minutos, a 65 °C durante 5 minutos y a 72 °C durante 2 minutos y de nuevo a 93 °C. Después, se utilizó una solución de TNT 200 μ l en tres lavados sucesivos de un minuto de duración a 93 °C. La mezcla de inicio se reemplazó después por 15 μ l de mezcla de crecimiento (la misma que la mezcla de inicio, pero sin ADN molde), y el crecimiento se realizó incubando los pocillos sellados durante 9 minutos a 93 °C y repitiendo 40 veces las siguientes condiciones: 93 °C durante 45 segundos, 65 °C durante 3 minutos, 72 °C durante 2 minutos. Después de finalizar este programa, los pocillos se conservaron a 6 °C hasta su utilización posterior. El control de la temperatura se realizó en un termociclador PTC 200.

20 **[0124]** Se aplicaron diferentes tratamientos a 6 conjuntos (A,B,C,D,E y F) de 3 pocillos (1,2,3), uno preparado con molde S1, uno con molde S1 y molde S2 y uno preparado solo con molde S2 (lo que produce A1,A2,A3,..., F1,F2,F3). El conjunto A no se trató, el conjunto B se incubó durante 10 minutos con exonucleasa BAL-31 (New England Biolabs, Beverly, MA) a 37 °C en tampón BAL-31 (BAL-31 digiere esencialmente ADN bicatenario que tiene ambos extremos libres), el conjunto C se incubó durante 10 minutos con nucleasa S1 (Pharmacia Uppsala, Suecia) a 37 °C en tampón S1 (la nucleasa S1 digiere esencialmente ADN monocatenario), el conjunto D, E y F se incubaron tanto con BAL-31 como con nucleasa S1. Las reacciones se detuvieron lavando los pocillos con tampón TNT.

[0125] La PCR (25 ciclos, 30 s a 94 °C, 45 s a 60 °C, 45 s a 72 °C) se realizó en los pocillos Nucleolink con los cebadores 0,25 μ M P70 (5'-CACCAACCCAAACCAACCCAAACCCACGACTCACTATAGGGCGAA) y P71 (5'-AGAAGGAGAAGGAAAGGGTAAAGGGGAACAAAAGCTGGA) en solución en los conjuntos A, B, C y D. P70 y P71 eran idóneos para la amplificación de S1 y S2, ya que el cebador P70 contiene la secuencia del cebador P1 y P71 contiene la de P2. En los pocillos del conjunto E, se realizó la PCR con un conjunto de cebadores directo (P150, 5'-GGTGCTGGTCTCAGTCTGT) e inverso (P151, 5'-CCCGCTTACCAGTTTCCATT) que están dentro de S1 y no dentro de S2 para producir un producto PCR de 321 pb, y en los pocillos del conjunto F, se realizó PCR con un conjunto de cebadores directo (P152, 5'-CTGGCCTTATCCCTAACAGC) e inverso (P153, 5'-CGATCTTGGCTCATCACAAT) que están dentro de S2 y no dentro de S1 para producir un producto PCR de 390 pb. Para cada una de las 18 reacciones de PCR, se utilizaron 3 μ l de solución para electroforesis en gel en agarosa al 1 % en presencia de bromuro de etidio 0,1 μ g/ml. En la Figura 9 se muestran los dibujos de los geles que muestran que, en las colonias, el ADN está protegido de la digestión por exonucleasas (conjuntos B, C y D en comparación con el conjunto A) y que tanto S1 como S2 pueden recuperarse de manera simultánea utilizando P1 y P2 (conjuntos A, B, C y D) o específica (conjunto E y F). En el conjunto E y F, donde los productos más cortos de la PCR se amplificaban de un modo más eficaz que los productos más largos de la PCR en los conjuntos A, B, C, D, pudo detectarse una contaminación cruzada entre los moldes S1 y S2 (véanse los carriles E2 y F1).

45 **E) Suministro de colonias secundarias**

[0126] También es posible modificar colonias inicialmente formadas para proporcionar colonias diferentes (es decir, proporcionar colonias que comprendan moléculas de ácido nucleico inmovilizadas con diferentes secuencias de aquellas moléculas presentes en las colonias inicialmente formadas). En este caso, las colonias inicialmente formadas se denominan "colonias primarias" y las últimas colonias formadas se denominan "colonias secundarias". Para convertir las colonias primarias en "cebadores secundarios" es necesario un procedimiento preliminar que será adecuado para la generación de colonias secundarias.

[0127] La Figura 10 muestra cómo se generan 'cebadores secundarios' utilizando colonias primarias existentes. Como un punto de partida, la colonia primaria (figura 10a) se deja en forma bicatenaria, hibridada por completo. Se podría utilizar una ADN exonucleasa específica monocatenaria para retirar todos los cebadores que no se hayan alargado. También se podría optar por proteger todos los extremos 3'-OH de los cebadores con didesoxirribunucleótido trifosfatos utilizando una ADN terminal transferasa (etapa (i), figura 10b).

60 **[0128]** En segundo lugar e independientemente, las moléculas de ADN que forman las colonias pueden escindirse utilizando endonucleasas. Por ejemplo, una enzima de restricción que reconozca un sitio específico dentro de la colonia (representado por la flecha 'RE' en la figura 10c) y escinda la colonia de ADN (etapa (ii), figura 10). Si se desea, la colonia escindida con enzimas (figura 10d) puede después digerirse parcialmente con una exonucleasa II específica bicatenaria en dirección 3' a 5' (por ejemplo, exonucleasa de E. coli III, representada por 'N', etapa (iii),

figura inferior). En cualquier caso, los cebadores secundarios se encuentran disponibles después de la desnaturalización (por ejemplo, por calor) y lavado (figura 10e).

5 **[0129]** Como alternativa, el ADN bicatenario que forma las colonias (figura 10f) puede digerirse con la exonucleasa 3'-5' específica de cadena doble, que solo digiere una cadena del ADN bicatenario. Un caso importante es cuando la exonucleasa digiere solo algunas bases de la molécula de ADN antes de liberarse en solución y cuando la digestión se puede realizar cuando otra enzima se une a la molécula de ADN (figura 10g). En este caso la digestión con exonucleasa continuará hasta que solo queden moléculas monocatenarias que, por término medio, tienen la mitad de la longitud del material de partida, y no tienen partes complementarias (que podrían formar dúplex parciales) que
10 quedan en las moléculas de cadena sencilla en una colonia (figura 10h).

[0130] En todos los casos, estos tratamientos dan lugar a fragmentos monocatenarios injertados sobre un soporte que corresponde a la secuencia del molde original y que se pueden utilizar para el crecimiento de nuevas colonias de ADN si se proporciona un molde nuevo apropiado para el inicio de la colonia (figuras 10e y 10h).

15 **[0131]** El resultado de dicho tratamiento, por tanto un soporte de sujeción de cebadores secundarios, se denominará "soporte para el crecimiento de colonias secundarias". Los moldes útiles para el crecimiento de colonias secundarias pueden incluir moléculas que tengan secuencias conocidas (o complementarias de dichas secuencias). Como alternativa los moldes pueden proceder de moléculas no secuenciadas (por ejemplo, fragmentos al azar). En
20 ningún caso los moldes deben proporcionarse con una o más regiones para el apareamiento con cadenas de ácido nucleico presentes en las colonias primarias.

[0132] La Figura 11 a-e muestra cómo puede generarse una colonia secundaria cuando se proporciona un molde (TP, Figura 11a) adecuado para una segunda ronda de generación de colonias de ADN sobre un soporte para el
25 crecimiento de colonias secundarias, que sujeta cebadores secundarios. En este ejemplo, el tratamiento de la colonia primaria, como se describe anteriormente, ha generado los cebadores secundarios, SP1 y SP2 (figura 11a). El molde TP, se hibridará con su cebador secundario complementario, SP1, y después de una reacción en extensión utilizando una ADN polimerasa, como se describe, extenderá como se representa (figura 12b). Después de una desnaturalización (etapa ii), reapareamiento (etapa iii) y de un ciclo de ADN polimerasa (etapa iv), se formará una
30 copia de la colonia primaria original (figura 11e).

[0133] El tamaño máximo de una colonia secundaria proporcionada por esta realización de la presente invención está limitado por el tamaño de la colonia primaria sobre la cual se desarrolla. Se pueden utilizar diversos procesos de crecimiento secundarios secuencialmente para crear colonias para aplicaciones específicas (es decir, una
35 primera colonia puede reemplazarse por una segunda colonia, la segunda colonia puede reemplazarse por una tercera colonia, etc.).

F) Suministro de cebadores extendidos

40 **[0134]** La Figura 12 muestra cómo se pueden generar cebadores extendidos en una matriz de oligonucleótidos. Se puede aplicar el mismo procedimiento a un soporte cubierto con colonias o cebadores secundarios como se describe en el apartado E.

45 **[0135]** En la Figura 12a) se proporciona un soporte que tiene una pluralidad de cebadores inmovilizados mostrados sobre el mismo. Se muestran diferentes cebadores inmovilizados presentes en diferentes regiones del soporte (representado por cuadrados). Los cebadores que tienen la secuencia 5'-AAA se muestran en un cuadrado y los cebadores que tienen la secuencia 5'-GGG se muestran en otro cuadrado.

50 **[0136]** Las Figuras 12b) a 12d) muestran cómo se modifican los cebadores iniciales presentes (cebadores iniciales) para dar diferentes cebadores (cebadores extendidos). En este ejemplo, los cebadores iniciales que tienen la secuencia 5'-AAA se modifican para producir dos tipos diferentes de cebadores extendidos, que tienen las secuencias 5'-AAAGCC y 5'-AAATAC respectivamente. Esto se consigue a través de la hibridación de moldes oligonucleotídicos 5'-GTATTT y 5'-GGCTTT con los cebadores primarios inmovilizados sobre la superficie (figura 12b), seguido por reacción de ADN polimerasa. Los cebadores iniciales que tienen la secuencia 5'-GGG se
55 modifican para producir dos tipos diferentes de cebadores extendidos, que tienen las secuencias 5'-GGGTAT y 5'-GGGTAA (figura 12d) de una manera similar.

60 **[0137]** La técnica de producir cebadores extendidos es útil para transformar oligonucleótidos inmovilizados proporcionados en una microplaca de ADN o en otra superficie en cebadores inmovilizados útiles en la amplificación de una secuencia de ácido nucleico diana particular y/o amplificando una cadena complementaria a la misma.

G) Preparación de fragmentos de ácido nucleico

65 **[0138]** Los aparatos de la presente invención se pueden utilizar para diversos procedimientos, algunos de los cuales se describirán más adelante. Los fragmentos de ácido nucleico para su utilización en la generación de

colonias se pueden preparar de diferente manera para los procedimientos diferentes (en el presente documento denominados “ácidos nucleicos preparados”). A continuación se describen diversos procedimientos de preparación:

(i) Preparación de fragmentos de ADN al azar

- 5
- [0139]** Aquí se describe un método para preparar ADN que se origina de una muestra (o de una pluralidad de muestras) biológica por amplificación en el caso en el que no sea necesario rastrear el origen del ADN cuando este se incorpora en una colonia.
- 10 **[0140]** El ADN de interés se extrae primero de la muestra biológica y se corta al azar en trozos “pequeños” (por ejemplo, de 50 a 10.000 bases de longitud, pero preferentemente de 500 a 1000 pares de bases de longitud, representado por la barra “I”, figura 13a). (Esto se puede realizar, por ejemplo, mediante una extracción con fenol-cloroformo seguido de tratamiento con ultrasonido, cizalla mecánica, digestión parcial con endonucleasas de restricción cortadoras frecuentes y otros métodos conocidos por los expertos en la técnica). Para normalizar las
- 15 condiciones experimentales, los fragmentos de ADN extraídos y cortados pueden fraccionarse según su tamaño, por ejemplo, mediante electroforesis en gel de agarosa, centrifugación en gradiente de sacarosa o cromatografía en gel. Los fragmentos obtenidos dentro de una sola fracción se pueden utilizar para proporcionar moldes para reducir la variabilidad del tamaño de los moldes.
- 20 **[0141]** En segundo lugar, los fragmentos de ADN molde extraídos, cortados y (opcionalmente) clasificados, pueden estar ligados con enlazadores oligonucleotídicos (IIa y IIb, figura 13a) que contienen la secuencia de uno o más cebadores que previamente se han injertado sobre un soporte. Esto se puede realizar, por ejemplo, utilizando ligamiento del “extremo romo”. Como alternativa, los fragmentos de ADN molde pueden insertarse en un vector biológico en un sitio que está flanqueado por la secuencia de los cebadores que están injertados sobre el soporte.
- 25 Este ADN clonado puede amplificarse dentro de un hospedador biológico y extraerse. Obviamente, si se está trabajando con un solo cebador injertado en el soporte sólido para la formación de colonias de ADN, la purificación de los fragmentos que contienen ambos cebadores P1 y P2 no supone ningún problema.
- [0142]** A continuación, los fragmentos de ADN obtenidos después de dicho proceso adecuado se denominan con la expresión: “ADN genómico preparado” (III, figura 13a).
- 30

(ii) Preparación de fragmentos de ADN al azar que se originan a partir de una pluralidad de muestras

- [0143]** Aquí se describe cómo preparar ADN que se origina de una pluralidad de muestras biológicas en el caso en el que sea necesario rastrear el origen del ADN cuando este se incorpora en una colonia.
- 35
- [0144]** El procedimiento es el mismo que el descrito en el apartado anterior, excepto que en este caso, los enlazadores oligonucleotídicos que se utilizan para encolar los fragmentos de ADN genómicos cortados al azar están constituidos ahora por dos partes; la secuencia de los cebadores injertados sobre la superficie (P1 y P2, figura 13b) y una secuencia “etiqueta” que es diferente para cada muestra y que se utilizará para identificar el origen de la colonia de ADN. Obsérvese que, para cada muestra, la etiqueta puede no ser exclusiva, pero que se puede utilizar una pluralidad de etiquetas. En adelante, los fragmentos de ADN obtenidos después de dicho proceso adecuado se denominarán con la expresión “ADN genómico etiquetado” (III, Figura 13b).
- 40
- [0145]** Este procedimiento de etiquetado se puede utilizar para proporcionar colonias que llevan un medio de identificación que es independiente de la secuencia llevada por el propio molde. Esto también puede ser útil cuando algunas colonias van a recuperarse específicamente (utilizando el procedimiento dado en el apartado D). Esto también podría ser útil en el caso en el que las colonias recuperadas se procesen también, por ejemplo, creando nuevas colonias primarias y si se desea una referencia cruzada entre las colonias originales y las nuevas colonias.
- 45
- 50

(iii) Preparación de fragmentos de ADN correspondientes a una pluralidad de secuencias de ADN que se originan de una muestra

- [0146]** El ADN de interés puede extraerse primero de una muestra biológica mediante cualquier medio conocido por los expertos en la técnica (como se ha mencionado anteriormente). Después, las secuencias específicas de interés pueden amplificarse con PCR (etapa (i), figura 13c) utilizando cebadores de PCR (IIa y IIb) constituidos por dos partes; 1) en el extremo 5', las secuencias corresponden a las secuencias de uno o más oligonucleótidos cebadores que se han injertado sobre una superficie (P1 y P2) y 2) en el extremo 3', a secuencias cebadoras específicas para la secuencia de interés (S1 y S2). En adelante, los fragmentos de ADN obtenidos después de dicho proceso adecuado se denominarán con la expresión: “ADN preparado” (III, figura 13c).
- 55
- 60

(iv) Preparación de una pluralidad de fragmentos de ADN que se originan de una pluralidad de muestras

- [0147]** El procedimiento es el mismo que el indicado en el aparato anterior excepto que en este caso los cebadores de ADN (IIa y IIb) utilizados para realizar la amplificación por PCR (etapa (i), figura 13d) están

constituidos ahora por tres partes; 1) la secuencia de los cebadores injertados sobre la superficie (P1 y P2), 2) una secuencia "etiqueta" que es diferente para cada muestra y que se utilizará para la identificación del origen de la colonia de ADN y 3) secuencias cebadoras que rodean la secuencia específica de interés (S1 y S2). Obsérvese que para cada muestra, se podría utilizar una pluralidad de etiquetas, como en el (ii), anterior.

5

[0148] En adelante, los fragmentos de ADN obtenidos después de dicho proceso adecuado se denominarán con la expresión: "ADN etiquetado" (III, figura 13d). Los usos posibles de las etiquetas son los mismos que los indicados en el punto (ii) anterior.

10 (v) Preparación de ARNm

[0149] El procedimiento es similar a los procedimientos descritos para preparar fragmentos de ADN en los apartados anteriores, excepto que el punto de partida es extraer ARNm mediante cualquier medio conocido por los expertos en la técnica (por ejemplo, mediante la utilización de kits de preparación de ARNm disponibles en el comercio). El ARNm puede copiarse en ADNc bicatenario mediante cualquier medio conocido por los expertos en la técnica (por ejemplo utilizando una transcriptasa inversa y una ADN polimerasa). Ciertamente, las etiquetas y los cebadores descritos anteriormente se pueden utilizar junto con los procesos de síntesis de ADNc bicatenario para permitir su incorporación en los moldes. En lo sucesivo, los fragmentos de ARNm obtenidos después de dichos procesos adecuados se denominarán con las expresiones: "ARNm total preparado" (cf. "ADN genómico preparado", como se describe en el apartado (I) anteriormente), "ARNm total etiquetado", (cf. "ADN genómico etiquetado", como se describe en el apartado (ii) anteriormente), "ARNm preparado" (cf. "ADN preparado", como se describe en el apartado (iii) anteriormente) y "ARNm etiquetado" (cf. "ADN etiquetado", como se describe en el apartado (iv) anteriormente).

25 **H) Ensayos de detección preferidos**

[0150] En los procedimientos de ensayo de la presente invención se pueden utilizar marcadores para proporcionar señales detectables. Los ejemplos incluyen:

30 a) un grupo fluorescente o un sistema de fluorescencia basado en transferencia de energía.

b) un sistema basado en biotina. En este caso las colonias pueden incubarse con estreptavidina marcada con un grupo fluorescente o una enzima (por ejemplo perlas de látex fluorescentes cubiertas con estreptavidina; estreptavidina marcada con grupos fluorescentes; enzimas para su uso con el ensayo de fluorescencia correspondiente).

c) un sistema basado en la detección de un antígeno o su fragmento, por ejemplo, un hapteno (incluyendo biotina y grupos fluorescentes). En este caso, las colonias pueden incubarse con anticuerpos (por ejemplo específicos para un hapteno). Los anticuerpos pueden marcarse con un grupo fluorescente o con una enzima (por ejemplo perlas de látex fluorescentes cubiertas con el anticuerpo; anticuerpos marcados con grupos fluorescentes; anticuerpos ligados a una enzima para su uso con un ensayo de fluorescencia o luminiscencia correspondiente, etc.).

d) un radiomarcador (por ejemplo, incorporado utilizando una 5'-polinucleótido quinasa y [γ - 32 P]adenosina trifosfato o una ADN polimerasa y [α - 32 P o α - 33 P] desoxirribonucleósido trifosfatos para añadir uno o más grupos fosfato radioactivos a un ácido nucleico). En este caso las colonias pueden incubarse con un líquido de centelleo.

e) un colorante u otro agente de tinción.

50 **[0151]** Los marcadores para su uso en la presente invención están preferentemente unidos a

a) ácidos nucleicos

b) proteínas que se unen específicamente a ADN bicatenario (por ejemplo, histonas, represores, potenciadores) y/o

c) proteínas que se unen específicamente a ADN monocatenario (por ejemplo, proteína de unión a ácido nucleico monocatenario).

60 **[0152]** Las colonias marcadas se detectan preferentemente mediante:

a) medición de la fluorescencia

b) medición de la luminiscencia.

65

c) medición de la radioactividad.

d) medición del flujo o de la anisotropía fluorescente inducida por el campo eléctrico y/o

5 e) medición del grosor de la capa polimérica.

[0153] En la presente invención se pueden utilizar agentes de tinción. Por tanto las colonias de ADN pueden incubarse con un agente de tinción específico de ADN adecuado, tal como colorantes intercalantes, bromuro de etidio, YO-YO, YO-PRO (Molecular Probes, Eugene, OR).

10 [0154] Con determinados ejemplos de tinción el resultado puede observarse con un aparato adecuado de formación de imágenes fluorescentes.

15 [0155] A continuación se describirán ejemplos de ensayos/procedimientos particulares con más detalle

I) Realizaciones preferidas de ensayos de la presente invención

i) Ensayo de hibridación de sondas de ácido nucleico

20 [0156] Las colonias de ADN se preparan primero para la hibridación. Después, se hibridan con una sonda (marcada o no). Si se requiere, la sonda hibridada se somete a ensayo, y se observa el resultado. Esto se puede realizar con un aparato de la presente invención (por ejemplo, como se ha descrito anteriormente).

Preparación para la hibridación

25 [0157] En una realización preferida de la presente descripción, las colonias se tratan con una endonucleasa de restricción de ADN que es específica de una secuencia proporcionada por una forma monocatenaria de uno de los cebadores originalmente injertados sobre la superficie donde se forman las colonias o de otra secuencia presente en una molécula de ADN molde (véase, por ejemplo, la figura 16c).

30 [0158] Después de la digestión con la enzima de restricción, las colonias pueden calentarse a una temperatura lo suficientemente alta como para separar las moléculas de ADN bicatenarias. Después de esta etapa de desnaturalización térmica, las colonias se lavan para retirar las cadenas de ADN monocatenario separadas, no hibridadas, dejando un ADN monocatenario restante unido.

35 [0159] En otra realización las colonias pueden digerirse parcialmente con una exonucleasa de ADN bicatenario de 3' a 5' (véase el apartado E, figura 10f) que retira una cadena de los dúplex de ADN comenzando desde el extremo 3', por tanto dejando una parte de una molécula de ADN en una forma monocatenaria.

40 [0160] Como alternativa, el ADN en las colonias puede primero desnaturalizarse con calor y después digerirse parcialmente con una exonucleasa de ADN monocatenario específica de 3' a 5' que digiere ADN monocatenario comenzando desde el extremo 3'.

45 [0161] Una alternativa adicional es simplemente desnaturalizar con calor el ADN en las colonias.

Hibridación de la sonda

50 [0162] Las sondas de ácido nucleico monocatenario (marcadas o no) pueden hibridarse con ADN monocatenario en las colonias en condiciones de temperatura y tampón apropiadas (que depende de la secuencia de cada sonda, y que se puede determinar utilizando protocolos conocidos por los expertos en la técnica).

Ensayo de sondas hibridas no marcadas

55 [0163] Una sonda hibridada proporcionada inicialmente en una forma no marcada se puede utilizar con un cebador para la incorporación de los diferentes (o un subconjunto de los diferentes) desoxirribonucleósido trifosfatos marcados (o una mezcla de marcados y no marcados) con una ADN polimerasa. Los nucleótidos marcados incorporados pueden después detectarse como se ha descrito anteriormente.

Ensayo cíclico de sondas marcadas o no marcadas

60 [0164] En primer lugar, las colonias de ADN pueden prepararse para la hibridación mediante los métodos descritos anteriormente. Después pueden hibridarse con una sonda (marcada o inicialmente no marcada). Si se desea, las sondas marcadas hibridadas se someten a ensayo y los resultados se observan con un aparato como se ha descrito anteriormente. Después, la sonda puede retirarse mediante desnaturalización con calor y puede hibridarse y
65 detectarse una sonda específica para una secuencia de ADN. Estas etapas pueden repetirse con nuevas sondas

tantas veces como se desee.

[0165] En segundo lugar, las sondas pueden someterse a ensayo como se ha descrito anteriormente para sondas no marcadas, excepto que en cada ciclo se utiliza únicamente un subconjunto (preferentemente 1 solo) de los diferentes nucleótidos (marcados o no). Después, las colonias pueden someterse a ensayo para monitorizar la incorporación de los nucleótidos. Este segundo proceso puede repetirse hasta que se determine una secuencia de una longitud deseada.

(ii) Ensayo de síntesis de ARN *in situ*

10

[0166] En esta realización, se pueden utilizar colonias de ADN como moldes para la síntesis de ARN *in situ* como se representa en la figura 14. Las colonias de ADN pueden generarse a partir de moldes y cebadores, de tal manera que un promotor de ARN polimerasa se coloca en un extremo del ADN bicatenario en la colonia. Después, las colonias de ADN pueden incubarse con ARN polimerasa y el ARN recién sintetizado (ARNc) puede someterse a ensayo según se desee. La detección se puede realizar de manera no específica (por ejemplo, con tinción) o de una manera dependiente de secuencia (por ejemplo, hibridación).

[0167] El molde de ADN (I, figura 14a) a amplificar en una colonia se genera mediante reacción de PCR utilizando cebadores (Ia y Ib) que tienen las cuatro partes siguientes; 1) una secuencia idéntica a las secuencias de los cebadores injertados sobre la superficie ('P1' y 'P2'), 2) una secuencia "etiqueta" que es diferente para cada muestra, 3) una secuencia correspondiente a un promotor de ARN polimerasa, es decir, los promotores T3, T7 y SP6 de ARN ('RPP', figura 14a) y 4) secuencias cebadoras que rodean la secuencia específica de interés ('S1' y 'S2'). En lo sucesivo, los fragmentos de ADN obtenidos después de dicho proceso adecuado se denominarán con la expresión: "síntesis de ADN y ARN etiquetado" (III, figura 14b).

25

[0168] Después de la amplificación del molde de ADN a partir de la muestra de ADN original, estos moldes se utilizan para generar colonias de ADN. Las colonias de ADN (IV, figura 14c) se incuban después con la ARN polimerasa específica para el promotor de la ARN polimerasa ('RPP', figura 14c). Esto generará una copia de ARN específico para el molde de la colonia de ADN (ARNc molde, V, figura 14d).

30

[0169] El ARNc así sintetizado se puede aislar y utilizar como sondas de hibridación, como moldes de ARN mensajero (ARNm) para la síntesis de proteínas *in vitro* o como moldes para el análisis de secuencias de ARN *in situ*.

35 (iii) Métodos de secuenciación

[0170] En otra realización de la presente invención, las colonias pueden analizarse para determinar secuencias de moléculas de ácido nucleico que forman las colonias. Dado que en cada colonia pueden proporcionarse cantidades muy grandes de las mismas moléculas de ácido nucleico, es probable que la fiabilidad de los datos de secuenciación obtenidos sea muy alta.

40

[0171] Las secuencias determinadas pueden ser parciales o completas. Las secuencias de los ácidos nucleicos presentes en una o más colonias pueden determinarse. Al mismo tiempo, puede determinarse una pluralidad de secuencias.

45

[0172] En algunas realizaciones, inicialmente se puede obtener la secuencia de una cadena complementaria con una cadena de ácido nucleico a secuenciar (o de una parte de la misma). Sin embargo esta secuencia puede convertirse utilizando normas de emparejamiento de bases para proporcionar la secuencia deseada (o aparte de la misma). Esta conversión la puede realizar un ordenador o una persona. Se puede realizar después de cada etapa de extensión de cebadores o se puede realizar en una fase posterior.

50

[0173] La secuenciación se puede realizar mediante diversos métodos. Por ejemplo, métodos basados en digestión con endonucleasas de restricción secuencial y se puede utilizar ligamiento con enlazador. Un método de este tipo se desvela, por ejemplo, en el documento WO95/27080. Este método comprende las etapas de: ligar una sonda con un extremo de un polinucleótido, teniendo la sonda un sitio de reconocimiento de nucleasas; identificar uno o más nucleótidos en el extremo del polinucleótido; y escindiendo el polinucleótido con una nucleasa que reconozca el sitio de reconocimiento de nucleasas de la sonda, de tal manera que el polinucleótido se acorte en uno o más nucleótidos.

55

[0174] Sin embargo en un método preferido de la presente invención, las moléculas de ácido nucleico amplificadas (preferentemente en forma de colonias, como se desvela en el presente documento) se secuencian permitiendo que los cebadores se hibriden con las moléculas de ácido nucleico, extendiendo los cebadores y detectando los nucleótidos utilizados en la extensión del cebador. Preferentemente, después de la extensión de un cebador mediante un solo nucleótido, el nucleótido se detecta antes de que se use un nucleótido adicional en una extensión de cebador (secuenciación gradual).

65

[0175] Uno o más de los nucleótidos que se utilizan en la extensión de cebadores puede marcarse. La utilización de nucleótidos marcados durante la extensión de cebadores facilita la detección. (El término “marcador” se usa en un sentido general para indicar cualquier resto que pueda identificarse utilizando un sistema de detección apropiado. Preferentemente el marcador no está presente en nucleótidos de origen natural). En el mejor de los casos, los marcadores son no radioactivos, tales como fluoróforos. Sin embargo se pueden utilizar marcadores radioactivos.

[0176] Cuando se proporcionan nucleótidos en forma marcada, los marcadores pueden ser los mismos para diferentes nucleótidos. Si se utiliza el mismo marcador, se puede utilizar cada incorporación de nucleótido para proporcionar un aumento acumulativo de la misma señal (por ejemplo, de una señal detectada a una longitud de onda particular). Como alternativa se pueden utilizar diferentes marcadores para cada tipo de nucleótido (que puede detectarse a diferentes longitudes de onda).

[0177] Por tanto, se pueden proporcionar cuatro marcadores diferentes para dATP, dTTP, dCTP y dGTP, o se puede proporcionarse el mismo marcador para todos ellos. De manera similar, se pueden proporcionar cuatro marcadores diferentes para ATP, UTP, CTP y GTP, o se puede proporcionar el mismo marcador para todos ellos). En algunas realizaciones de la presente invención se puede proporcionar una mezcla de nucleótidos marcados o no, como se describirá más adelante con más detalle.

[0178] En una realización preferida de la presente invención, la secuenciación de las moléculas de ácido nucleico presentes en al menos 2 colonias diferentes, se realiza de manera simultánea. Más preferentemente, la secuenciación de las moléculas de ácido nucleico presentes en más de 10, más de 100, más de 1000 o incluso más de 1.000.000 de colonias diferentes se realiza de manera simultánea. Por lo tanto si se proporcionan colonias que tienen diferentes moléculas de ácido nucleico, se pueden determinar muchas secuencias diferentes (completas o parciales) de manera simultánea, es decir, se pueden determinar más de 10, más de 100, más de 1000 o incluso más de 1.000.000 secuencias diferentes de manera simultánea.

[0179] Si se desea, se pueden proporcionar controles, mediante los cuales se proporciona una pluralidad de colonias que comprenden las mismas moléculas de ácido nucleico. Determinando si se obtienen o no las mismas secuencias para las moléculas de ácido nucleico en estas colonias se puede confirmar si el procedimiento de secuenciación es o no fiable.

[0180] En la figura 17 se ilustra un método de secuenciación de la presente invención, que se titula “secuenciación in situ”. En las colonias de ADN preparadas, hibridadas con un cebador de secuenciación apropiado, la adición cíclica de los desoxirribonucleótido trifosfatos individuales y de la ADN polimerasa permitiría la determinación de la secuencia de ADN inmediatamente en posición 3' con respecto al cebador de secuenciación. En el ejemplo indicado en la figura 17, la adición de dGTP permite la determinación de 1 colonia que contiene 'G'. En el segundo ciclo de adición de dATP se detecta en ambas colonias, determinando que las dos colonias tienen 'A' en la siguiente posición. Después de diversas repeticiones de la adición de desoxirribonucleósido trifosfatos sencillos, será posible determinar cualquier secuencia. Por ejemplo, se pueden determinar secuencias de al menos 10, al menos 20, al menos 50 o al menos 100 bases.

[0181] Si se proporcionan colonias inicialmente en una forma que comprenden moléculas bicatenarias, las colonias pueden procesarse para proporcionar moléculas monocatenarias para su uso en la secuenciación, como se ha descrito anteriormente. (Sin embargo se debe observar que las moléculas bicatenarias se pueden utilizar para la secuenciación sin dicho procesamiento. Por ejemplo, una molécula de ADN bicatenaria se puede proporcionar con una secuencia promotora y la secuenciación gradual se puede realizar después utilizando una ARN polimerasa y ribonucleótidos marcados (cf Figura 16 d). (Otra alternativa es introducir una muesca en una molécula de ADN bicatenaria de tal manera que se pueda realizar la traducción de la muesca utilizando desoxirribonucleótidos marcados y una ADN polimerasa con actividad exonucleasa de 5' a 3').

[0182] Una manera de procesar moléculas bicatenarias presentes en colonias para proporcionar colonias monocatenarias se describe más adelante con referencia a la Figura 19. En este caso, moléculas inmovilizadas bicatenarias presentes en una colonia (que pueden estar en forma de estructuras de tipo puente) se escinden y después de esto se realiza una etapa de desnaturalización. (Como alternativa, inicialmente se podría utilizar una etapa de desnaturalización y después se podría realizar una etapa de escisión). Preferentemente, la escisión se realiza enzimáticamente. Sin embargo, también son posibles otros medios de escisión, tal como escisión química. (Se puede proporcionar un sitio de escisión apropiado en dicha molécula). La desnaturalización se puede realizar mediante cualquier medio adecuado. Por ejemplo, se puede realizar calentando y/o cambiando la fuerza iónica de un medio en las proximidades de las moléculas de ácido nucleico.

[0183] Una vez que se han proporcionado las moléculas monocatenarias a secuenciar, con las mismas se pueden hibridar cebadores adecuados para la extensión de los cebadores. Como cebadores se prefieren oligonucleótidos que son moléculas de ácido nucleico que tienen típicamente una longitud de 6 a 60, por ejemplo, de 15 a 25 nucleótidos. Estos pueden comprender nucleótidos de origen natural y/o no natural. (Sin embargo, como alternativa, si se desea, se puede utilizar otras moléculas, por ejemplo, cadenas de ácido nucleico más largas, como

cebadores). Preferentemente, los cebadores para su uso en la secuenciación se hibridan con las mismas secuencias presentes en las moléculas de ácidos nucleicos amplificados así como lo hacen los cebadores que se utilizaron para proporcionar dichos ácidos nucleicos amplificados. (Los cebadores que tienen las mismas/similares secuencias se pueden utilizar para fines tanto de amplificación como de secuenciación).

5

[0184] Cuando los cebadores se proporcionan en solución y están apareados (hibridados) con las moléculas de ácido nucleico presentes en las colonias a secuenciar, aquellos cebadores que quedan en solución o que no se pueden aparear específicamente tienen que retirarse después del apareamiento. Las condiciones de apareamiento preferidas (temperatura y composición del tampón) impiden la hibridación no específica. Estas condiciones pueden ser rigurosas. Dichas condiciones serán típicamente temperaturas de apareamiento cercanas a la T_f (temperatura de fusión) del cebador a una concentración salina determinada (por ejemplo, cebador 50 nM en tampón NaCl 200 mM a 55 °C para un oligonucleótido de 20 unidades monoméricas (meros) con un contenido de GC del 50 %). (Un experto en la técnica puede determinar las condiciones rigurosas para un sistema determinado. Estas dependerán de la composición de bases, del contenido de GC, de la longitud del cebador que se utilice y de la concentración salina. Para un oligonucleótido de 20 bases de promedio calculado de GC al 50 %, la temperatura de apareamiento es de 55 a 60 °C, pero en la práctica puede variar entre 35 a 70 °C).

[0185] No es preciso que los cebadores utilizados para la extensión de cebadores tengan que proporcionarse en solución, ya que pueden proporcionarse en forma inmovilizada. En esta realización los cebadores deben proporcionarse cerca de las moléculas inmovilizadas con las que se van a aparear. (De hecho, dichos cebadores pueden estar ya presentes como cebadores inmovilizados en exceso que no se utilizaron en la amplificación de las moléculas de ácido nucleico durante la formación de colonias).

[0186] Las moléculas de ácido nucleico presentes en las colonias a secuenciar, incluirán una secuencia que se hibrida con los cebadores que se van a utilizar en la secuenciación (preferentemente en condiciones "rigurosas"). Esta parte puede añadirse a una molécula determinada antes de la amplificación (cuya molécula puede tener una secuencia total/parcialmente desconocida) utilizando técnicas conocidas para los expertos en la materia. Por ejemplo se puede sintetizar artificialmente y se puede añadir a una molécula determinada utilizando una ligasa.

[0187] Una vez que se proporciona una molécula de ácido nucleico apareada con un cebador, se puede realizar el cebador de extensión. Se pueden utilizar ARN o ADN polimerasas. Sin embargo, las ADN polimerasas son las enzimas de elección para las realizaciones preferidas. Algunas de estas se encuentran disponibles en el comercio. Se pueden utilizar polimerasas que no tengan actividad exonucleasa de 3' a 5', tales como la ADN polimerasa de T7 o se puede utilizar el fragmento pequeño (Klenow) de la ADN polimerasa I [por ejemplo la ADN polimerasa de T7 modificada Sequenase™ 2.0 (Amersham) o el fragmento Klenow (exo de 3' a 5', New England Biolabs)]. Sin embargo, no es esencial el uso de dichas polimerasas. De hecho, cuando se desee que las polimerasas tengan actividad correctora de errores, no se debería utilizar polimerasas que no tengan actividad exonucleasa de 3' a 5'. Determinadas aplicaciones pueden requerir el uso de polimerasas termoestables, tales como ThermoSequenase™ (Amersham) o Taquenase™ (ScienTech, St Louis, MO). Para las reacciones de extensión de cebadores se puede utilizar cualquier nucleótido (de origen natural o no natural). Los nucleótidos preferidos son los desoxirribonucleótidos: dATP, dTTP, dGTP y dCTP (aunque en algunas aplicaciones se prefiere dUPT, el análogo de dTTP) o los ribonucleótidos ATP, UTP, GTP y CTP; al menos algunos de los cuales se proporcionan en forma marcada.

[0188] Preferentemente, después de cada etapa de extensión de cebador se incorpora una etapa de lavado para retirar los nucleótidos no incorporados que pueden interferir con las etapas posteriores. La solución de lavado preferida debe ser compatible con la actividad polimerasa y tener una concentración salina que no interfiera con el apareamiento de las moléculas cebadoras con las moléculas de ácido nucleico a secuenciar. (En realizaciones menos preferidas, la solución de lavado puede interferir con la actividad polimerasa. En este caso la solución de lavado no requiere retirarse antes de la extensión del cebador adicional).

[0189] Considerando que en una colonia determinada se pueden proporcionar muchas copias de moléculas a secuenciar, se puede utilizar una combinación de nucleótidos marcados y no marcados. En este caso, incluso si una pequeña proporción de los nucleótidos está marcada (por ejemplo, marcada con fluorescencia), la cantidad de marcadores incorporados en cada colonia durante la extensión de cebadores puede ser suficiente para detectar con un dispositivo de detección. Por ejemplo, la proporción de nucleótidos marcados con respecto a no marcados, puede seleccionarse de tal manera que, por término medio, los nucleótidos marcados que se utilicen en la extensión de cebadores, sea menor de 50 %, menor de 20 %, menor de 10 % o incluso menor de 1 % del momento (es decir, por término medio, en una etapa de extensión de cebadores determinada, un nucleótido se incorpora en forma marcada en menos del 50 %, menos del 20 %, menos del 10 % o menos del 1 % de los cebadores extendidos).

[0190] Por tanto en una realización adicional de la presente descripción se proporciona un método de secuenciación de moléculas de ácido nucleico presentes en una colonia de la presente invención, comprendiendo el método las etapas de:

65

- a) proporcionar al menos una colonia que comprenda una pluralidad de moléculas de ácido nucleico monocatenario que tenga las mismas secuencias entre sí y que se hibride con cebadores de una manera que se permita la extensión de cebadores en presencia de nucleótidos y de una polimerasa de ácido nucleico;
- 5 b) proporcionar dicha al menos una colonia con una polimerasa de ácido nucleico y un nucleótido determinado en forma marcada o no marcada en condiciones que permitan la extensión de los cebadores si una base complementaria o si una pluralidad de dichas bases está presente en la posición apropiada en las moléculas de ácido nucleico monocatenario presentes en dicha al menos una colonia;
- 10 c) detectar si dicho nucleótido marcado se ha utilizado o no para la extensión de cebadores determinando si el marcador presente en dicho nucleótido se ha incorporado o no en los cebadores extendidos.

[0191] Las etapas b) y c) pueden repetirse una o más veces. Preferentemente, se proporciona una pluralidad de colonias diferentes y se determinan diversas secuencias diferentes de manera simultánea.

15

[0192] Esta realización adicional de la presente descripción se puede utilizar para reducir costes, ya que se necesitan relativamente pocos nucleótidos marcados. También se puede utilizar para reducir efectos de desactivación.

- 20 **[0193]** Sin embargo también es posible utilizar nucleótidos únicamente marcados para la extensión de cebadores o utilizar una mayor proporción de los mismos (por ejemplo, más del 50 %, más del 70 % o más del 90 % de los nucleótidos utilizados pueden estar marcados). Esto se puede realizar, por ejemplo, si se seleccionan marcadores para impedir o reducir efectos de desactivación. Como alternativa, pueden retirarse o neutralizarse marcadores a diversas fases que se vuelven problemáticos por efectos de desactivación (por ejemplo se puede realizar
- 25 blanqueamiento láser de fluoróforos). Sin embargo, esto puede aumentar el número de etapas necesarias y por lo tanto se prefiere no retirar los marcadores (o al menos no retirarlos después de haberse incorporado cada nucleótido sino que únicamente se retiren periódicamente). En otras realizaciones menos preferidas, el propio cebador y su producto de extensión se pueden retirar y reemplazar por otro cebador. Si se requiere, se pueden realizar diversas etapas secuenciales de adiciones de nucleótidos sin marcador antes de comenzar la secuenciación real en
- 30 presencia de nucleótidos marcados. Una alternativa adicional es utilizar un tipo de marcador diferente al que se utiliza inicialmente (por ejemplo, cambiando de fluoresceína a rodamina) que se vuelve problemático por efectos de desactivación.

[0194] En realizaciones preferidas de la presente descripción en un cebador extendido durante la secuenciación se incorpora una pluralidad de bases marcadas. Esto es ventajoso ya que puede aumentar la velocidad del procedimiento de secuenciación con respecto a los métodos en los que, una vez incorporada una base marcada en un cebador extendido, el marcador debe retirarse antes de que pueda incorporarse una base marcada adicional. (La pluralidad de bases marcadas puede estar en forma de uno o más tramos contiguos, aunque esto no es esencial).

40 **[0195]** Por lo tanto, la presente invención se refiere a un método para secuenciar moléculas de ácido nucleico, que comprende las etapas de:

a) utilizar una primera colonia para proporcionar una pluralidad de moléculas de ácido nucleico monocatenario que tengan las mismas secuencias entre si y que se hibriden con cebadores de una manera que se permita la

45 extensión de cebadores en presencia de nucleótidos y de una polimerasa de ácido nucleico;

b) utilizar una segunda colonia para proporcionar una pluralidad de moléculas de ácido nucleico monocatenario que tengan las mismas secuencias entre sí, y que también se hibriden con cebadores de una manera que se permita la extensión de cebadores en presencia de nucleótidos y de una polimerasa de ácido nucleico;

50

c) proporcionar a cada colonia una polimerasa de ácido nucleico y un nucleótido marcado determinado en condiciones que permitan la extensión de los cebadores si una base complementaria o una pluralidad de dichas bases están presentes en la posición apropiada en las molécula de ácido nucleico monocatenario;

55 d) detectar si dicho nucleótido marcado se ha utilizado o no para la extensión de cebadores en cada colonia, determinando si el marcador presente en dicho nucleótido se ha incorporado o no en los cebadores extendidos;

e) repetir las etapas c) y d) una o más veces de manera que se proporcionen cebadores extendidos que comprendan una pluralidad de marcadores.

60

[0196] Preferentemente las secuencias de las moléculas de ácido nucleico presentes en dicha primera localización y en dichas localizaciones son diferentes entre sí, es decir, se secuencian una pluralidad de colonias que comprenden moléculas de ácido nucleico diferentes.

65 **[0197]** A la vista de la descripción anterior, se apreciará que se puede utilizar una gran cantidad de métodos de

secuenciación diferentes utilizando colonias de la presente invención. En estos métodos se pueden utilizar diversos sistemas de detección para detectar marcadores utilizados en secuenciación (aunque en determinadas realizaciones la detección puede ser posible simplemente a ojo, de manera que no se necesite un sistema de detección). Un sistema de detección preferido para marcadores fluorescentes es una cámara de dispositivo de carga acoplada (CCD) que opcionalmente se puede acoplar a un dispositivo de aumento. Se puede utilizar cualquier otro dispositivo que permita la detección y, preferentemente, también la cuantificación de fluorescencia en una superficie. Se pueden seleccionar dispositivos tales como dispositivos formadores de imágenes fluorescentes o microscopios confocales.

[0198] En realizaciones menos preferidas, los marcadores pueden ser radioactivos y por tanto podría ser necesario un dispositivo de detección de radioactividad. En el mejor de los casos, dichos dispositivos serían sistemas de formación de imágenes de radioactividad en tiempo real. También, menos preferidos son otros dispositivos basados en pantallas de fósforo (Molecular Dynamics) o en películas de autorradiografía para la detección.

[0199] Dependiendo de la cantidad de colonias a monitorizar, se puede preferir un sistema de escaneo para recoger datos. (Aunque una alternativa es proporcionar una pluralidad de detectores para permitir recuperar todas las colonias). Dicho sistema permite que un detector se desplace con respecto a una pluralidad de colonias a analizar. Esto es útil cuando todas las colonias que proporcionan señales no están dentro del campo visual de un detector. El detector puede mantenerse en una posición fija y las colonias a analizar pueden desplazarse al campo visual del detector (por ejemplo, mediante una plataforma móvil). Como alternativa las colonias pueden mantenerse en una posición fija y el dispositivo de detección puede desplazarse para llevarlas a su campo visual.

[0200] El sistema de detección se utiliza preferentemente en combinación con un sistema de análisis para determinar la cantidad de bases (y preferentemente también la naturaleza) incorporada por la extensión de cebadores en cada colonia después de cada etapa. Este análisis se puede realizar inmediatamente después de cada etapa o más adelante, utilizando datos registrados. La secuencia de las moléculas de ácido nucleico presentes dentro de una colonia determinada puede después deducirse a partir del número y del tipo de nucleótidos añadidos después de cada etapa.

[0201] Preferentemente el sistema de detección es parte de un aparato que comprende otros componentes. La presente invención incluye un aparato que comprende una pluralidad de nucleótidos marcados, una polimerasa de ácido nucleico y medios de detección para detectar nucleótidos marcados cuando se incorporan en una molécula de ácido nucleico por extensión de cebadores, adaptándose los medios de detección para diferenciar entre señales proporcionadas por los nucleótidos marcados incorporados en las diferentes colonias.

[0202] El aparato también puede incluir medios de control de temperatura, de suministro de disolventes y de lavado. También puede ser automático.

[0203] Los métodos y aparatos dentro del ámbito de la presente invención se pueden utilizar en la secuenciación de:

- moléculas de ácido nucleico no identificadas (es decir secuenciación *de novo*);
- y moléculas de ácido nucleico que se van a secuenciar para verificar si están presentes una o más diferencias relativas a una secuencia conocida (por ejemplo, identificación de polimorfismos). Algunas veces esto recibe el nombre de "resecuenciación".

[0204] Tanto la secuenciación *de novo* como la resecuenciación se analizan más adelante con mayor detalle (véanse los siguientes apartados (v) y (vi)).

[0205] Para las aplicaciones de secuenciación *de novo*, el orden de los nucleótidos aplicados a una localización determinada puede seleccionarse según se desee. Por ejemplo se puede seleccionar la adición secuencial de los nucleótidos dATP, dTTP, dGTP, dCTP; dATP, dTTP, dGTP, dCTP; y así sucesivamente. (Generalmente se repetirá un solo orden de cuatro nucleótidos, aunque esto no es esencial). Para las aplicaciones de resecuenciación, el orden de los nucleótidos a añadir en cada etapa se selecciona preferentemente de acuerdo con una secuencia conocida.

[0206] La resecuenciación puede ser de particular interés para el análisis de una gran cantidad de moléculas molde similares para detectar e identificar diferencias de secuencias (por ejemplo para el análisis de plásmidos recombinantes en clones candidatos después de mutagénesis dirigida o, lo que es más importante, para la exploración de polimorfismos en una población). Las diferencias de una secuencia determinada pueden detectarse mediante la ausencia de incorporación de uno o más nucleótidos presentes en la secuencia determinada a fases particulares de extensión de cebadores. A diferencia de las técnicas más habitualmente usadas, el método de la presente invención permite la detección de cualquier tipo de mutación, tal como mutaciones, inserciones, o deleciones puntuales. Además, no solo mutaciones existentes conocidas, sino también mutaciones previamente no identificadas pueden caracterizarse con el suministro de información de secuencias.

[0207] En algunas realizaciones de la presente invención, pueden haberse secuenciado moléculas de ácido nucleico largas mediante diversas reacciones de secuenciación, permitiendo cada una de ellas la determinación de parte de la secuencia completa. Estas reacciones se pueden realizar en diferentes colonias (donde cada una de las diferentes colonias se proporciona con las mismas moléculas de ácido nucleico a secuenciar pero con diferentes cebadores), o en ciclos sucesivos aplicados a la misma colonia (donde entre cada uno de los ciclos, los cebadores y los productos de extensión se retiran mediante lavado y se reemplazan por cebadores diferentes).

(iv) Huella de ADN

10 **[0208]** Esta realización de la presente invención tiene por objeto resolver el problema de explorar una gran población para la identificación de características determinadas de genes determinados, tales como la detección de polimorfismos mononucleotídicos.

15 **[0209]** En una realización preferida, esto consiste en generar ADN genómico etiquetado (véase el apartado G(ii) anterior). (Por tanto cada muestra originada a partir de una muestra individual determinada se ha marcado con una sola etiqueta). Este ADN etiquetado se puede utilizar para generar colonias primarias en una superficie apropiada que comprende cebadores inmovilizados. Después se pueden realizar diversos ensayos de hibridación de sondas sucesivos en las colonias. Entre cada ensayo la sonda precedente puede retirarse, por ejemplo, mediante desnaturalización térmica y lavado.

20 **[0210]** Las ventajas de esta realización de la presente invención sobre las estrategias para resolver este problema se ilustran en el siguiente ejemplo de una posible aplicación práctica.

25 **[0211]** Se pretende detectar qué parte de un gen (de un tamaño, por ejemplo, de 2000 bases), si hubiere, está relacionada con un fenotipo de enfermedad en una población de típicamente 1.000 a 10.000 individuos. Para cada individuo, se puede realizar una amplificación por PCR para amplificar específicamente el gen de interés y ligar una etiqueta y un cebador que genere una colonia (en referencia al apartado G(iv), preparación de "ADN etiquetado").

30 **[0212]** Para obtener una matriz representativa de muestras, se podría querer disponer 500.000 colonias al azar (es decir una redundancia de 10 veces, para tener solamente una pequeña probabilidad de perder la detección de una muestra. Con una densidad de colonia de 10.000 colonias por mm², se puede utilizar una superficie de -7 mm x 7 mm. Esta es una superficie mucho más pequeña en comparación con cualquier otra tecnología disponible hasta ahora (por ejemplo, la estrategia HySeq utiliza 220 mm x 220 mm para el mismo número de muestras (50.000) sin redundancia). La cantidad de reactivos (una gran parte del coste) será proporcional a la superficie ocupada por la matriz de las muestras. Por tanto, la presente invención puede ofrecer una mejora de 800 veces sobre la tecnología actualmente disponible.

40 **[0213]** La utilización de un aparato para monitorizar el resultado de la secuenciación o de los ensayos de hibridación con sonda 'in situ', conlleva un tiempo del orden de 1 a 10 segundos para formar imágenes de una señal fluorescente de las colonias ensayadas utilizando fluorescencia presente en una superficie de -1 mm². Por tanto, se asume que el cuello de botella del método es el tiempo que se necesita para la formación de imágenes resultantes del ensayo, que lleva del orden de 10 minutos para la formación de imágenes resultantes de un ensayo de 50.000 muestras (500.000 colonias). Para proporcionar 200 ensayos que incluyen la formación de imágenes (en una o varias superficies de 7 mm x 8 mm), utilizando la presente invención puede llevar menos de 36 horas. Esto representa una mejora de 20 veces en comparación con el mejor método conocido hasta ahora (HySeq presume de 45 30 días para conseguir una tarea comparable).

50 **[0214]** Las mejoras (densidades de colonias 10 veces más altas y formación de imágenes en un tiempo de 1 segundo) permitirían un rendimiento mucho más alto y finalmente el rendimiento finalmente esperado podría ser de aproximadamente 2000 veces más rápido que la mejor tecnología, aunque no se ha demostrado del todo, disponible hasta ahora.

55 **[0215]** Otra ventaja de utilizar la presente invención se basa en el hecho de que supera el problema que se plantea con individuos que tienen mutaciones heterocigotas para un gen determinado. Aunque este problema puede abordarse mediante métodos de secuenciación existentes para determinar polimorfismos alélicos, los métodos de detección de mutaciones de alto rendimiento actuales basados en hibridación de sondas oligonucleotídicas puede conducir a dificultades en la interpretación de resultados debido a una hibridación de sondas desigual en casos de polimorfismos alélicos y por lo tanto pueden producirse errores. En esta realización de la presente invención, cada colonia surge de una sola copia de un gen de interés amplificado. Si se genera un promedio de 10 colonias para cada locus individual, habrá un promedio de 5 colonias correspondientes a una versión de un gen y 5 colonias correspondientes a la otra versión del gen. Por tanto pueden puntuarse mutaciones heterocigotas mediante el número de veces que se detecta un solo alelo por muestra genómica individual.

(v) Resecuenciación de ADN

65

[0216] Esta realización de la presente invención proporciona una solución al problema de identificar y caracterizar nuevos polimorfismos alélicos dentro de genes conocidos en una gran población de muestras biológicas.

[0217] En su realización preferida esto consiste en obtener ADN etiquetado (cada muestra que se origina de un individuo determinado se ha etiquetado con una etiqueta exclusiva - véase el apartado G(iv)). Este ADN codificado puede después utilizarse para generar colonias primarias en una superficie apropiada que comprende cebadores inmovilizados. Después se pueden realizar diversos ensayos de hibridación de sondas sucesivos con las colonias en los que entre cada ensayo cíclico la sonda precedente se puede retirar mediante desnaturalización térmica y lavado. Preferentemente, la secuencia de ADN en 3', con respecto a una sonda específica, se puede determinar directamente por secuenciación 'in situ' (apartado I(iii), Métodos de secuenciación).

[0218] Las ventajas de la presente invención sobre otras estrategias para resolver este problema se ilustran en el siguiente ejemplo de una posible aplicación práctica.

[0219] Se desea identificar la variabilidad de la secuencia de un gen (de un tamaño, por ejemplo, de 2000 bases), si hubiera, en una población de típicamente 4 000 individuos. Se supone que se conoce una secuencia de referencia del gen. Para cada individuo, se puede realizar una amplificación por PCR para amplificar específicamente el gen de interés y ligar una etiqueta y un cebador que genera colonias. Para obtener una matriz representativa de muestras, se podría querer disponer 40 000 colonias al azar (es decir una redundancia de 10 veces, para tener solamente una pequeña probabilidad de perder la detección de una muestra). Con una densidad de colonia de 10.000 colonias por mm², se puede utilizar una superficie de ~2 mm x 2 mm.

[0220] La utilización de un aparato con una cámara CCD (que tiene una placa de 2000 x 2000 píxeles) para monitorizar el resultado del ensayo, podría llevar del orden de 10 segundos para formar una imagen de una señal fluorescente de colonias en una superficie de 4 mm². Si se supone que es posible leer al menos 20 bases durante una ronda del ensayo, esto requiere 61 etapas de formación de imágenes (se necesitan 3n+1 etapas de formación de imágenes para leer n números de bases). Si se supone que el cuello de botella del método es el tiempo para formar las imágenes del resultado del ensayo, esto lleva del orden de 15 minutos para formar las imágenes del resultado de un ensayo de 4 000 muestras (40 000 colonias). Para realizar 100 ensayos (en una o varias superficies de 2x2 mm²) para cubrir todo el gen de interés, la presente invención puede permitir que todo el experimento de exploración se realice en aproximadamente un día, con un aparato. Esto puede compararse con los sistemas operativos más poderosos disponibles en la actualidad.

[0221] En esta realización de la presente invención, con suposiciones conservativas (densidad de colonia, tiempo de formación de imágenes, tamaño de la microplaca CCD), puede alcanzarse un rendimiento de 3,2x10⁶ bases por hora, es decir, una mejora de 400 veces cuando se compara con el sistema más comúnmente usado en la actualidad (los secuenciadores de ADN actuales tienen un rendimiento típico del orden de 8 000 bases de lectura/hora).

40 (vi) Secuenciación de ADN *de novo*

[0222] Esta realización de la presente invención se refiere a resolver el problema de secuenciación de nuevos genomas (o partes de los mismos) con bajo coste y en un corto periodo de tiempo, en el que la secuencia de ADN se desconoce. Puede prepararse un ADN genotipo, bien directamente a partir de ADN total de un organismo de interés o de un vector en el que se ha insertado ADN. El ADN genómico preparado (de cualquier fuente) se puede utilizar para generar colonias de ADN. Después, las colonias de ADN pueden someterse a digestión con una enzima de restricción cortadora poco frecuente, cuyo sitio está incluido en el enlazador, desnaturalizarse y secuenciarse.

[0223] La figura 16 representa un ejemplo de secuenciación de ADN *de novo*. En este ejemplo, el ADN genómico está fragmentado en trozos de 100 a 2 000 pares de bases (véase, preparación de fragmentos de ADN al azar, apartado G(i)). Estos fragmentos se ligarán a enlazadores oligonucleotídicos (IIa y IIIb, figura 16a) que incluyen secuencias específicas para los cebadores injertados sobre la superficie ('P1' y 'P2'), una secuencia que es reconocida por una nucleasa de restricción cortadora poco frecuente ('RE') y una secuencia correspondiente a un cebador de secuenciación ('SP'), produciendo moldes (III, figura 16b). Utilizando este ADN preparado como molde para la formación de colonias de ADN, se obtienen colonias primarias (IV, figura 16c). Esas colonias se someten después a digestión con la correspondiente endonucleasa de restricción y se desnaturalizan para retirar la cadena de ADN no unida (V, figura 16d). El cebador de secuenciación (SP) se aparea después con el molde monocatenario unido (figura 16e). La incorporación y detección de nucleótidos marcados se puede realizar después como se ha descrito anteriormente (véase el apartado I(iii), Métodos de Secuenciación).

[0224] En esta realización, el rendimiento que se puede obtener puede ser como mínimo 400 veces mayor que el que se puede obtener con los métodos disponibles en la actualidad.

(vii) Monitorización de la expresión génica de ARNm

65

[0225] Esta realización de la invención se refiere a resolver el problema de la monitorización de la expresión de una gran cantidad de genes de manera simultánea.

[0226] Su realización preferida se representa en la figura 17.

5

[0227] En primer lugar, se prepararon colonias primarias, como se representa en la figura 3. En esta forma preferida, el ADN utilizado para esta preparación es 'ADN genómico preparado' o 'ADN genómico etiquetado', como se describe en los apartados G(i) y G(iii), respectivamente, y en el que el ADN es del genoma completo de uno (o varios) organismo(s) o de un subconjunto de los mismos (por ejemplo, de una biblioteca de genes previamente aislados). En la figura 17, las letras mayúsculas "A", "B" y "D" representan colonias que han surgido de genes que presentan niveles de expresión alto, medio y bajo respectivamente, y "E" representa colonias que surgen de genes no expresados (en casos reales, todas estas situaciones pueden no estar necesariamente presentes de manera simultánea).

15 **[0228]** En segundo lugar, las colonias se trataron para cambiarlas en soportes (es decir cebadores secundarios) para el crecimiento de colonias secundarias (etapa i en la figura 17a), como se describe en el apartado E. En esta fase (figura 17b), las colonias tratadas se representan con caracteres subrayados (A, B, D o E).

[0229] En tercer lugar, (etapa ii en la figura 17b) este soporte para el crecimiento de colonias secundarias se usa para regenerar colonias de moldes de ARNm (o ADNc) extraídos de una muestra biológica, como se describe en el apartado C. Si el molde es ARNm, la etapa de cebado de regeneración de colonias se realizará con una transcriptasa inversa. Después de una cantidad de ciclos de amplificación de colonias determinado, preferentemente de 1 a 50, la situación será como se representa en la figura 17c: las colonias correspondientes a genes altamente expresados (representados por la letra "A") están totalmente regeneradas, ya que su regeneración se ha iniciado por muchas copias del ARNm; las colonias correspondientes a genes con niveles de expresión medios (representados por las letras "b" y "B"), se han regenerado solo parcialmente; solo algunas de las colonias correspondientes a genes poco frecuentes (representados por la letra "d") se han regenerado parcialmente; las colonias correspondientes a las secuencias no expresadas (representadas por la letra "E"), no se ha regenerado del todo.

30 **[0230]** Finalmente, (etapa iii en la figura 17c), se realizaron ciclos adicionales de crecimiento de colonias (preferentemente de 2 a 50), y las colonias que no se habían regenerado totalmente durante las etapas previas finalmente se regeneraron por completo, "b" se convierte en "B", "d" se convierte en "D" (figura 17d): las colonias correspondientes a genes con niveles de expresión altos y medios se regeneraron del todo "A" y "B" o "B"; las colonias correspondientes a genes con niveles de expresión bajos no se regeneraron del todo "D" y "D"; las colonias correspondientes a secuencias no expresadas no se regeneran del todo "E".

35

[0231] Los niveles de expresión relativos de los genes se pueden obtener mediante los siguientes métodos preferidos:

40 - En primer lugar, los niveles de expresión pueden monitorizarse siguiendo la tasa de regeneración de las colonias (es decir, midiendo la cantidad de ADN dentro de una colonia después de diferentes números de ciclos de crecimiento de colonias durante la etapa (iii)) así como la tasa a la cual una colonia se regenera estará relacionada con el número de moléculas de ARNm (o de ADNc) que iniciaron la regeneración de esa colonia (en primera aproximación, el número de moléculas de ADN después de n ciclos, indicado como M(n), en una colonia que se somete a regeneración debe darse mediante $M(n) = M_0 r^{(n-1)}$, en la que M_0 es el número de moléculas que inicia la regeneración de colonia, r es la tasa de crecimiento y n es el número de ciclos);

45

- En segundo lugar, los niveles de expresión pueden monitorizarse contando, para cada gen, el número de colonias que se ha regenerado y comparando este número con el número total de colonias correspondientes para ese gen. Estas mediciones generalmente darán acceso a los niveles de expresión relativos de los genes representados por las colonias. La identificación de las colonias se realiza preferentemente mediante huella de identificación, de una manera esencialmente similar a la de la realización, apartado I(iv). Obsérvese que no se requiere la codificación de las muestras de ADN, pero puede considerarse como una alternativa a la identificación directa del ADN en las colonias. Esto puede ser de interés práctico porque con la codificación se pueden utilizar los mismos códigos (por tanto los mismos oligonucleótidos implicados en el ensayo del código) para cualquiera de los conjuntos de genes, mientras que sin código, se debe utilizar un conjunto diferente de oligonucleótidos específicos para cada uno de los conjuntos de genes.

55

[0232] Esta realización de nuestra invención tiene muchas ventajas si se compara con el estado actual de la técnica entre las que se incluyen: un rendimiento muy alto; no se requiere amplificación previa del ARNm (aunque la amplificación previa es compatible con nuestra invención); se requieren pequeñas cantidades de muestras y reactantes debido a la alta densidad de las muestras con nuestra invención; la presencia de genes altamente expresados no tiene incidencia sobre la capacidad de monitorizar genes con bajos niveles de expresión; la capacidad de monitorizar de manera simultánea niveles de expresión bajos y altos dentro del conjunto de genes de interés.

65

[0233] Cuando el ADN inicial en la generación de la colonia de ADN primario se realiza a partir del ADN de un genoma completo, esta realización también proporciona las siguientes características: no hay interferencia entre genes que se expresan a alto nivel y a bajo nivel aun cuando no se ha realizado una amplificación específica de los genes de interés. Esta es una característica exclusiva del uso de la presente invención: la amplificación específica no es posible porque la suposición inicial de esta realización es monitorizar los genes de expresión que pueden incluso no haberse aislado, por tanto que son desconocidos, y por tanto para los que no se conocen secuencias específicas (exclusivas) y cuyas secuencias específicas deberían haber sido necesarias para la amplificación de genes específicos. La capacidad de nuestra invención para realizar este tipo de monitorización de la expresión de ARNm se debe a que cuando se preparan las colonias primarias, estadísticamente, cada parte del genoma inicial se representará mediante el mismo número de colonias. Por tanto, el ADN frecuente y poco frecuente iniciará el mismo número de colonias (por ejemplo, una colonia por molécula de genoma añadida). La información cuantitativa podría obtenerse a partir de ARNm frecuentes y poco frecuentes monitorizando la tasa de crecimiento de las colonias.

(viii) Aislamiento y caracterización de nuevos genes expresados

[0234] Esta realización de nuestra invención se refiere a resolver el problema de aislar genes que están inducidos específicamente en condiciones determinadas, por ejemplo, en tejidos específicos, cepas diferentes de una especie determinada o bajo una activación específica. Un ejemplo práctico es la identificación de genes que se regulan positiva y negativamente después de la administración de un fármaco.

[0235] La realización preferida para el aislamiento de genes a partir de una muestra biológica específica o activada (en lo sucesivo en el presente documento denominada muestra diana) que se regula positivamente en comparación con una muestra biológica de referencia (en lo sucesivo en el presente documento denominada muestra de referencia) se representa en la figura 18.

[0236] En primer lugar, se preparan colonias primarias (figura 18a). En esta forma preferida, el ADN que se utiliza para esta preparación es ADN genómico preparado o ADN genómico etiquetado, como se describe en los apartados G(i) y G(ii), respectivamente, en el que el ADN es del genoma completo de uno (o varios) organismos o de un subconjunto de los mismos (por ejemplo de una biblioteca de genes previamente aislados), y en el que los dos cebadores utilizados para la generación de colonias (en lo sucesivo en el presente documento denominados P1 y P2) contienen un sitio de restricción de endonucleasas. En la figura 18a, "A" representa colonias que han surgido de genes expresados en la muestra tanto de referencia como diana, "B" representa colonias que surgen de genes expresados únicamente en la muestra de referencia, "C" representa colonias que han surgido de genes expresados únicamente en la muestra diana, y "D" representa colonias que surgen de genes no expresados (en casos reales, todas estas situaciones pueden no estar necesariamente presentes de manera simultánea).

[0237] En segundo lugar, las colonias primarias se añaden después para generar cebadores secundarios como el soporte para el crecimiento de la colonia secundaria (etapa i en la figura 18a). En esta fase (b) las colonias se representan como caracteres subrayados (A,B,C,D).

[0238] En tercer lugar, (etapa ii en la figura 21b) los cebadores secundarios se utilizan para regenerar colonias utilizando ARNm o ADNc (representado por "mA + mB") extraído de la muestra de referencia biológica como un molde, como se describe en G(v). Si el molde es ARNm, la primera etapa de alargamiento de la regeneración de colonias se realizará con una transcriptasa inversa. Después de suficientes ciclos de crecimiento de colonias, preferentemente de 5 a 100, solamente se regenerarán las colonias correspondientes a genes expresados en la muestra de referencia ("A" y "B"), como se representa en la (figura 18c).

[0239] En la etapa (iii) las colonias se digieren con una enzima de restricción (representada por RE) que reconoce un sitio en las secuencias cebadoras flanqueantes, P1 y P2, que se injertan sobre el soporte y que son la base de la generación de colonias primarias. Hay que destacar que, únicamente las colonias que se han regenerado durante la etapa (ii) se someterán a digestión. Esto es porque el soporte para el crecimiento de las colonias secundarias se realiza en moléculas de ADN monocatenario, que no pueden digerirse por la enzima de restricción. Únicamente las colonias regeneradas están presentes en una forma bicatenaria, y se someten a digestión. Después de la digestión, la situación es la que se representa en la figura 18d. Las colonias correspondientes a los genes expresados en la muestra de referencia han desaparecido totalmente, es decir, ya no están presentes como un soporte para el crecimiento de colonias secundarias, y las colonias correspondientes a los genes que se expresan únicamente en la muestra diana "C" y las colonias correspondientes a genes no expresados "D" aún están presentes como un soporte para la generación de colonias secundarias.

[0240] En la etapa (iv), el ARNm (o ADNc) (representado por "mA + mC") extraído de la muestra diana, se usa para generar colonias secundarias. Dado que las colonias correspondientes a mA y mB ya no existen, únicamente las colonias correspondientes a mC pueden regenerarse (es decir únicamente el ARNm se expresa específicamente en la muestra diana). Después de suficientes ciclos de crecimiento de colonias (preferentemente de 5 a 100) la situación es tal que únicamente se regeneran las colonias correspondientes a los genes expresados específicamente en la muestra diana ("C", figura 18e).

[0241] En la etapa (v), las colonias regeneradas "C" se usan para generar copias del ADN que contienen mediante la realización de diversos ciclos de crecimiento de colonias (preferentemente de 1 a 20) en presencia de los cebadores P1 y P2, como se describe en el apartado D de la presente invención. Después se realiza una amplificación por PCR utilizando P1 y P2 en solución (descrito en el apartado D) y el ADN amplificado se caracteriza mediante métodos clásicos.

[0242] La realización preferida para aislar genes de una muestra biológica específica o activada que se expresan menos que en una muestra biológica de referencia se representa en la figura 19. Las diferentes etapas implicadas en este procedimiento son muy similares a las implicadas en el aislamiento de genes que están más regulados que en la muestra de referencia, y la notación es la misma que la de la figura 18. La única diferencia es invertir el orden usado para regenerar las colonias: en la etapa (ii), el ARNm que se utiliza es el extraído de la muestra biológica diana ("mA + mC") en lugar del ARNm extraído de la muestra biológica de referencia ("mA + mB"), y en la etapa (iv), el ARNm que se utiliza es el extraído de la muestra biológica de referencia ("mA + mB") en lugar del extraído de la muestra diana ("mA + mC"). Como resultado, únicamente se recupera y amplifica el ADN procedente de las colonias correspondientes a los genes que se expresan en la muestra de referencia, pero no en la muestra diana ("B", figura 19f).

[0243] Las realizaciones preferidas se indican en los siguientes párrafos:

20 **Realizaciones**

[0244]

1. Un método de secuenciación de moléculas de ácido nucleico amplificadas, inmovilizadas, que comprende:
 - (a) permitir que los cebadores se hibriden con las moléculas de ácido nucleico amplificadas;
 - (b) extender los cebadores por adición de un nucleótido; y
 - (c) detectar el nucleótido utilizado en la extensión del cebador.
2. Un método de acuerdo con el párrafo 1 en el que el nucleótido añadido la etapa (b) se marca y la etapa (c) comprende detectar si el nucleótido marcado se ha utilizado o no para la extensión del cebador determinando si el marcador presente en dicho nucleótido se ha incorporado o no en los cebadores extendidos.
3. Un método de acuerdo con el párrafo 1 o con el párrafo 2 en el que la etapa (b) comprende proporcionar a las moléculas de ácido nucleico amplificadas una polimerasa de ácido nucleico y un nucleótido determinado en forma marcada y no marcada o en forma marcada solo en condiciones que permitan la extensión de los cebadores si una base complementaria o si una pluralidad de dichas bases está presente en la posición apropiada en las moléculas de ácido nucleico amplificadas.
4. Un método de acuerdo con cualquier párrafo anterior en el que las moléculas de ácido nucleico amplificadas se proporcionan en forma de al menos una colonia de ácido nucleico, comprendiendo la colonia una pluralidad de moléculas de ácido nucleico monocatenario que tienen la misma secuencia entre sí y que se hibridan con cebadores de manera que se permite la extensión de cebadores en presencia del nucleótidos.
5. Un método de acuerdo con el párrafo 4 en el que las moléculas de ácido nucleico amplificadas se proporcionan como una pluralidad de colonias comprendiendo cada una de ellas una pluralidad de moléculas de ácido nucleico monocatenario que tienen la misma secuencia entre sí y que se hibridan con cebadores de una manera que se permite la extensión de cebadores en presencia de nucleótidos.
6. Un método de acuerdo con el párrafo 4 o con el párrafo 5 en el que la etapa (b) comprende proporcionar al menos una colonia con una polimerasa de ácido nucleico y un nucleótido determinado en forma marcada y no marcada o en forma marcada solo en condiciones que permitan la extensión de los cebadores si una base complementaria o si una pluralidad de dichas bases está presente en una posición apropiada en las moléculas de ácido nucleico monocatenario presentes en la dicha al menos una colonia y la etapa (c) comprende detectar si el nucleótido marcado se ha utilizado o no para la extensión de cebadores determinando si el marcador presente en dicho nucleótido se ha incorporado o no en los cebadores extendidos.
7. Un método para la secuenciación de moléculas de ácido nucleico presentes en una colonia que comprende una pluralidad de cadenas de ácido nucleico inmovilizadas, comprendiendo el método las etapas de:
 - (a) proporcionar al menos una colonia que comprende una pluralidad de moléculas de ácido nucleico monocatenario que tienen la misma secuencia entre sí y que se hibridan con cebadores de una manera que se permite la extensión de cebadores en presencia de nucleótidos y una polimerasa de ácido nucleico;
 - (b) proporcionar dicha al menos una colonia con una polimerasa de ácido nucleico y un nucleótido determinado en forma marcada y no marcada en condiciones que permitan la extensión de los cebadores si una base complementaria o si una pluralidad de dichas bases está presente en la posición apropiada en las moléculas de ácido nucleico monocatenario presentes en la dicha al menos una colonia;
 - (c) detectar si dicho nucleótido marcado se ha utilizado o no para la extensión de cebadores determinando si el marcador presente en dicho nucleótido se ha incorporado o no en los cebadores extendidos.

8. Un método de acuerdo con el párrafo 7 en el que en la etapa (a) se proporciona una pluralidad de colonias, comprendiendo cada una de ellas una pluralidad de moléculas de ácido nucleico monocatenario que tienen la misma secuencia entre sí y que se hibridan con cebadores de una manera que se permite la extensión de cebadores en presencia de nucleótidos.
- 5 9. Un método de acuerdo con el párrafo 7 o párrafo 8 en el que las moléculas de ácido nucleico monocatenario están presentes en forma amplificada en al menos una colonia.
- 10 10. Un método de acuerdo con cualquier párrafo anterior en el que las etapas (b) y (c) se repiten una o más veces.
11. Un método de acuerdo con cualquier párrafo anterior en el que el marcador unido a cualquier nucleótido añadido por extensión de cebador en la parte (b) y detectado en la etapa (c) se retira después de la etapa (c).
12. Un método de acuerdo con el párrafo 11 en el que las etapas (b) y (c) se repiten una o más veces y el marcador unido a cualquier nucleótido añadido por extensión de cebador en la parte (b) y detectado en la etapa (c) se retira después de cada etapa (c).
- 15 13. Un método de acuerdo con cualquiera de los párrafos 4 a 8 en el que dichas moléculas de ácido nucleico monocatenario y dichos cebadores están inmovilizados.
14. Un método de acuerdo con el párrafo 5 o párrafo 8 que se usa para secuenciar, completa o parcialmente, de manera simultánea, las moléculas de ácido nucleico amplificadas presentes en al menos 2 colonias diferentes que tienen diferentes moléculas de ácido nucleico.
- 20 15. Un método de acuerdo con el párrafo 14 que se usa para secuenciar, completa o parcialmente, de manera simultánea, las moléculas de ácido nucleico presentes en 10 o más colonias diferentes que tienen diferentes moléculas de ácido nucleico.
16. Un método de acuerdo con el párrafo 14 que se usa para secuenciar, completa o parcialmente, de manera simultánea, las moléculas de ácido nucleico presentes en 100 o más colonias diferentes que tienen diferentes moléculas de ácido nucleico.
- 25 17. Un método de acuerdo con el párrafo 14 que se usa para secuenciar, completa o parcialmente, de manera simultánea, las moléculas de ácido nucleico presentes en 1000 o más colonias diferentes que tienen diferentes moléculas de ácido nucleico .
18. Un método de acuerdo con el párrafo 14 que se usa para secuenciar, completa o parcialmente, de manera simultánea, las moléculas de ácido nucleico presentes en 1.000.000 o más colonias diferentes que tienen diferentes moléculas de ácido nucleico.
- 30 19. Un método de acuerdo con cualquier párrafo anterior, en el que después de la etapa (b) el exceso de nucleótidos que no se ha usado en la extensión de cebador se retira.
20. Un método de acuerdo con cualquier párrafo anterior, en el que la etapa (c) utiliza una cámara de dispositivo de carga acoplada, que se acopla a un dispositivo de aumento.
- 35 21. Un método de acuerdo con cualquier párrafo anterior, en el que cada uno de los cuatro nucleótidos diferentes se usa en la extensión de cebadores.
22. Un método de acuerdo con el párrafo 21, en el que dichos cuatro nucleótidos diferentes se utilizan en un orden predeterminado en ciclos repetidos.
- 40 23. Un método de acuerdo con el párrafo 21 o párrafo 22, en el que los nucleótidos son dATP, dTTP, dGTP y dCTP en forma marcada.
24. Un método de acuerdo con el párrafo 21 o párrafo 22, en el que los nucleótidos son ATP, UTP, GTP y CTP en forma marcada.
25. Un método como se describe en cualquier párrafo anterior con la excepción de que las moléculas de ácido nucleico bicatenario que tienen muescas en su interior se proporcionan en la etapa (a) en vez de moléculas de ácido nucleico que se hibridan con cebadores.
- 45 26. El método de acuerdo con cualquier párrafo anterior en el que las moléculas de ácido nucleico amplificadas o las moléculas de ácido nucleico monocatenario son ADN o ARN.
27. Un método de acuerdo con cualquier párrafo anterior en el que las moléculas de ácido nucleico amplificadas o las moléculas de ácido nucleico monocatenario y/o los cebadores comprenden bases de origen natural y/o no natural.
- 50 28. Un método de acuerdo con uno cualquiera de los párrafos 1 a 27 en el que las moléculas de ácido nucleico amplificadas o las moléculas de ácido nucleico monocatenario se secuencian para determinar si en las moléculas de ácido nucleico monocatenario hay una o más diferencias de secuencia con respecto a una secuencia de referencia conocida.
- 55 29. Un método de acuerdo con el párrafo 28 para su uso en la identificación de polimorfismos o mutaciones en las moléculas de ácido nucleico amplificado o en las moléculas de ácido nucleico monocatenario.

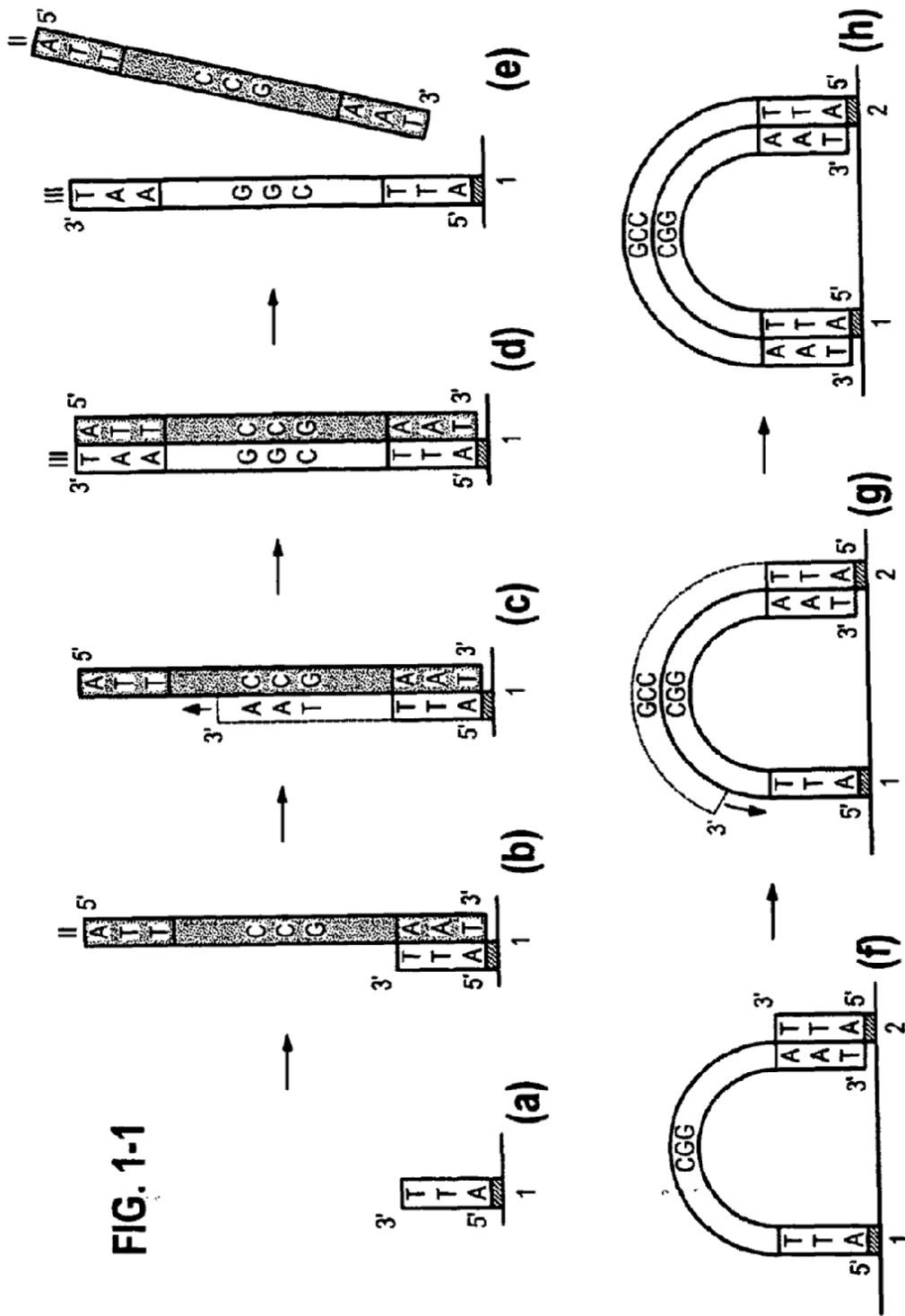
REIVINDICACIONES

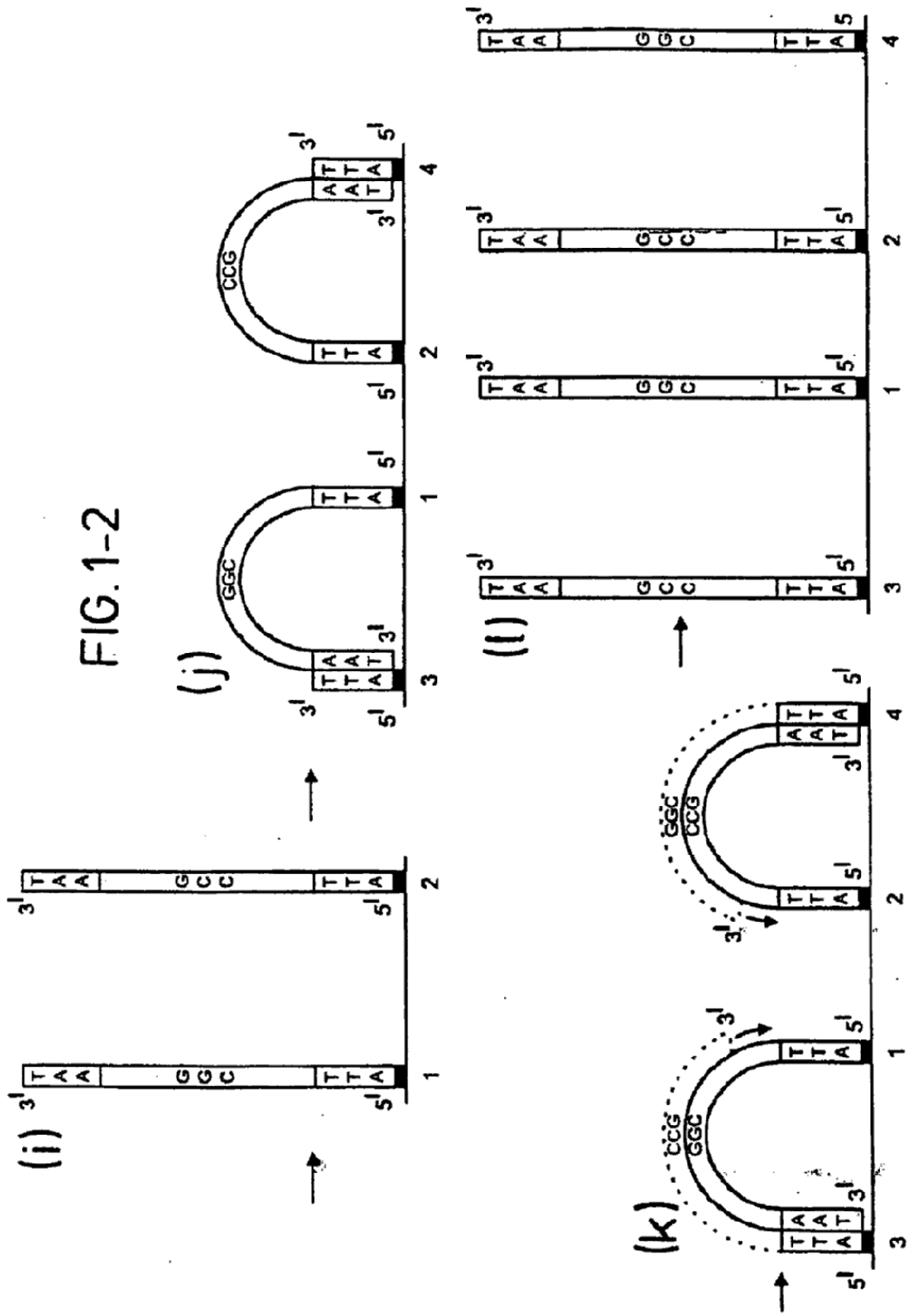
1. Un método de secuenciación de una molécula de ácido nucleico que comprende:
- 5 a. utilizar una primera colonia para proporcionar una pluralidad de moléculas de ácido nucleico monocatenario que tienen la misma secuencia y que hibridan con cebadores de una manera que permite la extensión de cebadores;
- b. utilizar una segunda colonia para proporcionar una pluralidad de moléculas de ácido nucleico monocatenario que tienen la misma secuencia y que hibridan con cebadores de una manera que permite la extensión de
- 10 cebadores;
- c. proporcionar a cada colonia una polimerasa de ácido nucleico y un nucleótido marcado en condiciones que permitan la extensión de los cebadores si una base o una pluralidad de bases complementarias están presentes en la posición apropiada en las moléculas de ácido nucleico monocatenario en las colonias;
- d. detectar si el nucleótido marcado se ha incorporado o no en los cebadores extendidos; y
- 15 e. repetir las etapas c. y d. una o más veces de manera que se proporcionen cebadores extendidos que comprendan una pluralidad de marcadores.
2. El método de la reivindicación 1, en el que las moléculas de ácido nucleico monocatenario comprenden una primera parte que tiene una secuencia que puede aparearse con un primer cebador de amplificación y una segunda
- 20 parte que tiene una secuencia complementaria a una secuencia que puede aparearse con un segundo cebador de amplificación y, opcionalmente, una tercera parte que comprende una secuencia desconocida, estando la tercera parte situada entre la primera parte y la segunda parte.
3. El método de la reivindicación 2, en el que las colonias se forman por amplificación de moléculas de ácido
- 25 nucleico monocatenario en una superficie sólida; o
- en el que las colonias se forman proporcionando una superficie sólida, apareando moléculas de ácido nucleico monocatenario en la superficie sólida, y amplificando las moléculas de ácido nucleico monocatenario.
4. El método de la reivindicación 1, en el que las secuencias de las moléculas de ácido nucleico presentes en la
- 30 primera y segunda colonias son diferentes entre sí.
5. El método de la reivindicación 1, en el que se secuencian una pluralidad de primera y segunda colonias que comprenden diferentes moléculas de ácido nucleico.
- 35 6. El método de la reivindicación 5, en el que las colonias se monitorizan mediante un detector en una posición fija y las colonias que van a analizarse se desplazan en el campo visual del detector; o
- en el que las colonias se monitorizan manteniendo las colonias en una posición fija y desplazando el detector para llevar las colonias al campo visual del detector.
- 40 7. El método de la reivindicación 1, en el que los nucleótidos marcados se detectan mediante un sistema de detección.
8. El método de la reivindicación 6, en el que el sistema de detección detecta marcadores fluorescentes.
- 45 9. El método de la reivindicación 6, en el que el sistema de detección es una cámara de dispositivo de carga acoplada (CCD, *Charge Coupled Device*), opcionalmente acoplada a un dispositivo de aumento.
10. El método de la reivindicación 6, en el que el sistema de detección se utiliza en combinación con un sistema de análisis para determinar el número y tipo de nucleótidos incorporados por extensión de cebadores en cada colonia
- 50 después de cada etapa.
11. El método de la reivindicación 1, en el que la secuencia de las moléculas de ácido nucleico en cada colonia se deduce a partir del número y tipo de nucleótidos añadidos después de cada etapa.
- 55 12. Una superficie que comprende distintas pluralidades de moléculas de ácido nucleico inmovilizadas en forma de áreas distintas, comprendiendo cada área una pluralidad de cadenas de ácido nucleico idénticas y una pluralidad de cadenas complementarias idénticas a las mismas; en la que cada cadena de ácido nucleico dentro de dicha área se localiza de tal manera que otra cadena de ácido nucleico se localiza en la superficie a una distancia de la longitud de esa cadena; en la que diferentes áreas están constituidas por diferentes cadenas de ácido nucleico amplificadas y
- 60 por cadenas complementarias a las mismas amplificadas, pero en la que se proporcionan secuencias idénticas en el primer y segundo extremos de cada dichas cadenas de ácido nucleico en cada área y se proporcionan secuencias idénticas en el primer y segundo extremos de dichas cadenas de ácido nucleico complementarias en cada área; y en la que cada molécula de ácido nucleico amplificado inmovilizado comprende una primera parte que tiene una secuencia que puede aparearse con un primer cebador de amplificación, una segunda parte que tiene una
- 65 secuencia complementaria a una secuencia que puede aparearse con un segundo cebador de amplificación y una

tercera parte que comprende una secuencia desconocida localizada entre la primera y segunda partes.

13. La superficie de la reivindicación 12, en la que las moléculas de ácido nucleico amplificadas, inmovilizadas son moléculas de ácido nucleico bicatenario que no tienen puente.

5





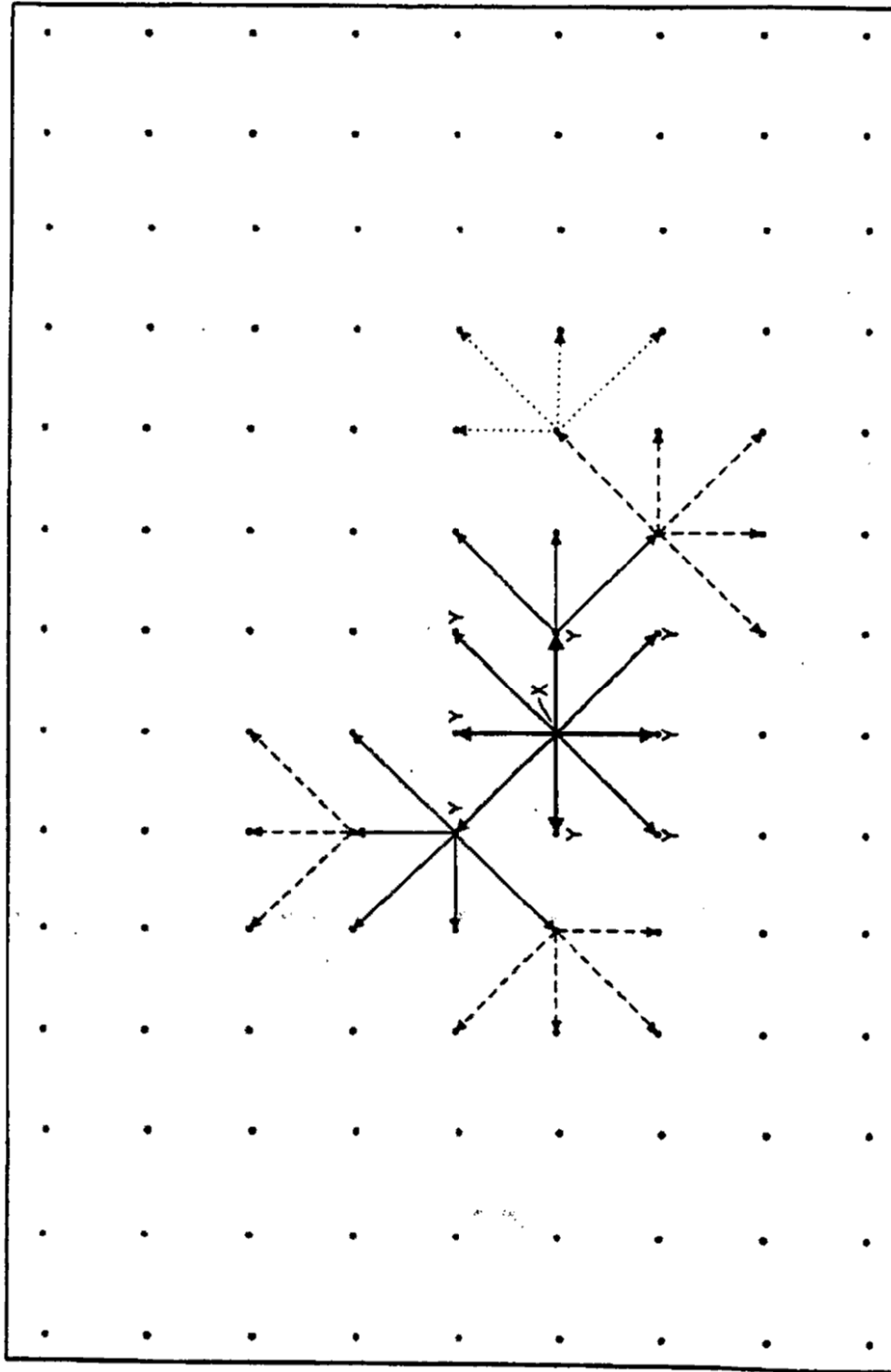
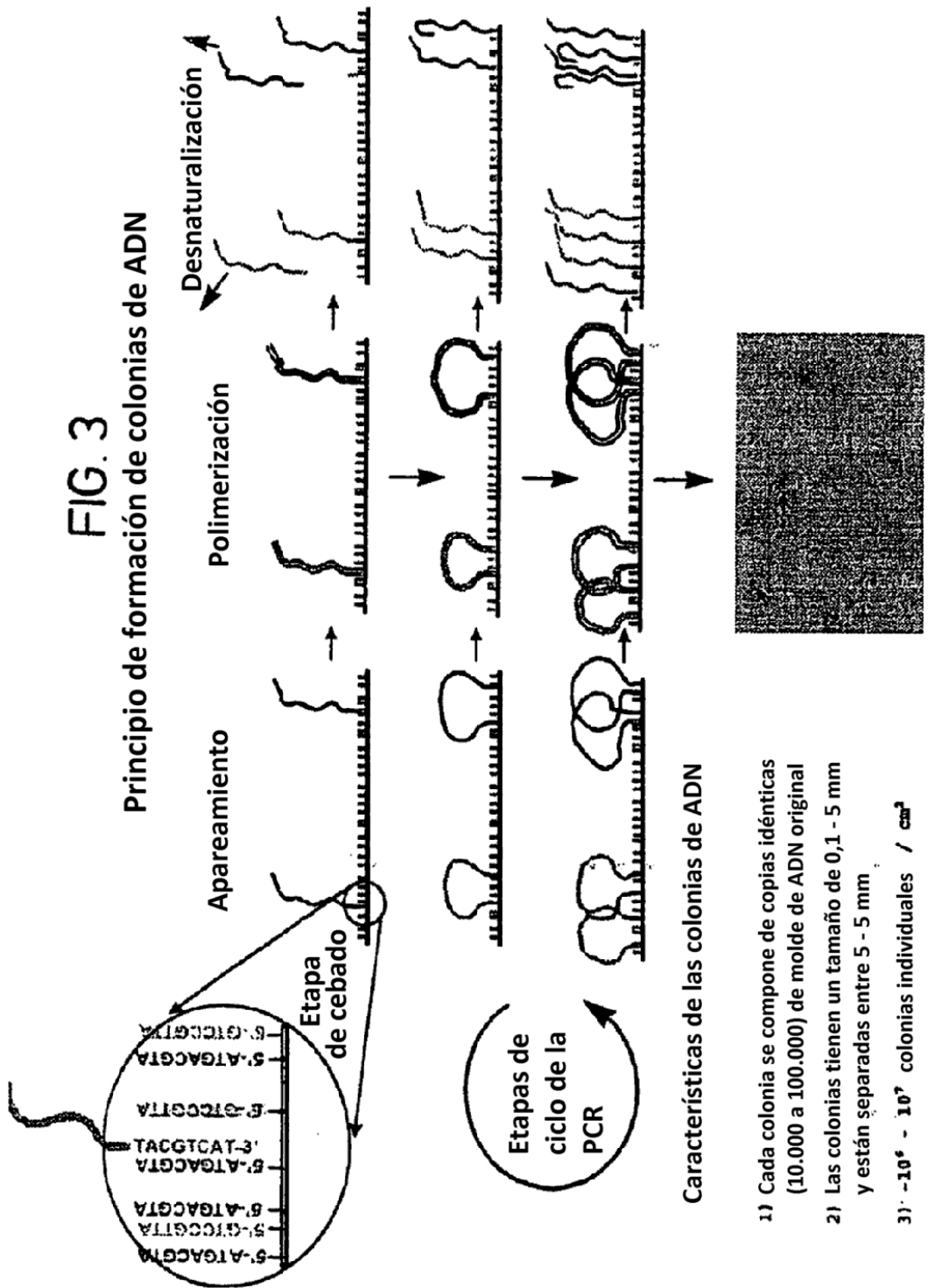


FIG. 2 MUESTRA UNA VISTA ESQUEMÁTICA DEL CRECIMIENTO DE LAS COLONIAS



- 1) Cada colonia se compone de copias idénticas (10.000 a 100.000) de molde de ADN original
- 2) Las colonias tienen un tamaño de 0,1 - 5 mm y están separadas entre 5 - 5 mm
- 3) $10^5 - 10^7$ colonias individuales / cm^2

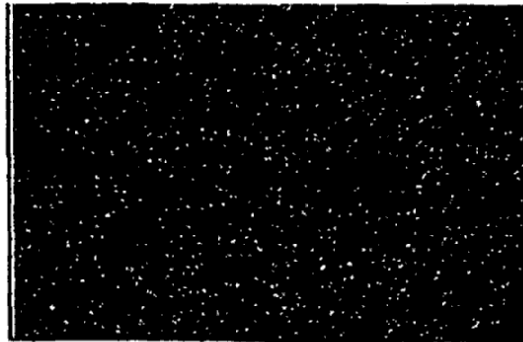
FIG. 4a

Generación de colonias en tubos funcionalizados con el oligonucleótido $\rho 57$ (control).
Resultados de visualización de microesferas fluorescentes. La hibridación con sonda detecta
0,10 fmol/tubo.



FIG. 4b

Generación de colonias en tubos funcionalizados con el oligonucleótido $\rho 58$ (control).
Resultados de visualización de microesferas fluorescentes. La hibridación con sonda detecta
6,2 fmol/tubo.



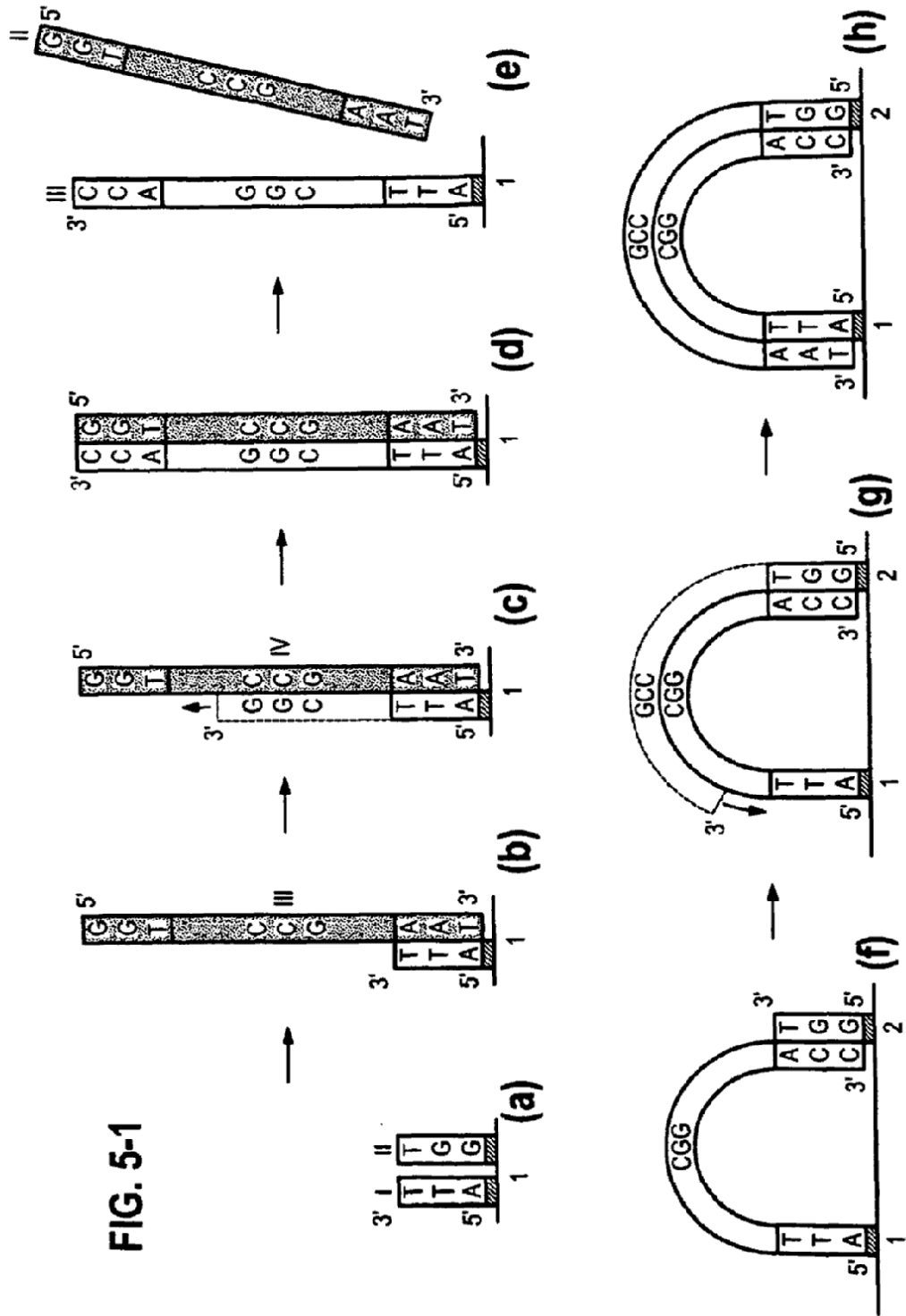


FIG. 5-1

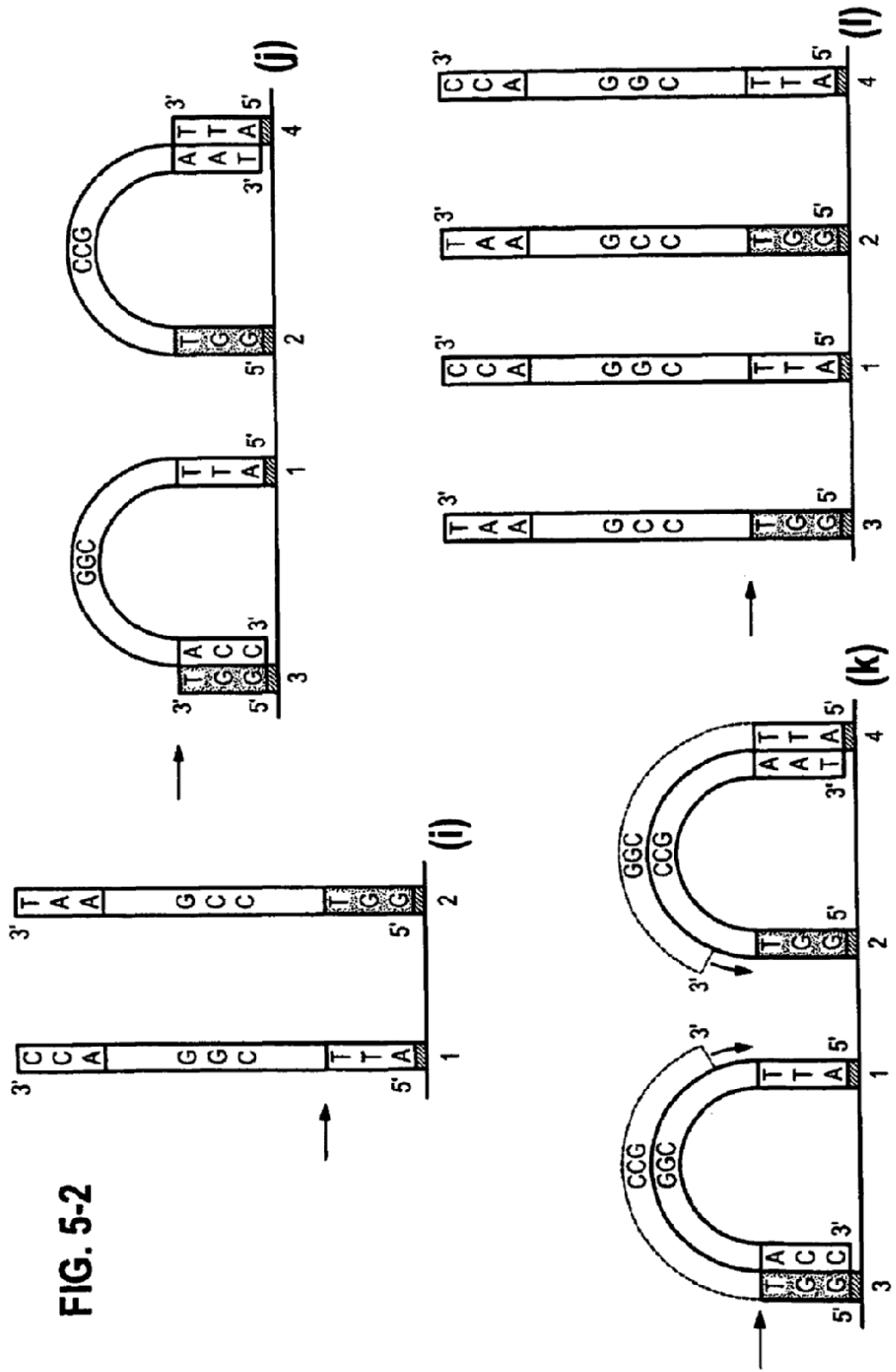


FIG. 6(a)

Amplificación de ADN - detección de colonias con colorantes intercalantes de ADN

i) control



ii) con molde

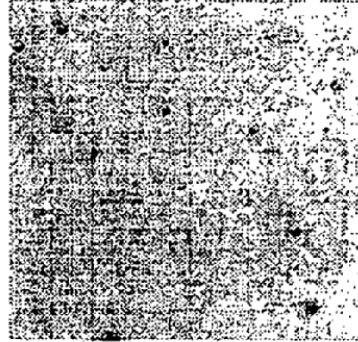


FIG. 6(b)

ADN amplificación: detección de colonias por hibridación con sondas específicas

i) proporción de molde S1/S2 = 1/0



ii) proporción de molde S1/S2 = 1/10



iii) proporción de molde S1/S2 = 0/1



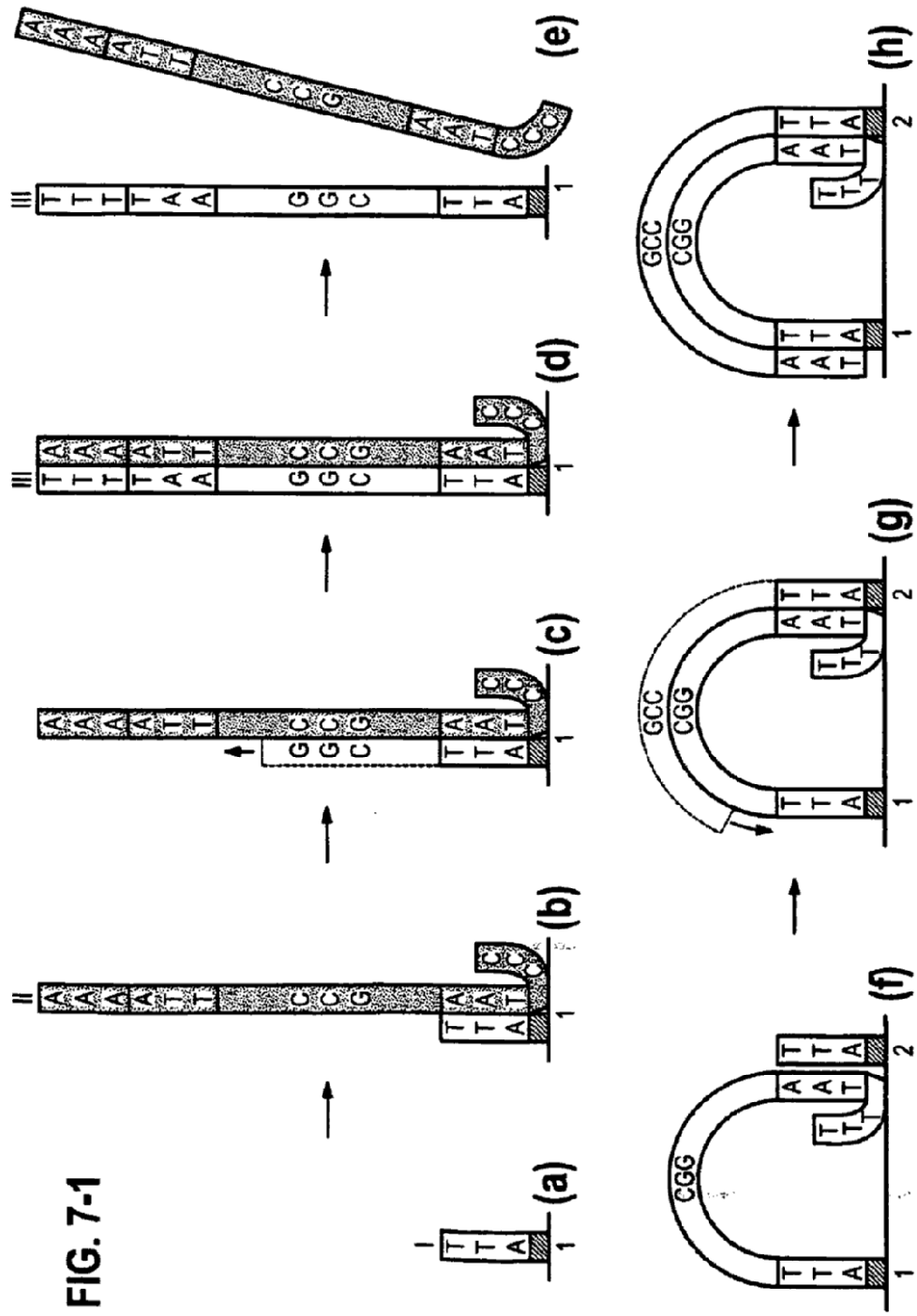
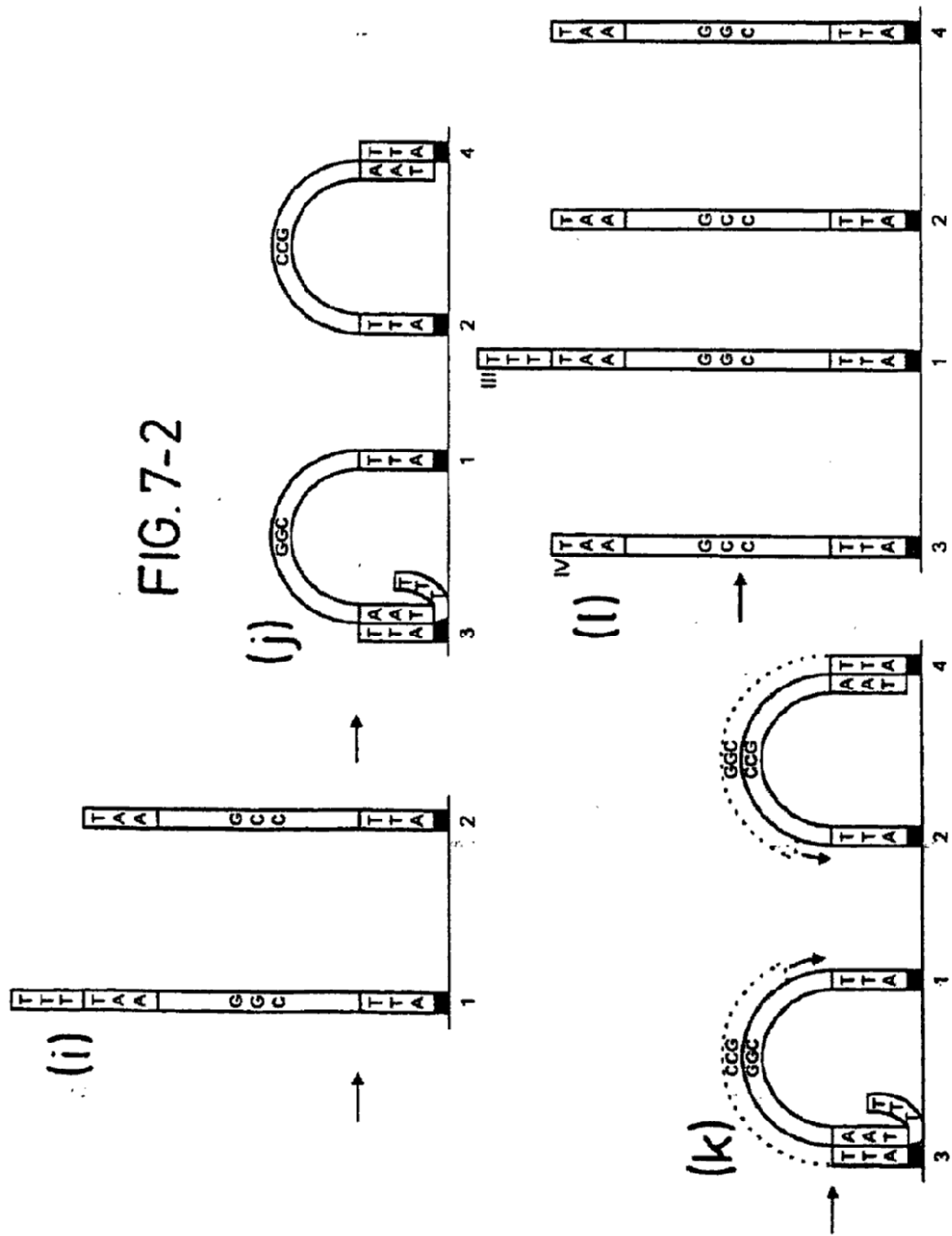


FIG. 7-1



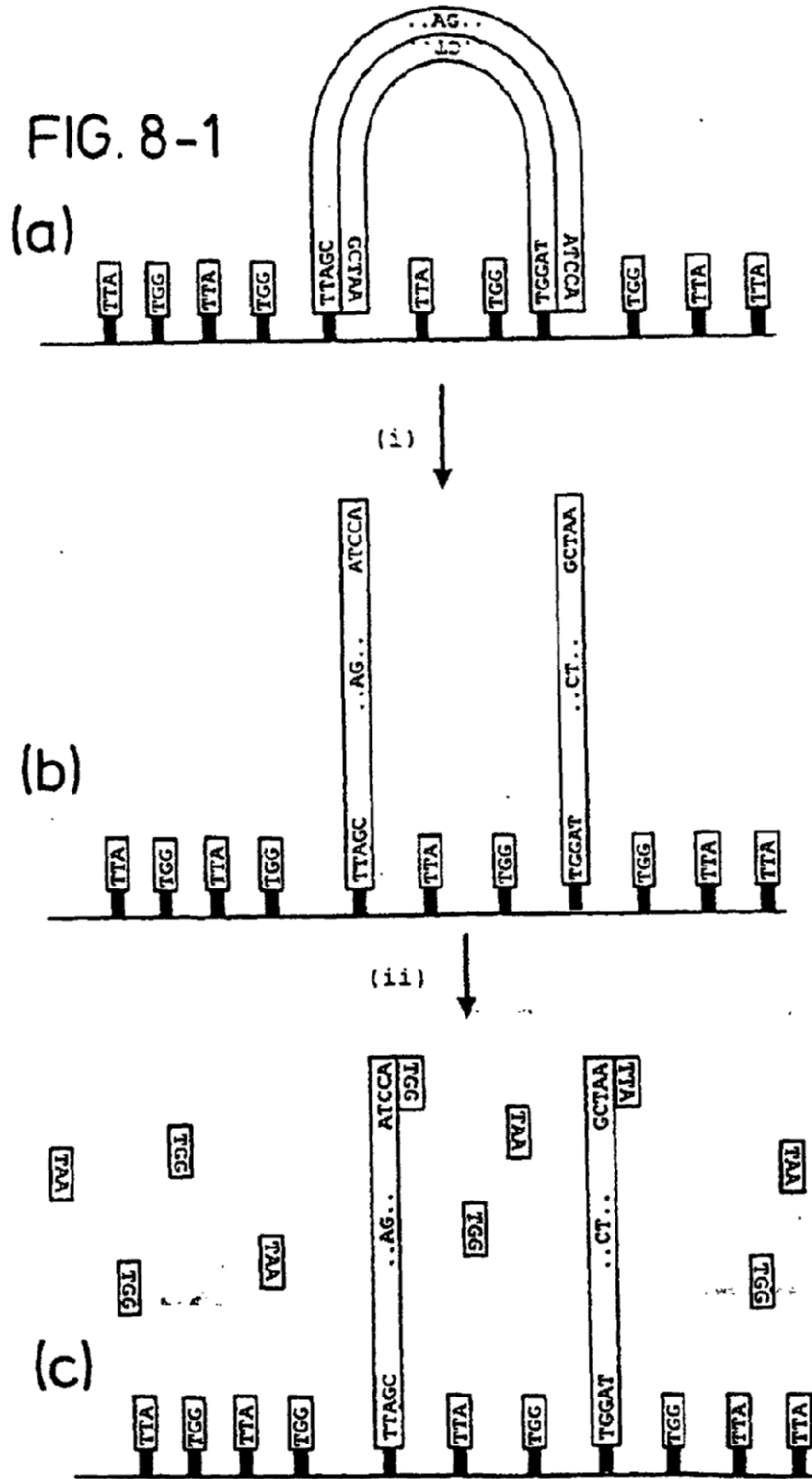


FIG. 8-2

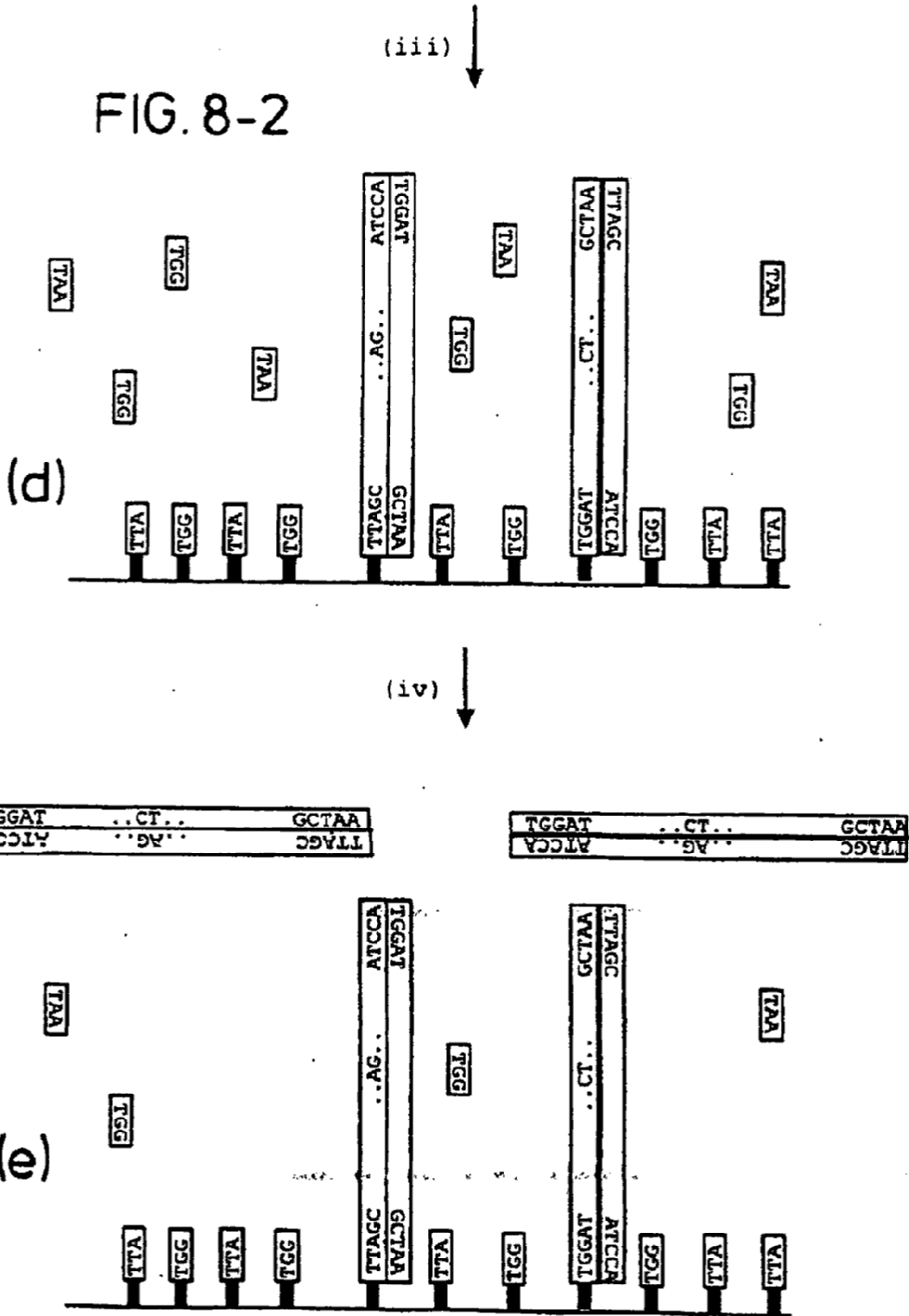
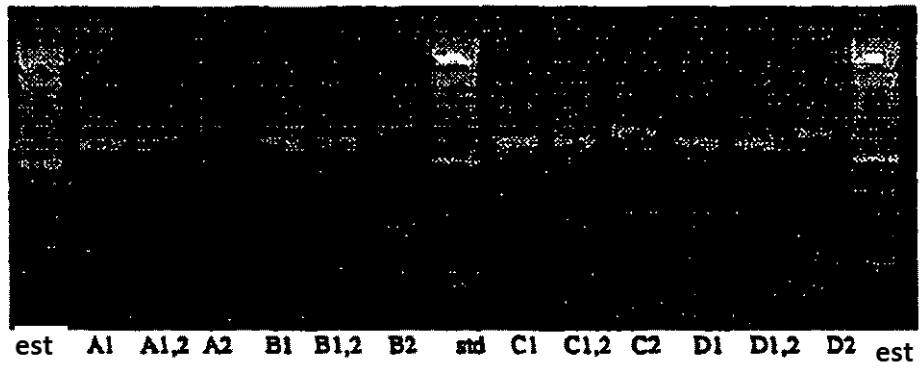


FIG. 9

A)



B)

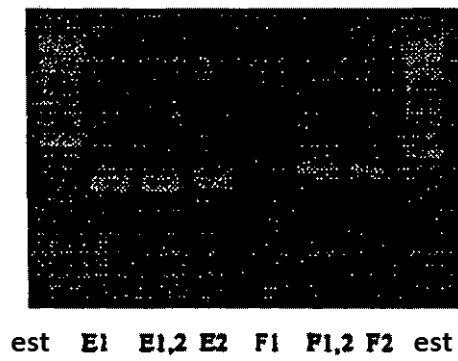
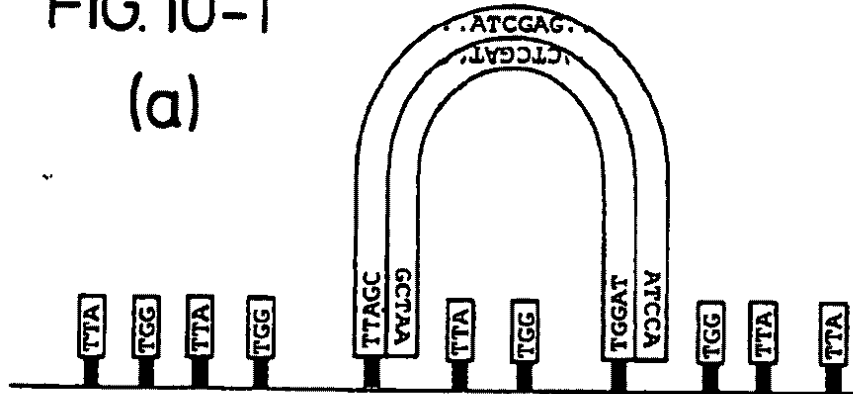


FIG. 10-1
(a)



(i)



(b)

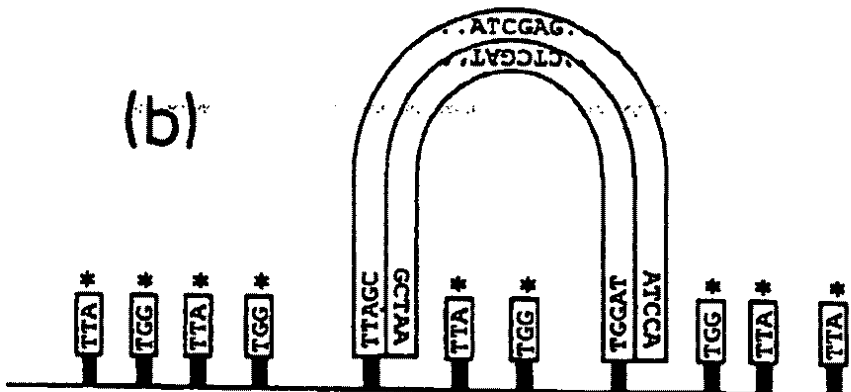
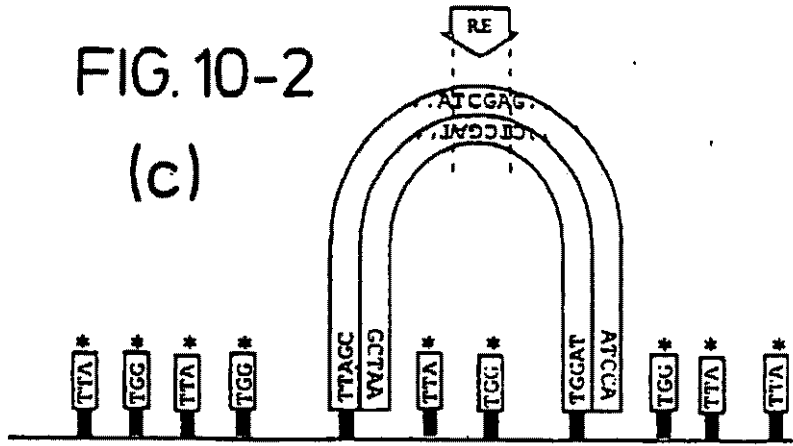


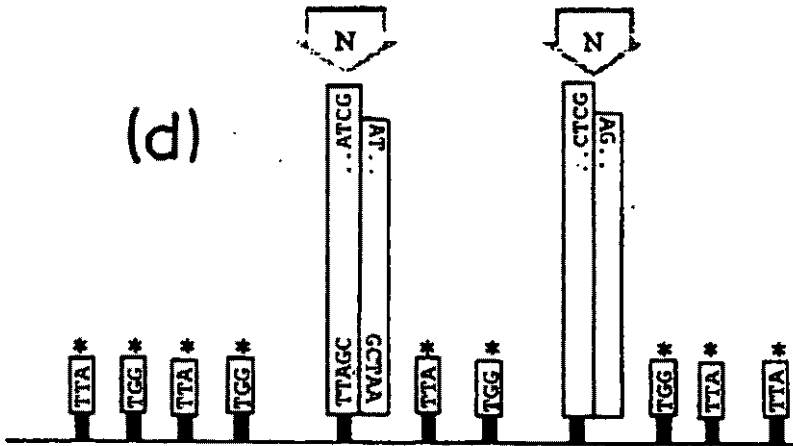
FIG. 10-2

(c)



(i)

(d)



(iii)

(e)

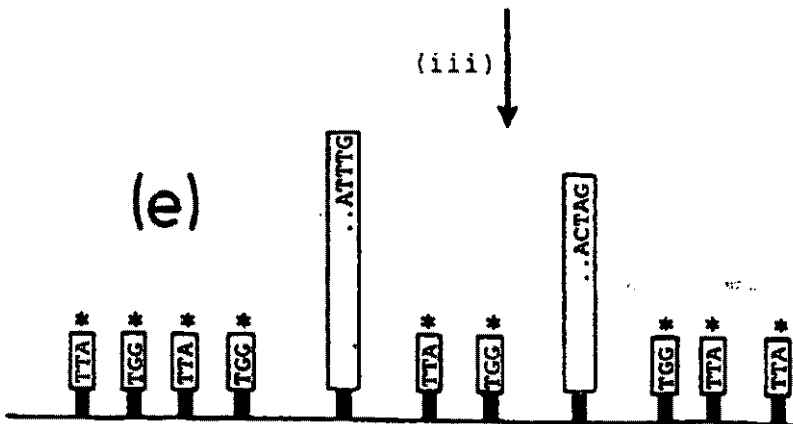
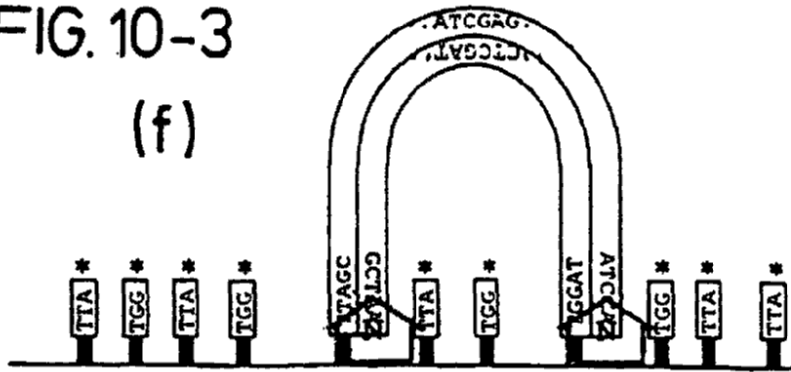


FIG. 10-3

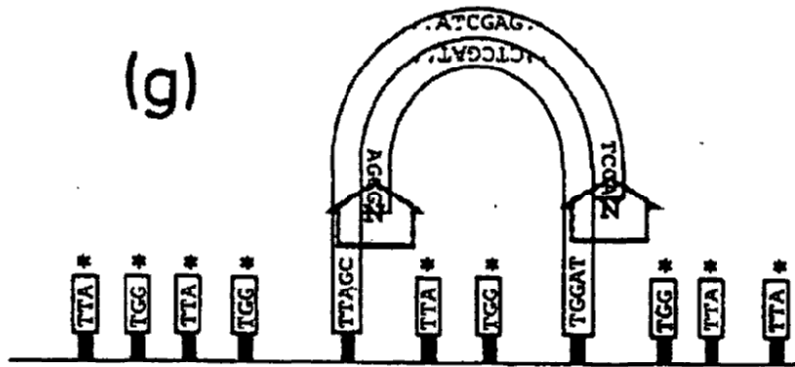
(f)



(iv)



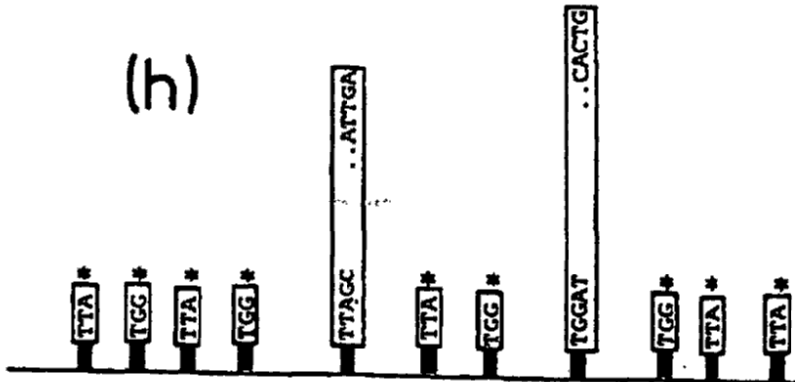
(g)



(v)



(h)



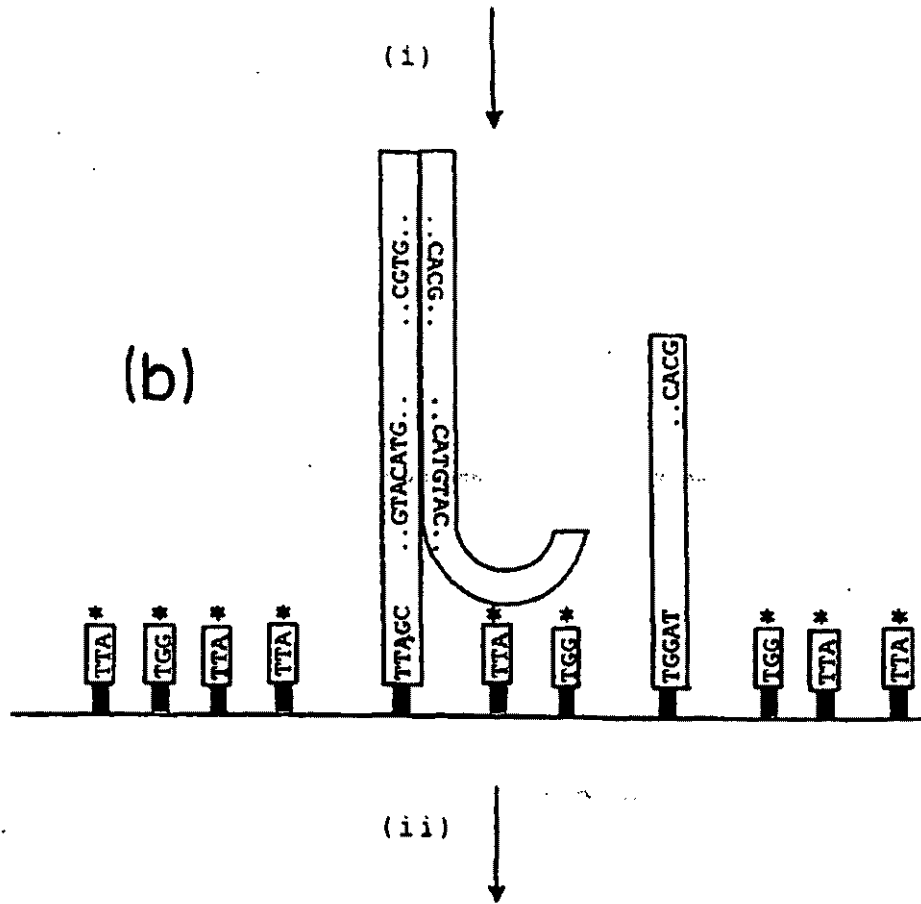
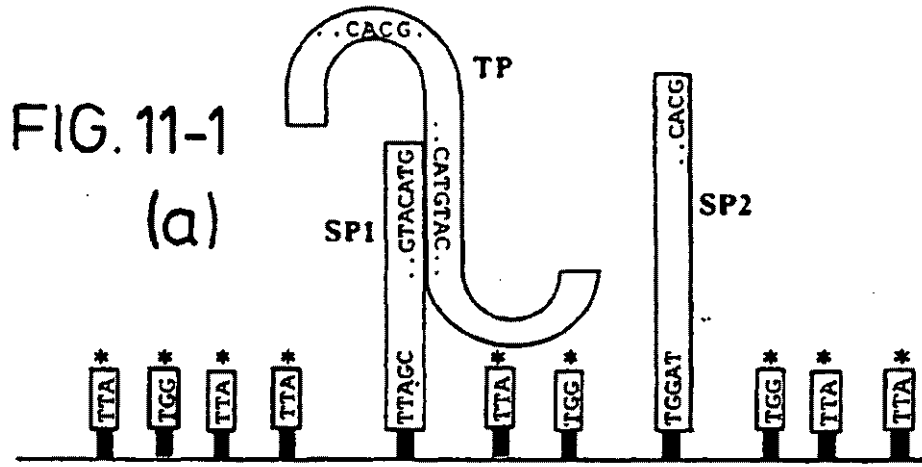


FIG. 11-2

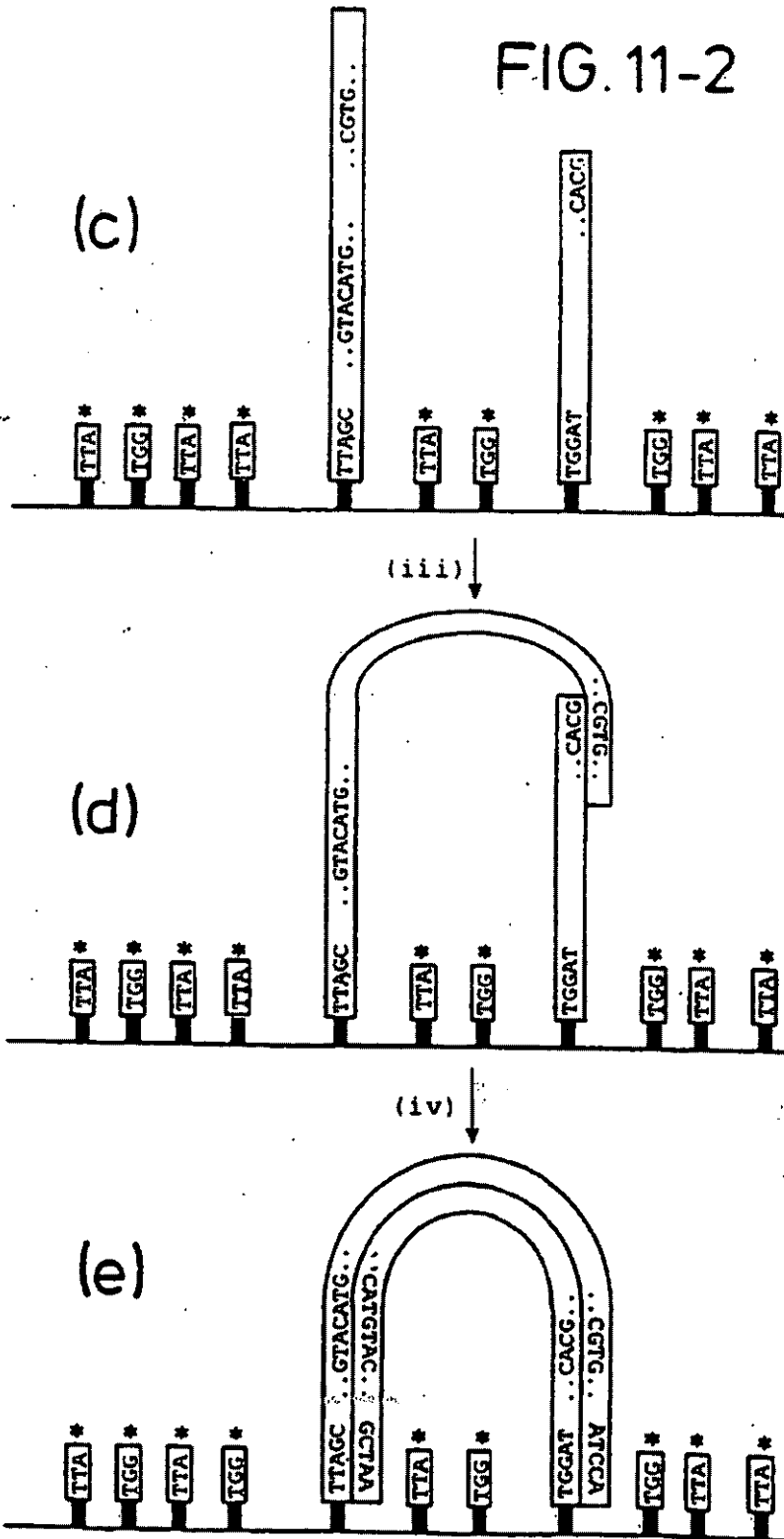
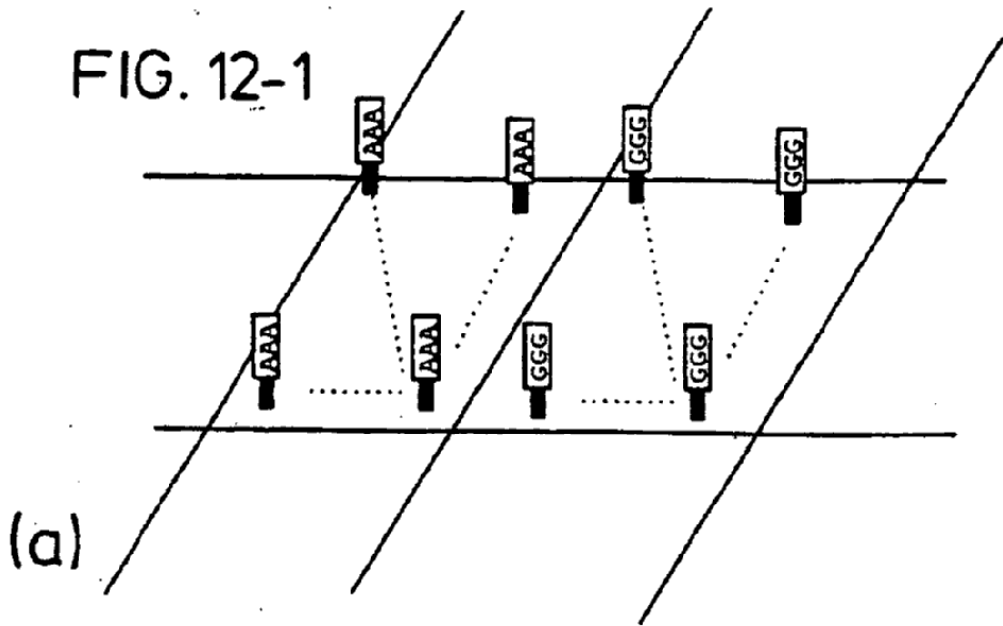
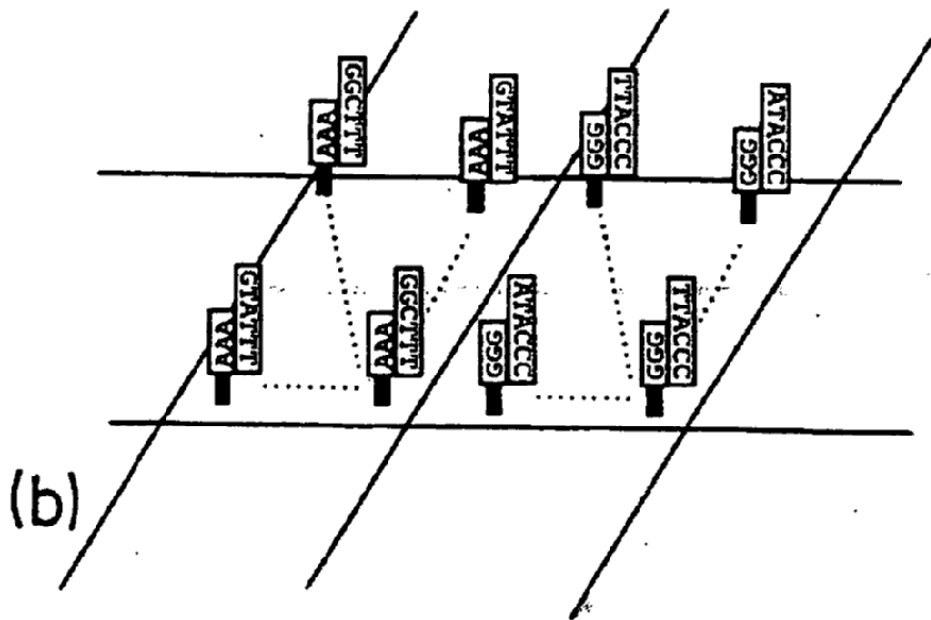


FIG. 12-1

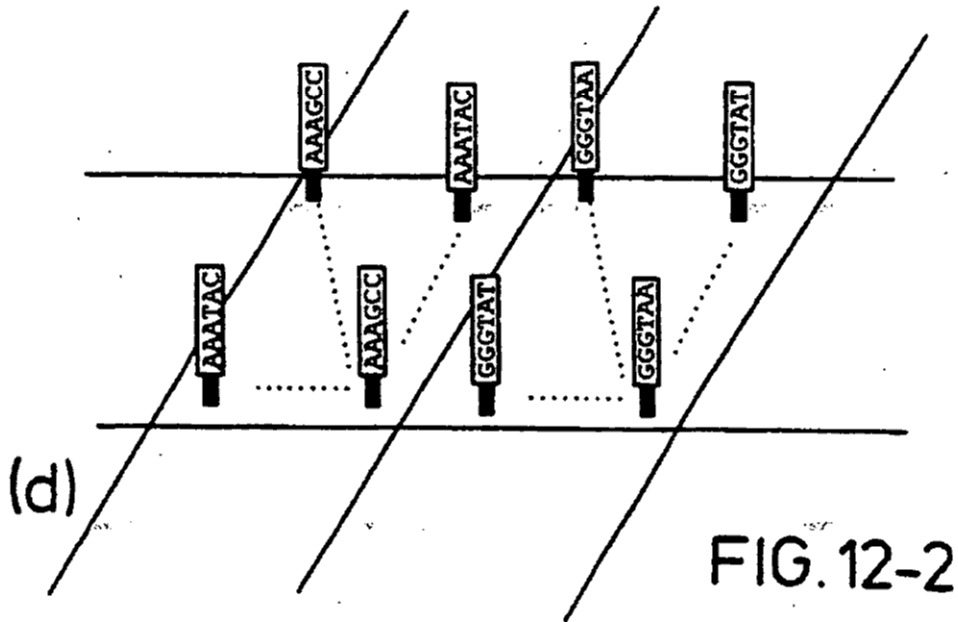
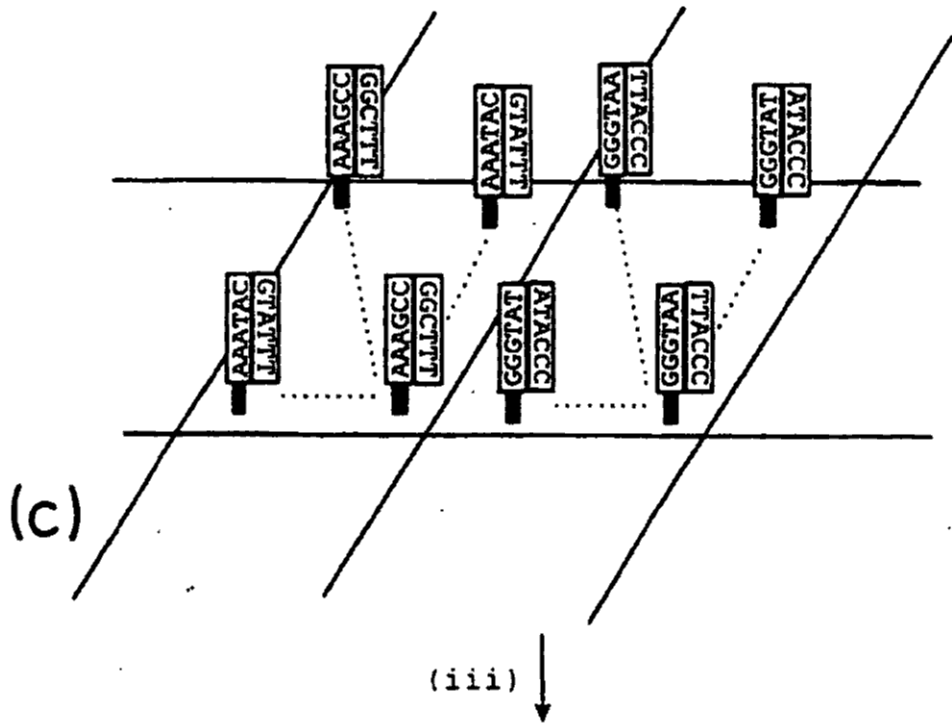


(i)



(ii)





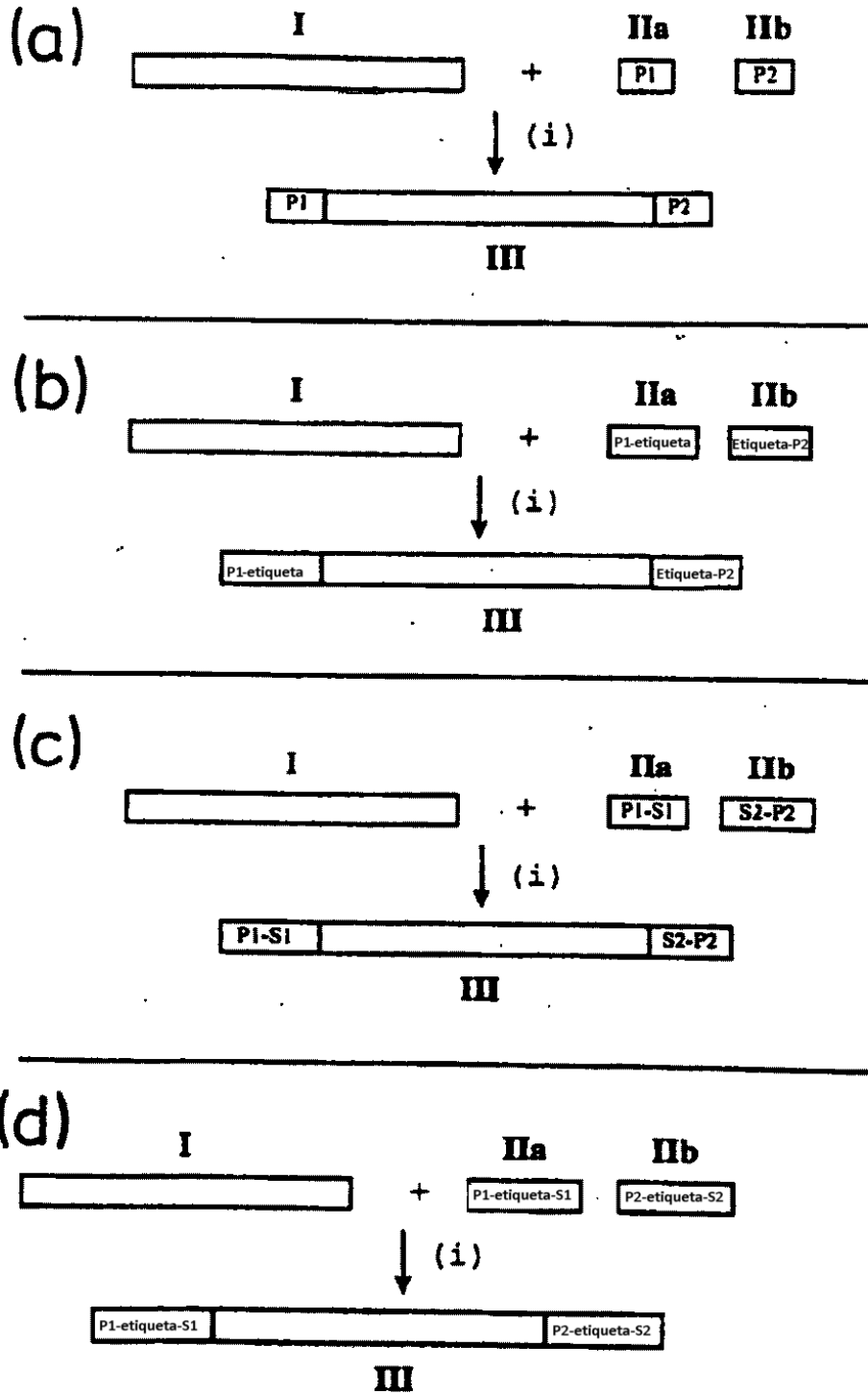


FIG. 13

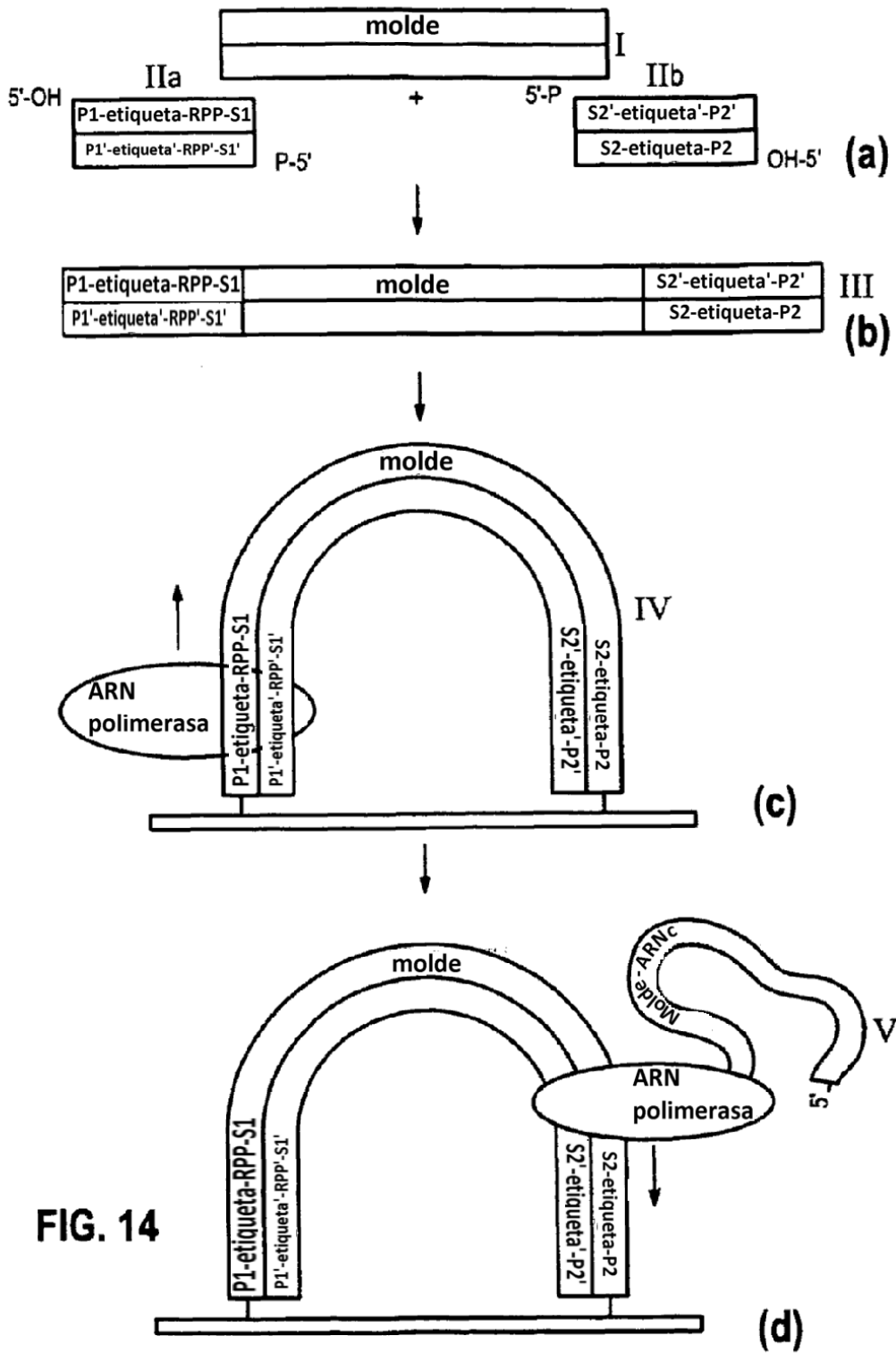
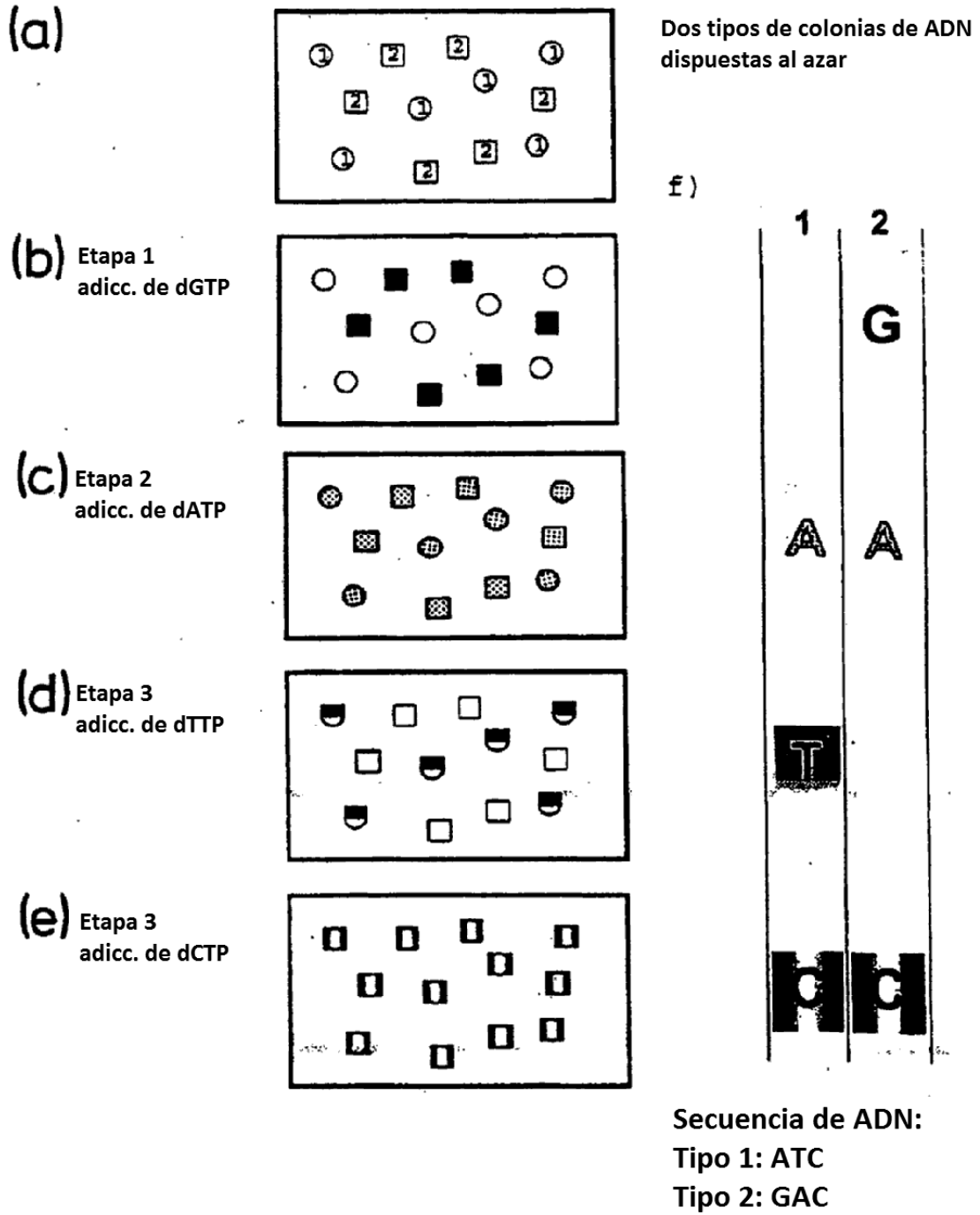


FIG. 14

FIG. 15

Secuencia *in situ*



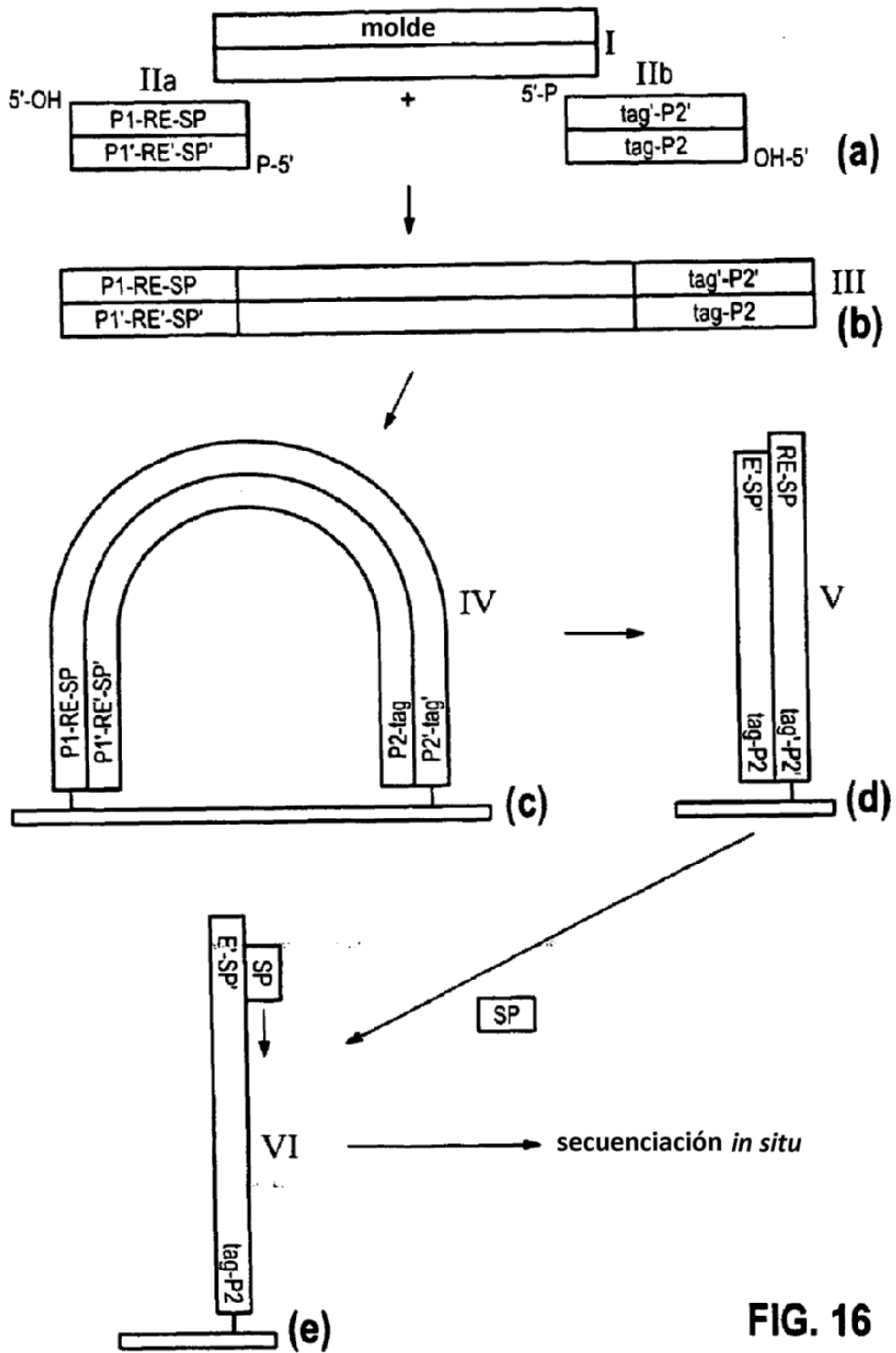


FIG. 16

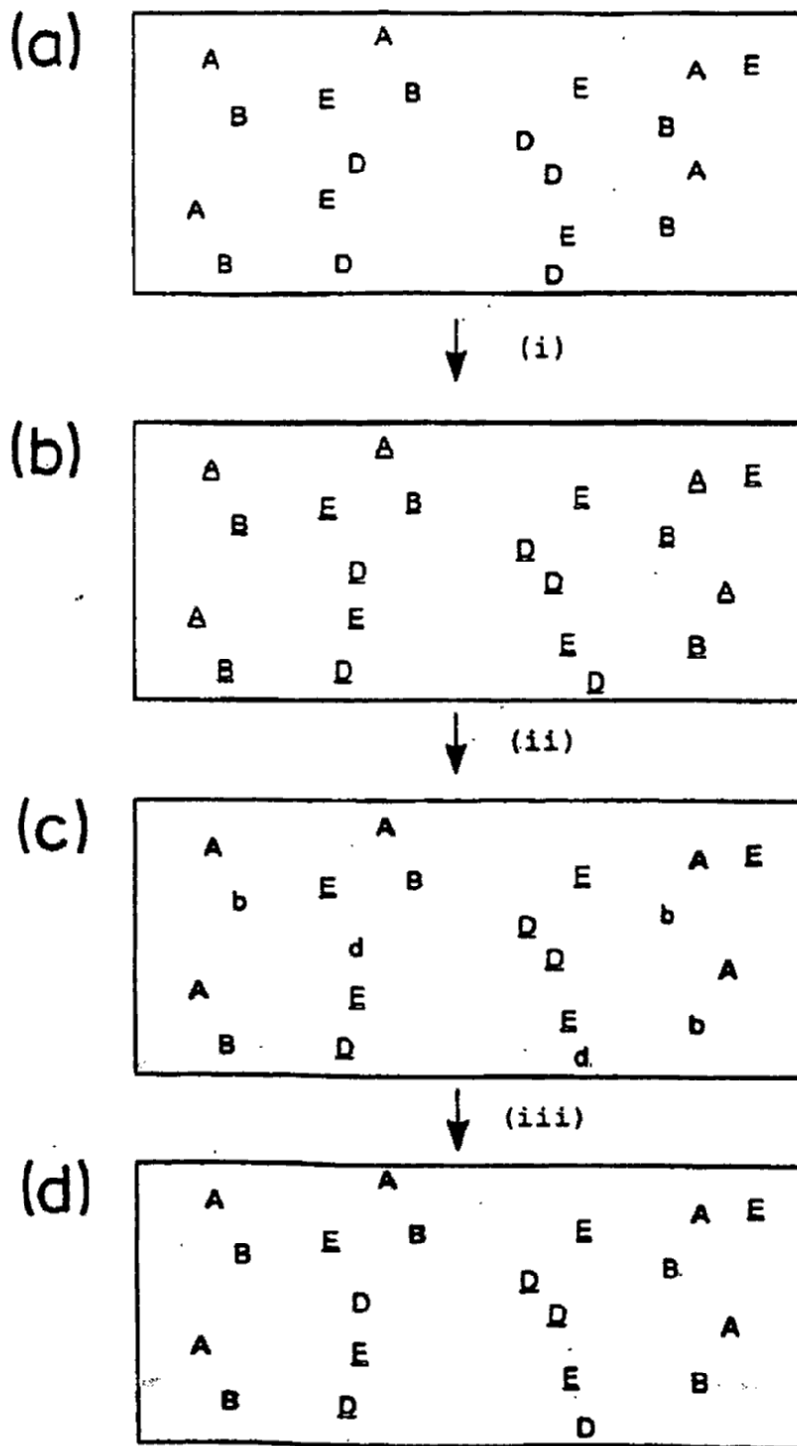
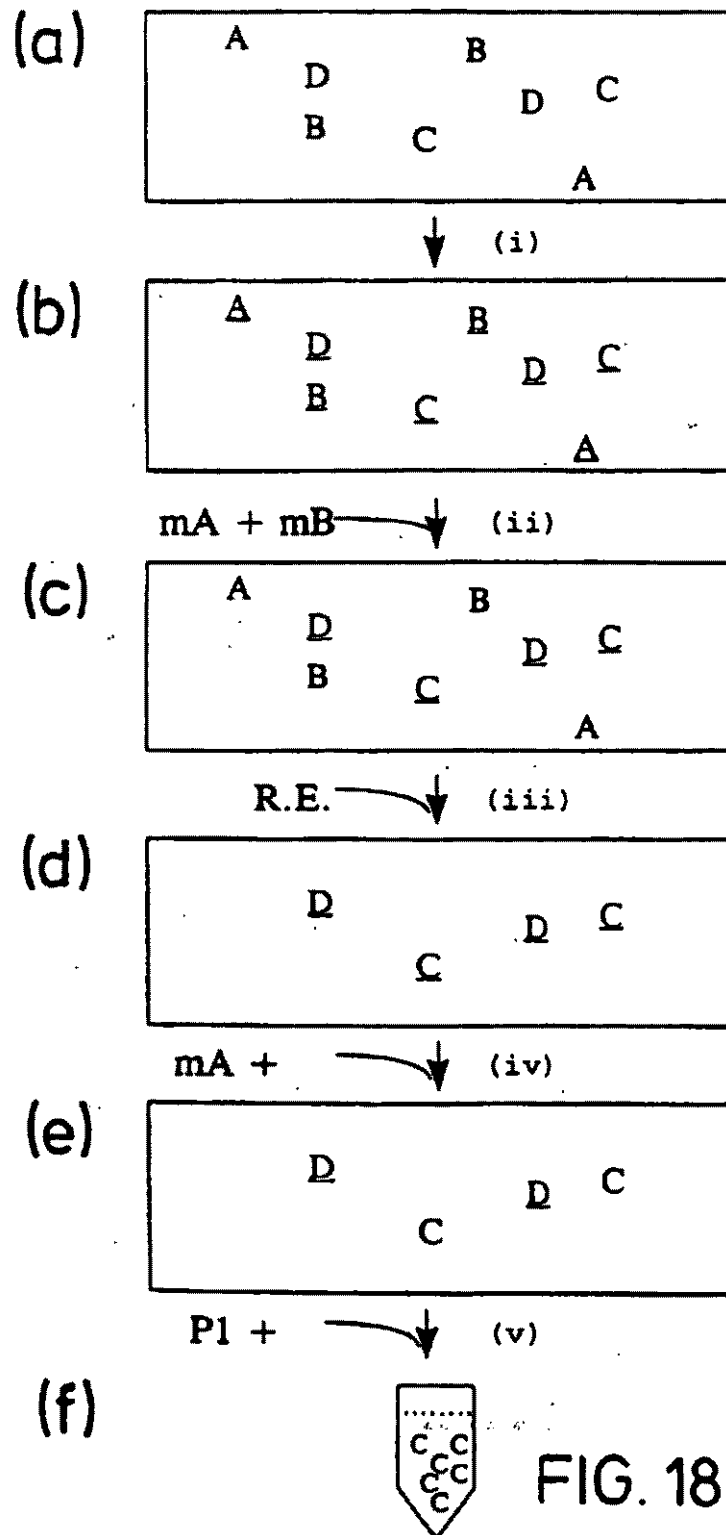


FIG. 17



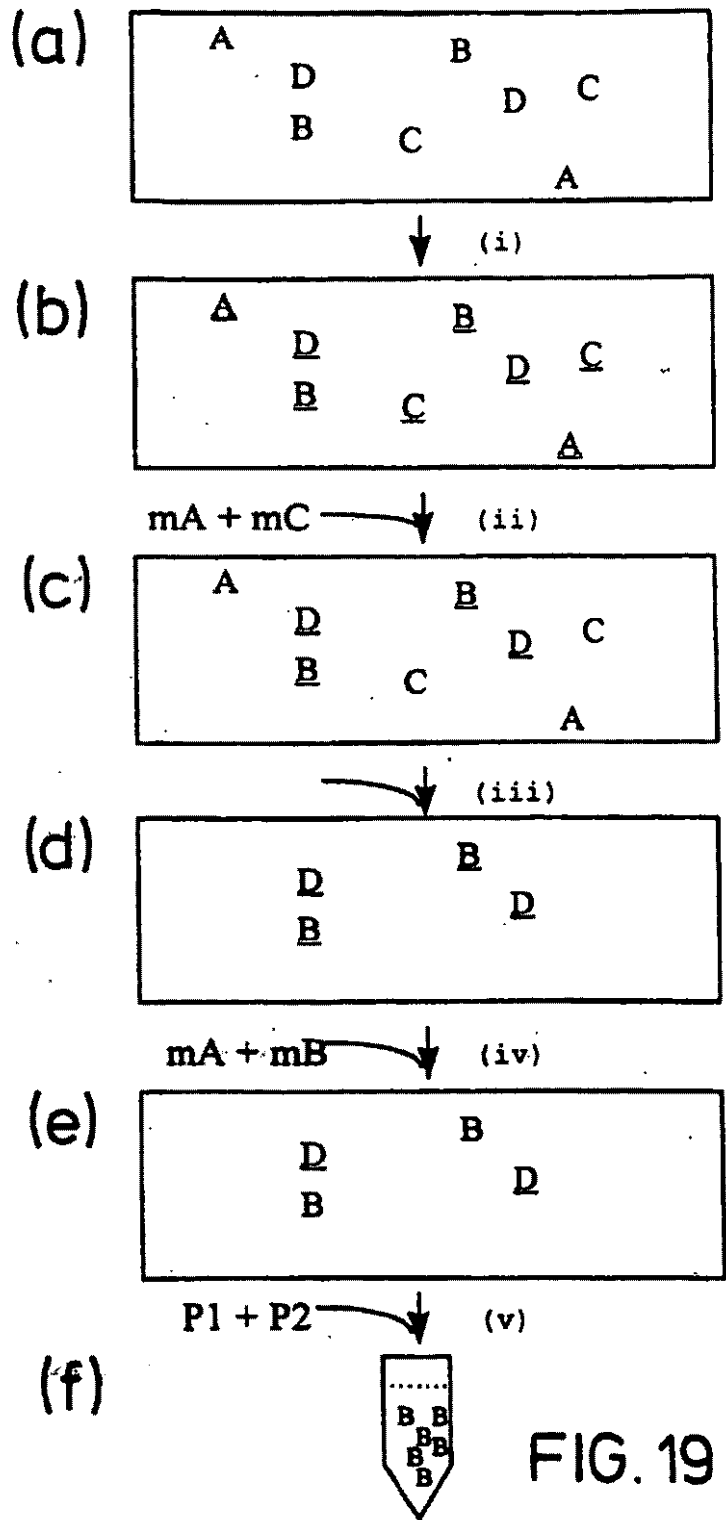


FIG. 19

REFERENCIAS CITADAS EN LA DESCRIPCIÓN

Esta lista de referencias citadas por el solicitante es únicamente para la comodidad del lector. No forma parte del documento de la patente europea. A pesar del cuidado tenido en la recopilación de las referencias, no se pueden excluir errores u omisiones y la EPO niega toda responsabilidad en este sentido.

Documentos de patentes citados en la descripción

- WO 9604404 A [0028]
- WO 9527080 A [0173]

Literatura diferente de patentes citada en la descripción

- 15 • FRIEDRICH, G.A. Moving beyond the genome projects. *Nature Biotechnology*, 1996, vol. 14, 1234 [0002]
- VENTER, J.C. ; H.O SMITH ; L. HOOD. A new strategy for genome sequencing. *Nature*, 1996, vol. 381, 364 [0002]
- 20 • MAIER, E. ; S. MEIEREWER ; A.R. AHMADI ; J. CURTIS ; H. LEHRACH. Application of robotic technology to automated sequence fingerprint analysis by oligonucleotide hybridization. *Journal Of Biotechnology*, 1994, vol. 35, 191 [0003]
- CHEE, M. ; R.YANG ; E. HUBBELL ; A. BERNO ; X.C. HUANG ; D. STERN ; J. WINKLER ; D.J. LOCKHART ; M.S. MORRIS ; S.P.A. FODOR. Accessing genetic information with high-density DNA arrays. *Science*, 1996, vol. 274 (5287), 610-614 [0003]
- 25 • MAIER, E. Robotic technology in library screening. *Laboratory Robotics and Automation*, 1995, vol. 7, 123 [0003]
- GRUNSTEIN, M. ; D.S. HOGNESS. Colony Hybridization: A method for the isolation of cloned DNAs that contain a specific gene. *Proceedings of the National Academy of Science, USA*, 1975, vol. 72, 3961 [0006]
- 30 • SAIKI, R. ; S. SCHARF ; F. FALOONA ; K. MULLIS ; G. HORN ; H. ERLICH ; N. ARNHEIM. *Science*, 1985, vol. 230, 1350-1354 [0008]
- OROSKAR, A.A. ; S.E. RASMUSSEN ; H.N. RASMUSSEN ; S.R. RASMUSSEN ; B.M. SULLIVAN ; A. JOHANSSON. Detection of immobilised amplicons by ELISA-like techniques. *Clinical Chemistry*, 1996, vol. 42, 1547 [0010]
- 35 • BLANCHARD, A.P. ; L. HOOD. Oligonucleotide array synthesis using ink jets. *Microbial and Comparative Genomics*, 1996, vol. 1, 225 [0012]
- 40 • YERSHOV, G. et al. DNA analysis and diagnostics on oligonucleotide microchips. *Proceedings of the National Academy of Science, USA*, 1996, vol. 93, 4913-4918 [0012]
- FODOR, S.P.A. et al. Light directed, spatially addressable parallel chemical synthesis. *Science*, 1991, vol. 251, 767 [0013]
- LOCKHART, D.J. et al. Expression monitoring by hybridisation to high-density oligonucleotide arrays. *Nature Biotechnology*, 1996, vol. 14, 1675 [0013]
- CHEE, M. et al. Accessing genetic information with high-density DNA arrays. *Science*, 1996, vol. 274, 610 [0013]
- MASKOS, U. ; E.M SOUTHERN. Oligonucleotide hybridizations on glass supports: a novel linker for oligonucleotide synthesis and hybridization properties of oligonucleotides synthesised in situ. *Nucleic Acids Research*, 1992, vol. 20 (7), 1679-84 [0030]
- LAMTURE, J.B. Direct-detection of nucleic-acid hybridization on the surface of a charge-coupled-device. *Nucleic Acids Research*, 1994, vol. 22 (11), 2121-2125 [0030]
- CHRISEY, L.A. ; G.U. LEE ; C.E. OFERRALL. Covalent attachment of synthetic DNA to self-assembled monolayer films. *Nucleic Acids Research*, 1996, vol. 24 (15), 3031-3039 [0030]
- S. TERRETTAZ et al. Protein binding to supported lipid membranes. *Langmuir*, 1993, vol. 9, 1361 [0032]
- GOVES, J.T. et al. Micropatterning Fluid Bilayers on Solid Supports. *Science*, 1997, vol. 275, 651 [0032]
- LARSEN, A.E. ; D.G. GRIER. Like charge attractions in metastable colloidal crystallites. *Nature*, 1997, vol. 385, 230 [0032]
- PIRRUNG, M.C. ; BRADLEY, J.C. Comparison of methods for photochemical phosphoramidite-based DNA-synthesis. *Journal Of Organic Chemistry*, 1995, vol. 60 (20), 6270-6276 [0034]
- SAMBROOK et al. *Molecular Cloning*. Cold Spring Harbor Laboratory Press, 1989 [0043]
- PCR: *Clinical Diagnostics and Research*. Springer-Verlag, 1992 [0047]

60