

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 742 310**

51 Int. Cl.:

C12Q 1/68 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **16.12.2014 PCT/CN2014/093940**

87 Fecha y número de publicación internacional: **09.07.2015 WO15101175**

96 Fecha de presentación y número de la solicitud europea: **16.12.2014 E 14877282 (5)**

97 Fecha y número de publicación de la concesión europea: **29.05.2019 EP 3090068**

54 Título: **Método y kit para la detección no invasiva de mutaciones del gen EGFR**

30 Prioridad:

31.12.2013 CN 201310756037

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

13.02.2020

73 Titular/es:

**BERRY GENOMICS CO., LTD. (100.0%)
Room 801, Floor 8, Building No. 5, Yard 4,
Shengmingyuan Road, Science Park, Changping
District
Beijing 102206, CN**

72 Inventor/es:

**WANG, LINA;
RU, LANLAN;
LI, TIANCHENG;
ZHANG, YAXI y
ZHANG, JIANGUANG**

74 Agente/Representante:

VEIGA SERRANO, Mikel

ES 2 742 310 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 y kit para la detección no invasiva de mutaciones del gen EGFR

5 Sector de la técnica

La presente invención se refiere al campo del diagnóstico genético. Más específicamente, la presente invención se dirige a un método para detectar mutaciones del gen del receptor del factor de crecimiento epidérmico (EGFR), así como a un kit para su uso en la detección de mutaciones del gen EGFR.

10

Estado de la técnica

El EGFR (receptor del factor de crecimiento epidérmico) normalmente está incluido en la membrana celular sobre la superficie celular. El gen EGFR tiene 118 kb de longitud, incluye 28 exones y codifica una glucoproteína de 170 kD compuesta por 1186 aminoácidos. Es una proteína de membrana que juega un papel importante en la regulación de la proliferación, crecimiento, reparación y supervivencia de células tumorales. Actualmente, la terapia dirigida se ha convertido en una herramienta importante en el tratamiento clínico del Carcinoma Pulmonar No Microcítico (CPNM). Iressa (Gefitinib, AstraZeneca) y Tarceva (Erlotinib, Roche), que funcionan como inhibidores de la tirosina quinasa EGFR (TKI), son los principales medicamentos aprobados por la FDA para la terapia dirigida del CPNM. Sin embargo, los experimentos clínicos mostraron que Iressa y Tarceva solo tenían efectos terapéuticos significativos en pacientes con CPNM al 10-30%. Otros estudios indicaron que las mutaciones del gen EGFR se relacionan con los efectos terapéuticos de la terapia dirigida del CPNM, y la mayoría de los pacientes que portaban mutaciones del gen EGFR mostraron efectos terapéuticos significativos. Una gran cantidad de documentos de investigación indicaron que las mutaciones del gen EGFR se localizan principalmente en el dominio de codificación de la tirosina quinasa (exones 18-21), en donde la delección en el exón 19 (746-753) representa aproximadamente el 45% de todas las mutaciones, y la sustitución en el exón 21 (principalmente L858R) representa aproximadamente el 40% de todas las mutaciones. En este momento, se cree que de manera general estas dos mutaciones novedosas pueden mejorar la sensibilidad de las células tumorales contra TKI, y pueden usarse como un índice eficaz para predecir el tratamiento con TKI. Por lo tanto, la detección de mutaciones del gen EGFR tiene un valor referencial importante para guiar la administración clínica en pacientes con CPNM. Los métodos para detectar mutaciones del gen EGFR actualmente utilizados en el tratamiento clínico incluyen: 1) secuenciación tradicional. Este método tiene una alta precisión. Sin embargo, el alto requisito en cuanto a la fuente de la muestra, el tiempo de secuenciación largo, la necesidad de análisis de secuencia y el alto coste de este método limitan su uso en clínica. 2) Reacción en cadena de la polimerasa de polimorfismo de conformación de cadena simple (PCR-SSCP). Es un método clásico para detectar mutaciones genéticas y puede detectar mutaciones desconocidas. Tiene las ventajas de un funcionamiento sencillo y de una alta sensibilidad. Sin embargo, también tiene una clara desventaja, que es la necesidad de un control de estándar paralelo. Por otra parte, este método tiene una alta tasa de falsos positivos: la tasa de detección es del 75-95 % cuando el producto de PCR del ensayo es inferior a 200 pb, mientras que la tasa de detección es solo de aproximadamente el 50 % cuando la muestra de PCR analizada es de más de 400 pb. 3) PCR enriquecida con mutación: una PCR de dos pasos que usa una endonucleasa de restricción para digerir de forma selectiva el gen EGFR de tipo salvaje. Después de la primera PCR, el gen EGFR de tipo salvaje se digiere de forma selectiva y los genes EGFR mutados se enriquecen, y luego se realiza la segunda PCR. El producto de PCR se detecta mediante electroforesis, y se determina si el gen EGFR está mutado en función de los resultados de detección del producto de PCR. Este método es altamente sensible y puede detectar una mutación entre 10^3 - 10^4 genes EGFR de tipo salvaje. Sin embargo, este método necesita realizar dos veces la PCR y la digestión enzimática, por tanto es complejo y requiere tiempo. Además, existen tecnologías como el AMRS y la PCR micro digital, pero su aplicación en la clínica aún necesita tiempo. El documento WO2013138510 detecta mutaciones en el gen EGFR, utilizando ADN plasmático (ejemplo 1, página 28 y siguientes). El documento WO2013138510 usa además cebadores etiquetados para una primera amplificación y una amplificación adicional (véase también la Figura 1). Estos amplicones etiquetados se usaron luego en secuenciación paralela masiva para analizar la mutación.

Por tanto, existe la necesidad urgente de un método rápido y eficaz para detectar clínicamente mutaciones del gen EGFR. Durante la investigación de detección de fragmentos de ADN, el inventor descubrió un nuevo método para detectar fragmentos de ADN, que incluye ciclar y luego amplificar fragmentos de ADN. Basándose en este descubrimiento y en combinación con la tecnología de secuenciación de alto rendimiento de segunda generación, el inventor mejoró el método antes mencionado y diseñó cebadores optimizados específicamente basados en los genes EGFR, y desarrolló un método y un kit del mismo para secuenciar y analizar mutaciones de genes EGFR en ADN plasmáticos.

60 Objeto de la invención

En un aspecto, la presente invención proporciona un método para detectar de forma no invasiva mutaciones del gen EGFR en sujetos, que comprende las siguientes etapas:

- 65 --diseñar cebadores de acuerdo con los exones del gen EGFR;
--extraer ADN plasmáticos de los sujetos;

ES 2 742 310 T3

- conectar los ADN plasmáticos extraídos con conectores de etiquetado;
 - pre-amplificar los ADN plasmáticos conectados con conectores de etiquetado;
 - ciclar los ADN amplificados para obtener los ADN ciclados;
 - amplificar por PCR los ADN ciclados utilizando los cebadores diseñados; y
- 5 --realizar secuenciación de alto rendimiento del producto amplificado por PCR y analizar las mutaciones del gen EGFR.

Además, los cebadores son un par de cebadores que están adyacentes y ampliados de forma regresiva.

- 10 Además, los cebadores del par de cebadores ampliado de forma regresiva se encuentran en el extremo 5' o 3' del exón del gen EGFR.

Además, El espacio del par de cebadores ampliado de forma regresiva es 0-1/2 de los pares de bases totales de los fragmentos de ADN.

- 15 Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de los genes EGFR.

- 20 Además, el extremo 5' de los cebadores ampliados de forma regresiva contiene secuencias enlazadoras para la biblioteca de secuenciación de alto rendimiento.

Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20, o exón 21 de los genes EGFR, y las secuencias de los cebadores son las siguientes:

E18-1-3F: CCCAGCTTGTGGAGCCTC

E18-1-3R: GACAAGAACACAGAGACAAGGGT

E18-2-F: GCAGGGCCTCTCATGGTC

E18-2-R: CCTGTGCCAGGGACCTTAC

E19-1-F: ACGTCTTCCTTCTCTCTGTCAT

E19-1-R: GTGAGATGGTGCCACATGCT

E19-2-F: GTCCATGGCTCTGAACCTCA

E19-2-R: CCACACAGCAAAGCAGAAAC

E20-1-F: CACTGACGTGCCTCTCC

E20-1-R: CTTGCATGGTGGCCAGA

E20-2-1F: CCTCCCCGTATCTCCCT

E20-2-1R: GGAGATAAGGAGCCAGGAT

E21-1-F: AGCAGGGTCTTCTCTGTTTCA

E21-1-R: GAGGGACAGATCATCATGGG

E21-2-F: TTTCTGACACCAGGGACC

- 25 E21-2-R: TGACCTAAAGCCACCTCCTT.

Además, la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN.

- 30 Además, los genes EGFR en los ADN plasmáticos tienen mutaciones de inserción, supresión, sustitución o fusión génica.

ES 2 742 310 T3

Además, las tecnologías de secuenciación de alto rendimiento se seleccionan de Roche/454 FLX, Illumina/Hiseq/Miseq, Applied Biosystems SOLID y Life Technologies/Ion Torrent/Proton.

5 En otro aspecto, La presente invención proporciona un kit para la detección no invasiva de mutaciones del gen EGFR, que comprende: reactivos para extraer ADN plasmático, una ADN ciclasa, cebadores y reactivos para amplificar ADN dianas.

Además, el kit comprende cebadores y reactivos para pre-amplificar regiones que se analizarán en los genes EGFR.

10 Además, el kit comprende reactivos para la secuenciación de alto rendimiento.

Además, los reactivos para la secuenciación de alto rendimiento son aplicables a las siguientes tecnologías de secuenciación de alto rendimiento: Roche/454 FLX, Illumina /Hiseq/Miseq, Applied Biosystems SOLID y Life Technologies/Ion Torrent/Proton.

15 Además, los cebadores para las regiones de amplificación que se analizarán en los genes EGFR son un par de cebadores que están adyacentes y están ampliados de forma regresiva.

20 Además, los cebadores del par de cebadores ampliado de forma regresiva se encuentran en el extremo 5' o 3' de los sitios o regiones a analizar en los genes EGFR.

Además, el espacio del par de cebadores ampliado de forma regresiva es 0-1/2 del total de pares de bases de los ADN plasmáticos.

25 Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de los genes EGFR.

Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20, o exón 21 de los genes EGFR, y las secuencias de los cebadores son las siguientes:

30

E18-1-3F: CCCAGCTTGTGGAGCCTC

E18-1-3R: GACAAGAACACAGAGACAAGGGT

E18-2-F: GCAGGGCCTCTCATGGTC

E18-2-R: CCTGTGCCAGGGACCTTAC

E19-1-F: ACGTCTTCCTTCTCTCTGTGCAT

E19-1-R: GTGAGATGGTGCCACATGCT

E19-2-F: GTCCATGGCTCTGAACCTCA

E19-2-R: CCACACAGCAAAGCAGAAAC

E20-1-F: CACACTGACGTGCCTCTCC

E20-1-R: CTTCGCATGGTGGCCAGA

E20-2-1F: CCTCCCCGTATCTCCCT

E20-2-1R: GGAGATAAGGAGCCAGGAT

E21-1-F: AGCAGGGTCTTCTCTGTTTCA

E21-1-R: GAGGGACAGATCATCATGGG

E21-2-F: TTTCTGACACCAGGGACC

E21-2-R: TGACCTAAAGCCACCTCCTT.

Además, los conectores de ADN en plasma contienen secuencias de etiquetado.

Además, los ADN plasmáticos se pre-amplifican antes de ciclarse.

Además, la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN.

- 5 Además, los sitios o regiones a analizar en los genes EGFR tienen mutaciones de inserción, supresión, sustitución o fusión génica.

10 En otro aspecto más, la presente invención proporciona un uso de cebadores contra exones del gen EGFR en la preparación de reactivos de diagnóstico o kits para detectar de forma no invasiva mutaciones del gen EGFR en sujetos, que se caracteriza por que los reactivos o kits de diagnóstico son aplicables a un método para detectar de forma no invasiva mutaciones del gen EGFR en sujetos que comprende los siguientes pasos:

- 15 --extraer ADN plasmáticos de los sujetos;
--conectar los ADN plasmáticos extraídos con conectores de etiquetado;
--pre-amplificar los ADN plasmáticos conectados con conectores de etiquetado;
--ciclar los ADN amplificados para obtener los ADN ciclados;
--amplificar por PCR los ADN ciclados utilizando los cebadores diseñados; y
--realizar secuenciación de alto rendimiento del producto amplificado por PCR y analizar las mutaciones del gen EGFR.

20 Además, los cebadores son un par de cebadores que están adyacentes y ampliados de forma regresiva.

Además, los cebadores del par de cebadores ampliado de forma regresiva se encuentran en el extremo 5' o 3' de los exones del gen EGFR.

25 Además, el espacio del par de cebadores ampliado de forma regresiva es 0-1/2 del total de pares de bases de los fragmentos de ADN.

30 Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de los genes EGFR.

Además, el extremo 5' de los cebadores ampliados de forma regresiva contiene secuencias enlazadoras para la biblioteca de secuenciación de alto rendimiento.

35 Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20, o exón 21 de los genes EGFR, y las secuencias de los cebadores son las siguientes:

E18-1-3F: CCCAGCTTGTGGAGCCTC

E18-1-3R: GACAAGAACACAGAGACAAGGGT

E18-2-F: GCAGGGCCTCTCATGGTC

E18-2-R: CCTGTGCCAGGGACCTTAC

E19-1-F: ACGTCTTCTTCTCTCTGTGCAT

E19-1-R: GTGAGATGGTGCCACATGCT

E19-2-F: GTCCATGGCTCTGAACCTCA

E19-2-R: CCACACAGCAAAGCAGAAAC

E20-1-F: CACTGACGTGCCTCTCC

E20-1-R: CTTGCATGGTGGCCAGA

E20-2-1F: CCTCCCCGTATCTCCCT

E20-2-1R: GGAGATAAGGAGCCAGGAT

E21-1-F: AGCAGGGTCTTCTCTGTTTCA

E21-1-R: GAGGGACAGATCATCATGGG

E21-2-F: TTTCTGACACCAGGGACC

E21-2-R: TGACCTAAAGCCACCTCCTT.

Además, la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN.

Además, los genes EGFR en los ADN plasmáticos tienen mutaciones de inserción, supresión, sustitución o fusión génica.

5 Además, la secuenciación de alto rendimiento se selecciona de Roche/454 FLX, Illumina/Hiseq/Miseq, Applied Biosystems SOLID y Life Technologies/Ion Torrent/Proton.

Descripción de las figuras

10 La figura 1 muestra el diseño experimental de la presente invención. El ADN plasmático se pre-amplifica primero para aumentar el rendimiento del producto de detección, y luego el ADN pre-amplificado se cicla para aumentar las plantillas disponibles para la detección. El gen EGFR se amplifica para preparar una biblioteca para la secuenciación de alto rendimiento.

15 La figura 2 muestra el principio de diseño de la ciclación de la presente invención. El ADN pre-amplificado es un ADN de doble cadena con conectores universales en ambos extremos. El puente de ADN con secuencias complementarias inversas está diseñado en base a las secuencias de los conectores universales en ambos extremos, y se forma una doble cadena parcial después del templado y la hibridación. El producto ciclado de cadena sencilla finalmente se obtiene mediante ligación usando la Tap ligasa. La figura 2A muestra un esquema general, y la figura 2B muestra el principio de ciclación.

20 La figura 3 muestra la velocidad de ciclación de la presente invención. Esta figura muestra los resultados de imagen del gel de desnaturalización al 6%. 1: un producto de PCR antes de la ciclación; 2: un producto después de la ciclación (no digerido por exonucleasa), M: un marcador; 3: un producto digerido con exonucleasa después de la ciclación (no purificado); 4: un producto digerido y purificado por exonucleasa después de la ciclación.

25 La figura 4 muestra la biblioteca de muestras de EGFR final. Esta figura muestra los resultados del gel de agarosa al 2%. 1: LC2; 2: LC3; 3: LC113; 4: LC314; 5: LC320; 6: NTC (un control negativo). La muestra es una muestra de cáncer pulmonar (CP).

Descripción detallada de la invención

35 Con los avances en la tecnología de secuenciación, la secuenciación tradicional de Sanger no ha podido cumplir completamente con los requisitos de la investigación. Por tanto, aparece la tecnología de secuenciación de segunda generación, con menor coste, mayor rendimiento, velocidad más rápida, y con la capacidad de secuenciar todo el genoma. El principio principal de la tecnología de secuenciación de segunda generación es la secuenciación de alto rendimiento por síntesis, concretamente, la determinación de secuencias de ADN mediante la captura del etiquetado final recién sintetizado. Las plataformas tecnológicas disponibles incluyen principalmente Roche/454 FLX, Illumina/Genome Analyzer/Hiseq/Miseq, Applied Biosystems SOLID, life Technologies/Ion Torrent, y similares. Tomando el producto de Illumina como ejemplo, El HiSeq 2000 puede alcanzar un rendimiento de secuenciación de 30X de cobertura en 6 genomas humanos por ciclo, es decir, aproximadamente 600 G/ciclo, y el tiempo de operación se reduce a 30 minutos. Asimismo, con el desarrollo de la tecnología de secuenciación de segunda generación, la investigación sobre su aplicación clínica se ha desarrollado rápidamente. La investigación muestra que la salud genética del feto se puede juzgar mediante la secuenciación de ADN plasmáticos maternos, y que la secuenciación de ADN plasmáticos en los sujetos se puede utilizar para la detección temprana del cáncer, lo cual tendrá una amplia aplicación en el futuro.

50 Los ADN plasmáticos, también conocidos como ADN circulantes, son ADN extracelulares en la sangre y tienen de decenas a cientos de nucleótidos de longitud (un pico principal de aproximadamente 167 pb). Se presenta en forma de complejo de ADN-proteína, o como fragmentos de ADN libres. Normalmente, los ADN plasmáticos proceden de la liberación de ADN de una pequeña cantidad de células senescentes y de células muertas. En condiciones saludables, la generación y eliminación de los ADN circulantes se encuentra en un equilibrio dinámico y se mantienen a un nivel bajo relativamente estable. 1 ml de plasma de una persona normal contiene alrededor de 2000 ADN genómicos. Los ADN circulantes pueden reflejar el estado metabólico de las células en el cuerpo humano, por tanto, es un índice importante para juzgar la salud. El cambio en la cantidad y calidad de los ADN circulantes en la sangre periférica está estrechamente relacionado con varias enfermedades (incluyendo tumor, traumatismo severo complejo, trasplante de órganos, enfermedades relacionadas con el embarazo, enfermedades infecciosas, fallo de órganos y similares). Como índice de detección no invasivo, se espera que sea un marcador molecular importante para el diagnóstico temprano de algunas enfermedades, realizando un seguimiento de las afecciones y evaluando los efectos terapéuticos y el pronóstico de las enfermedades. Por ejemplo, las investigaciones muestran que EGFR regula el progreso del ciclo celular, la reparación y supervivencia de células tumorales, y al mismo tiempo se relaciona con metástasis tumoral. Recientemente, la terapia molecular dirigida utilizando EGFR como diana de la terapia ha recibido una atención generalizada de las comunidades de cáncer dentro y fuera del país, y se ha aprobado un inhibidor de la tirosina quinasa EGFR, Iressa (Genfitinib), por la Administración de Alimentos y Medicamentos de los Estados Unidos (FDA) para tratar el CPNM avanzado. La característica destacada de los medicamentos moleculares dirigidos es que su efecto

terapéutico depende en gran medida de la diana: el efecto terapéutico es significativamente fuerte en pacientes con la "diana", mientras que el efecto terapéutico es débil o nulo en pacientes sin la "diana", lo que retrasa otros tratamientos y empeora las condiciones. Por lo tanto, las administraciones a ciegas sin detección de dianas no solo pueden tener como resultado una gran pérdida económica, sino que también pueden retrasar el tiempo valioso para el tratamiento o incluso agravar las condiciones. No hay que escatimar en juzgar rápida y exactamente si el paciente tiene la diana específica para el tratamiento farmacológico dirigido. La detección tradicional de EGFR detecta principalmente la sección de tejido de la lesión mediante FISH o qPCR. Sin embargo, se ha descubierto que hay más ADN libres en el plasma de pacientes con CPNM, aproximadamente 10 veces más que en el de las personas normales. Una gran cantidad de ADN libre en plasma deriva de la liberación de ADN de células tumorales senescentes y muertas. Son similares a los ADN genómicos tumorales en características genéticas, y sus mutaciones incluyen delección, mutación puntual y mayor número de copias. Las mutaciones del gen EGFR se pueden detectar examinando los ADN plasmáticos en pacientes con CPNM y, por lo tanto, permiten detectar la expresión de EGFR de forma no invasiva. La presente invención detecta rápidamente las expresiones y mutaciones de EGFR en los ADN plasmáticos mediante la tecnología de secuenciación de segunda generación, de manera precisa, de manera no invasiva y con alta sensibilidad, y por lo tanto proporciona varias bases de diagnóstico para los pacientes.

En vista de la importancia clínica de la detección no invasiva por secuenciación de ADN plasmáticos y el rápido desarrollo de la secuenciación de alto rendimiento de segunda generación, el inventor descubrió que la secuenciación de ADN plasmáticos a gran escala puede detectar las expresiones y mutaciones del gen EGFR más rápidamente, con precisión y de forma no invasiva. Es aplicable a una variedad de plataformas de secuenciación de alto rendimiento de segunda generación, incluyendo, pero sin limitarse a, Roche/454 FLX, Illumina/Genome Analyzer/HiSeq/Miseq, Applied Biosystems SOLID, life Technologies/Ion Torrent, y similares.

La presente invención se basa en los siguientes dos hechos: 1) los ADN libres del plasma en pacientes son similares a los ADN genómicos en cuanto a características genéticas. Los ADN libres del plasma de los pacientes tienen un contenido más alto que los de las personas normales, y a menudo contienen muchas mutaciones, aunque cada mutación puede ser de baja frecuencia; 2) La secuenciación de alto rendimiento de segunda generación puede obtener la información de ADN plasmáticos libres rápidamente, con precisión y con alto rendimiento. La combinación de estos dos hechos permite la aplicación no invasiva a gran escala de la detección en regiones específicas del genoma. Las investigaciones muestran que los ADN plasmáticos existen como fragmentos con poca cantidad (1 ml de plasma contiene alrededor de 2000 genomas) y de corta longitud (principalmente alrededor de 167 pb), lo que dificulta que la PCR tradicional enriquezca las mutaciones de manera eficaz utilizando ADN plasmático como plantillas, dando como resultado una disminución rápida en la sensibilidad de detección. La presente invención difiere de los métodos tradicionales en que los fragmentos de ADN conectados con los conectores de etiquetado de secuencia se amplifican y se ciclan en una sola cadena, luego por medio de amplificación de cebador consecutiva, las plantillas se usan al máximo y la biblioteca se secuencia por secuenciación de alto rendimiento de extremos emparejados. Las plantillas amplificadas originales se ensamblan en base al análisis de secuencia original, y se registran las secuencias de etiquetado. Las secuencias con la misma posición en el genoma y las mismas secuencias de etiquetado se calculan como una plantilla. Se calcula el número de plantillas amplificadas por cada par de cebadores y se cuenta y registra el número de plantillas mutadas. La presente invención mejora el método de ciclación y optimiza los cebadores para los genes EGFR. El diseño único de secuencias de etiquetado reduce el efecto de fondo y previene la contaminación de manera eficaz. El conjunto único de plantillas se cuenta restaurando las plantillas en los sistemas con precisión y, por lo tanto, finalmente se logra una detección de una sola molécula con alta precisión.

De acuerdo con una realización específica de la presente invención, esta proporciona un método para la detección no invasiva de mutaciones del gen EGFR en sujetos, que comprende las siguientes etapas: 1) diseñar cebadores de acuerdo con los exones del gen EGFR; (2) extraer ADN plasmático de los sujetos; (3) conectar los ADN plasmáticos extraídos con conectores de etiquetado; (4) pre-amplificar por PCR el ADN plasmático conectado por conectores de etiquetado; (5) ciclar los ADN amplificados para obtener ADN ciclados; (6) amplificar por PCR los ADN ciclados utilizando los cebadores diseñados; y (7) realizar secuenciación de alto rendimiento del producto amplificado por PCR y analizar las mutaciones del gen EGFR. "Detección no invasiva" en la presente invención significa que, en comparación con los métodos de detección de rutina en histología, tales como la cirugía, biopsia de tejido y similares, que están directamente contra los tejidos cancerosos y producirán daños corporales en los sujetos, la presente invención solo detecta muestras de sangre de los sujetos. Los métodos tradicionales para detectar ADN o fragmentos de genes requieren la amplificación por PCR de las regiones que se analizarán antes de la detección y, por lo tanto, los ADN o fragmentos de genes que se analizarán deben estar completos. Sin embargo, la mayoría de los fragmentos de ADN en plasma están incompletos y, por lo tanto, los fragmentos de ADN que pueden usarse como plantillas en la amplificación por PCR son pocos en número y son difíciles de detectar mediante una PCR de rutina. Por lo tanto, La amplificación por PCR en la presente invención adopta la tecnología de ciclación del ADN para transformar fragmentos de ADN en ADN cíclicos utilizando secuencias conectoras y enzimas. Se diseñan cebadores basados en las regiones a analizar; la biblioteca de secuenciación se amplifica y construye, y luego se secuencia mediante una tecnología de secuenciación de alto rendimiento; y se analizan las mutaciones del gen EGFR. La figura 1 muestra el principio de la presente invención. El ADN del plasma se extrae primero y se conecta con conectores de etiquetado, el cual está en forma de Y. La región de doble cadena contiene secuencias de etiquetado que se pueden distinguir entre diferentes plantillas. Para aumentar la cantidad de fragmentos de ADN amplificables, se puede realizar pre-amplificación, de forma que se puede anillar la secuencia universal en los cebadores y conectores, y luego se ciclan los ADN pre-

amplificados, por ejemplo, usando un puente de ADN. La figura 2 muestra el principio de ciclación del ADN de la presente invención. La secuencia del puente de ADN y la región del cebador pre-amplificado son complementarias. Los cebadores están diseñados según los exones de EGFR, por ejemplo, el exón 18, 19, 20 y 21, y se lleva a cabo una amplificación por PCR específica de acuerdo con los genes EGFR para generar una biblioteca de secuenciación, que luego se secuencia por la tecnología de secuenciación de alto rendimiento actual para obtener las regiones específicas en los genes EGFR analizados, concretamente, se analizan las secuencias génicas de los exones y luego las mutaciones génicas en la región específica. La secuenciación de alto rendimiento de la presente invención se refiere a la tecnología de secuenciación de segunda generación disponible, tal como Roche/454 FLX, Illumina/Genome Analyzer/Hiseq/Miseq, Applied Biosystems SOLID, life Technologies/Ion Torrent y similares. Las diferentes plataformas difieren en los principios de secuenciación. Tomando IlluminaMiseq como ejemplo, la biblioteca de secuenciación generada se desnaturaliza y se secuencia en alto rendimiento en un secuenciador Miseq. La longitud de lectura es de 250 pb en ambos extremos, y cada sitio de cebador produce aproximadamente 100.000 secuencias. Las secuencias de doble extremo se combinan en una secuencia de extremo único basada en regiones repetidas terminales. Se buscan regiones conectoras en la secuencia de extremo único, y la secuencia se restaura a los fragmentos de ADN plasmático originales, mientras se registra la secuencia de etiquetado en cada secuencia. Los fragmentos de ADN restaurados se comparan con el genoma humano, con las coordenadas iniciales y terminales de la secuencia en el genoma registrado. Se obtiene una secuencia de plantilla única después de la eliminación de la redundancia utilizando las coordenadas de inicio y terminal y secuencias de etiquetado como patrones. El tipo y la proporción de mutaciones se determinan contando las secuencias de plantilla en la zona diana.

De acuerdo con otra realización específica adicional de la presente invención, los cebadores del par de cebadores ampliado de forma regresiva se encuentran en el extremo 5' o 3' de los exones del gen EGFR. Además, el espacio del par de cebadores ampliado de forma regresiva es 0-1/2 del total de pares de bases de los fragmentos de ADN. Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de los genes EGFR. Además, el extremo 5' de los cebadores ampliados de forma regresiva contiene secuencias enlazadoras para la biblioteca de secuenciación de alto rendimiento. Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de genes EGFR, y las secuencias de los cebadores son las siguientes:

E18-1-3F: CCCAGCTTGTGGAGCCTC
 E18-1-3R: GACAAGAACACAGAGACAAGGGT
 E18-2-F: GCAGGGCCTCTCATGGTC
 E18-2-R: CCTGTGCCAGGGACCTTAC
 E19-1-F: ACGTCTTCCTTCTCTCTGTGCAT
 E19-1-R: GTGAGATGGTGCCACATGCT
 E19-2-F: GTCCATGGCTCTGAACCTCA
 E19-2-R: CCACACAGCAAAGCAGAAAC
 E20-1-F: CAACTGACGTGCCTCTCC
 E20-1-R: CTTGCGATGGTGGCCAGA
 E20-2-1F: CCTCCCCGTATCTCCCT
 E20-2-1R: GGAGATAAGGAGCCAGGAT
 E21-1-F: AGCAGGGTCTTCTCTGTTTCA
 E21-1-R: GAGGGACAGATCATCATGGG
 E21-2-F: TTTCTGACACCAGGGACC
 E21-2-R: TGACCTAAAGCCACCTCCTT.

Además, la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN.

Además, las tecnologías de secuenciación de alto rendimiento se seleccionan de Roche/454 FLX, Illumina /Hiseq/Miseq, Applied Biosystems SOLID y Life Technologies/Ion Torrent/Proton. En la presente invención se usa la tecnología Illumina.

De acuerdo con otra realización específica de la presente invención, esta proporciona un kit para la detección no invasiva de mutaciones del gen EGFR, que comprende: reactivos para extraer ADN plasmático, una ADN ciclasa, cebadores y reactivos para amplificar ADN dianas. En la extracción de ADN plasmático se pueden usar los reactivos convencionales o los kits disponibles comercialmente. Además, los cebadores de amplificación de las regiones que se analizarán en los genes EGFR son un par de cebadores que están adyacentes y están ampliados de forma regresiva. Además, los cebadores del par de cebadores ampliado de forma regresiva se encuentran en el extremo 5' o 3' de los sitios o regiones a analizar en los genes EGFR. Además, el espacio del par de cebadores ampliado de forma regresiva es 0-1/2 del total de pares de bases de los ADN plasmáticos. Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de los genes EGFR. Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20, o exón 21 de los genes EGFR, y las secuencias de los cebadores son las siguientes:

ES 2 742 310 T3

E18-1-3F: CCCAGCTTGTGGAGCCTC

E18-1-3R: GACAAGAACACAGAGACAAGGGT

E18-2-F: GCAGGGCCTCTCATGGTC

E18-2-R: CCTGTGCCAGGGACCTTAC

E19-1-F: ACGTCTTCCTTCTCTCTGTCAT

E19-1-R: GTGAGATGGTGCCACATGCT

E19-2-F: GTCCATGGCTCTGAACCTCA

E19-2-R: CCACACAGCAAAGCAGAAAC

E20-1-F: CACTGACGTGCCTCTCC

E20-1-R: CTTGCATGGTGGCCAGA

E20-2-1F: CCTCCCCGTATCTCCCT

E20-2-1R: GGAGATAAGGAGCCAGGAT

E21-1-F: AGCAGGGTCTTCTCTGTTTCA

E21-1-R: GAGGGACAGATCATCATGGG

E21-2-F: TTTCTGACACCAGGGACC

E21-2-R: TGACCTAAAGCCACCTCCTT.

Además, el kit comprende cebadores y reactivos para pre-amplificar las regiones que se analizarán en los genes EGFR. De manera específica, los reactivos y cebadores para la pre-amplificación incluyen la ADN polimerasa Taq y su tampón, y los cebadores para la pre-amplificación que son complementarios a los conectores en forma de Y.

Además, el kit comprende reactivos para la secuenciación de alto rendimiento. Además, los reactivos para la secuenciación de alto rendimiento son aplicables a las siguientes tecnologías de secuenciación de alto rendimiento: Roche/454 FLX, Illumina /Hiseq/Miseq, Applied Biosystems SOLID y Life Technologies/Ion Torrent/Proton. Además, los conectores de conexión de ADN plasmático contienen secuencias de etiquetado. Además, los ADN plasmáticos se pre-amplifican antes de ciclarse. Además, la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN.

De acuerdo con otra realización específica de la presente invención, esta proporciona un uso de cebadores de acuerdo con exones del gen EGFR en la preparación de reactivos de diagnóstico o kits para detectar de forma no invasiva mutaciones del gen EGFR en sujetos, que se caracteriza por que los reactivos o kits de diagnóstico son aplicables a un método para detectar de forma no invasiva mutaciones del gen EGFR en sujetos que comprende los siguientes pasos:

- extraer el ADN plasmático de los sujetos;
- conectar los ADN plasmáticos extraídos con conectores de etiquetado;
- pre-amplificar los ADN plasmáticos conectados con conectores de etiquetado;
- ciclar los ADN amplificados para obtener los ADN ciclados;
- amplificar por PCR los ADN ciclados utilizando los cebadores diseñados; y
- realizar secuenciación de alto rendimiento del producto amplificado por PCR y analizar las mutaciones del gen EGFR.

Además, los cebadores son un par de cebadores que están adyacentes y ampliados de forma regresiva. Además, los cebadores del par de cebadores ampliado de forma regresiva se encuentran en el extremo 5' o 3' de los exones del gen EGFR. Además, el espacio del par de cebadores ampliado de forma regresiva es 0-1/2 del total de pares de bases de los fragmentos de ADN. Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón

20 o exón 21 de los genes EGFR. Además, el extremo 5' de los cebadores ampliados de forma regresiva contiene secuencias enlazadoras para la biblioteca de secuenciación de alto rendimiento. Además, los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20, o exón 21 de los genes EGFR, y las secuencias de los cebadores son las siguientes:

5

E18-1-3F: CCCAGCTTGTGGAGCCTC
 E18-1-3R: GACAAGAACACAGAGACAAGGGT
 E18-2-F: GCAGGGCCTCTCATGGTC
 E18-2-R: CCTGTGCCAGGGACCTTAC
 E19-1-F: ACGTCTTCCTTCTCTCTGTCAT
 E19-1-R: GTGAGATGGTGCCACATGCT
 E19-2-F: GTCCATGGCTCTGAACCTCA
 E19-2-R: CCACACAGCAAAGCAGAAAC
 E20-1-F: CACACTGACGTGCCTCTCC
 E20-1-R: CTTGCGATGGTGGCCAGA
 E20-2-1F: CCTCCCCGTATCTCCCT
 E20-2-1R: GGAGATAAGGAGCCAGGAT
 E21-1-F: AGCAGGGTCTTCTCTGTTTCA
 E21-1-R: GAGGGACAGATCATCATGGG
 E21-2-F: TTTCTGACACCAGGGACC
 E21-2-R: TGACCTAAAGCCACCTCCTT.

Además, la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN. Además, las tecnologías de secuenciación de alto rendimiento se seleccionan de Roche/454 FLX, Illumina /Hiseq/Miseq, Applied Biosystems SOLID y Life Technologies/Ion Torrent/Proton.

10

Ejemplos

Ejemplo 1

15

La plantilla de ADN plasmático se amplificó usando conectores de diseño propio y de acuerdo con un método para construir una biblioteca de secuenciación de alto rendimiento de ADN plasmático (es decir, pre-amplificación por PCR usando cebadores fosforilados después de la conexión del conector). El producto de PCR se purificó mediante un corte de gel y se cicló mediante una conexión por puente de ADN. El producto ciclado se digirió por Exo III, se purificó y se cribó por PCR multiplex con 8 pares de cebadores consecutivos (los cebadores contienen secuencias universales para construir una biblioteca de secuenciación). Los sitios de mutación deben estar cerca del cebador directo o del cebador inverso. La biblioteca se obtuvo finalmente purificando el producto de amplificación amplificado por los cebadores universales.

20

1. Diseño del conector. Se templa para obtener una doble cadena, en donde X es una secuencia de etiquetado:

25

ssCycADT-1: GTCTCATCCCTGCGTGXXXX1
 ssCycADT-2: pXXXXCACGCAGGGTACGTGT

La estructura del producto de conexión:

ES 2 742 310 T3

Superior: GTCTCATCCCTGCGTGXXXXTNNN AXXXXCACGCAGGGTACGTGT
Inferior: TGTGCATGGGACGCACXXXXA NNN TXXXGTGCGTCCCTACTCTG

Cebadores:

ssCycUniprimer-F: pGTCTCATCCCTGCGTG
ssCycUniprimer-R: pACACGTACCCTGCGTG

5 La estructura de la biblioteca después de la pre-amplificación:

pGTCTCATCCCTGCGTGXXXXTNNN AXXXXCACGCAGGGTACGTGT
CAGAGTAGGGACGCACXXXXA NNN TXXXGTGCGTCCCATGCACA_f

Cebadores consecutivos para la amplificación en zonas diana:
EXON18 (123 pb)

CAAGTGCCGTGCTCTGGCACCCAAGCCCATGCCGTGGCTGCTGGTCCCCCTGCTGGGCCATGTCTGGCACTGCTTTCCAGCATGGTG
AGGGCTGAGGTGACCCTTGTCTCTGTGTTCTTGTCCCCCCAG CTTGTGGAGCCTTTACACCCAGTGGAGAAGCTCCCAACCAAGC
TCTCTTGAGGATCTTGAAGGAAACTGAATTCAAAAAGATCAAAGTGCCTGGGCTCCGGTGCCTTCGGCACGGTGTATAAGTAAGGTC
CCTGGCACAGGCTCTGGGCTGGGCCGAGGGCCTCTCATGGTCTGGTGGGAGCCCAGAGTCCTTGAAGCTGTATATTTCCATCA
TCTACTTTACTCTTTGTTTCACTGAGTGTGG

10

E18-1-3F: CCCAGCTTGTGGAGCCTC
E18-1-3R: GACAAGAACACAGAGACAAGGGT
E18-2-F: GCAGGGCCTCTCATGGTC
E18-2-R: CCTGTGCCAGGGACCTTAC

EXON19 (99 pb)

15

GCAATATCAGCCTTAGGTGCGGCTCCACAGCCCCAGTGTCCCTCACCTTCGGGGTGCATCGCTGGTAACATCCACCCAGATCACTGG
GCAGCATGTGGCACCATCTCACAATTGCCAGTTAACGTCTTCCTTCTCTCTGTGCATAG GGACTCTGGATCCAGAAAGGTGAGAAA
GTTAAAAATCCCGTCGCTATCAAGGAATTAAGAGAAGCAACATCTCCGAAAGCCAAGGAAATCCTCGATGTGAGTTTCTGCTTT
GCTGTGTGGGGTCCATGGCTCTGAACCTCAGGCCACCTTTTCTCATGTCTGGCAGCTGCTCTGCTCTAGACCCTGCTCATCTCCA
CATCCTAAATGTTCACTTTCTATG

E19-1-F: ACGTCTTCCTTCTCTCTGTCA
E19-1-R: GTGAGATGGTGCCACATGCT
E19-2-F: GTCCATGGCTCTGAACCTCA
E19-2-R: CCACACAGCAAAGCAGAAAC

EXON20 (186 pb)

20

CCATGAGTACGTATTTGAAACTCAAGATCGCATTATGCGTCTTACCTGGAAGGGTCCATGTGCCCTCCTTCTGGCCACCATG
CGAAGCCACACTGACGTGCCTCTCCCTCCCTCCAG GAAGCCTACGTGATGGCCAGCGTGGACAACCCCCAGTGTGCCGCTGCTGG
GCATCTGCCCTACCTCCACCGTGCAGCTCATCAGCAGCTCATGCCCTTCGGCTGCCTCCTGGACTATGTCCGGGAACACAAAGACA
ATATTGGCTCCAGTACCTGCTCAACTGGTGTGTGCAGATCGCAAAGTAATCAGGGAAGGGAGATACGGGGAGGGGAGATAAGGAG
CCAGGATCCTCACATGCGGTCTGCGCTCCTGGGATAGCAAGAGTTTGCATGGGGATATG

E20-1-F: CACACTGACGTGCCTCTCC
E20-1-R: CTTGCGCATGGTGGCCAGA

ES 2 742 310 T3

E20-2-1F: CCTCCCCGTATCTCCCT
 E20-2-1R: GGAGATAAGGAGCCAGGAT

EX0N21 (156 pb)

CTAACGTTCCGCCAGCCATAAGTCCTCGACGTGGAGAGGCTCAGAGCCTGGCATGAACATGACCCTGAATTCGGATGCAGAGCTTCTT
CCCATGATGATCTGTCCCTCACAGCAGGGTCTTCTCTGTTTCAG *GGCATGAACTACTTGGAGGCCGTCGCTTGGTGCACCGCGACC*
TGGCAGCCAGGAACGTACTGGTGA AAAACACCGCAGCATGTCAAGATCACAGATTTTGGGCTGGCCAAACTGCTGGGTCCGGAAGAGA
AAGAATACCATGCAGAAGGAGGCAAAGTAAGGAGGTGGCTTTAGGTACGCCAGCATTTCCTGACACCAGGGACCAGGCTGCCTTCC
 CACTAGCTGTATTGTTTAAACACATGCAGGGGAGGATGCTCTCCAGACATTCTGGGTGAGCTCGCAGC

5

E21-1-F: AGCAGGGTCTTCTCTGTTTCA
 E21-1-R: GAGGGACAGATCATCATGGG
 E21-2-F: TTTCTGACACCAGGGACC
 E21-2-R: TGACCTAAAGCCACCTCCTT

2. Los ADN plasmáticos libres se extrajeron de 2 ml de plasma.

3. Rellenado de los extremos:

10 La mezcla de reacción se preparó del siguiente modo:

Tabla 1	
Solución de ADN plasmático	38,5 µl
Tampón de fosforilación de ADN T4 (10X)	5 µl
Mezcla de dNTP 10 mM	2 µl
ADN polimerasa T4	2 µl
ADN fosforilasa T4	2 µl
Enzima Klenow	0,5 µl
H ₂ O estéril	0 µl
Volumen total	50 µl

15 La mezcla se depositó en un baño caliente a 20 °C durante 30 minutos. La muestra de ADN se purificó en una columna de purificación y se eluyó con 42 µl de dH₂O estéril o con un tampón de elución.

4. Adición de cola de poli-adenina en el extremo 3' de los fragmentos de ADN:

La mezcla de reacción se preparó del siguiente modo:

Tabla 2	
ADN con los extremos rellenos	32 µl
Tampón de reacción Klenow (10X)	5 µl
Solución de dATP	10 µl
Ex-enzima de Klenow (sin actividad exonucleasa 3'-5')	3 µl
H ₂ O estéril	0 µl
Volumen total	50 µl

20

La mezcla se depositó en un baño caliente a 37 °C durante 30 minutos. La muestra de ADN se purificó en una columna y se eluyó con 25 µl de dH₂O estéril o un tampón de elución.

5. Conexión de conectores a los fragmentos de ADN

La mezcla de reacción se preparó del siguiente modo:

25

Tabla 3	
ADN con extremos rellenos y con colas de dA	33 µl
Tampón de reacción para conexión rápida (5X)	10 µl
5 µm de conector de ADN	2 µl
ADN ligasa rápida T4 (NEB)	5 µl
Volumen total	50 µl

La mezcla se depositó en un baño caliente a 20 °C durante 15 minutos. La muestra de ADN se purificó en una columna de Qiagen y se eluyó con 25 µl de dH₂O estéril o con un tampón de elución.

6. Enriquecimiento de los fragmentos de ADN modificados por los conectores mediante pre-amplificación por PCR

30 La mezcla de reacción de PCR se preparó del siguiente modo:

ES 2 742 310 T3

Tampones EB	14 uL
Tampón de ligasa Taq 10X	5 uL
Oligonucleótido dividido (10 µM)	4 uL
Pre-lib	25 uL
Ligasa Taq	2 uL
Volumen total	50 uL

Programas de PCR:

95 °C	30 s	30 ciclos
50 °C	2 min	
4 °C	parada	

5		Tabla 4
	ADN	12,5 µl
	ADN polimerasa Phusion (mezcla de ADN polimerasa Phusion)	25 µl
	Mezcla de cebadores de PCR	2 µl
	Agua ultrapura	10,5 µl
	Volumen total	50 µl

Amplificación utilizando los siguientes programas de PCR:

- 10
- a. 98 °C 30 s;
 - b. 18 ciclos del siguiente modo:
98 °C 10 s, 65 °C 30 s, 72 °C 30 s;
 - c. 72 °C 5 min;
 - d. preservado a 4 °C

- 15
6. El producto de PCR se analizó por electroforesis en gel de agarosa al 2%, y los resultados se muestran en la Figura 4. El gel con una banda diana de 200 pb se cortó y se recicló usando un kit Qiagen (Figura 4) y se eluyó con 20 µl de tampón de elución.

7. Ciclación

El sistema de ciclación se preparó del siguiente modo (Tabla 5)

20

Plantilla de ADN	12 ul
Tampón de reacción Circ Ligase II 10X	2 ul
MnCl ₂ 50 mM	1 ul
Betaína 5 M (opcional):	4 ul
ligasa de ADNcs Circ Ligasa II (100 U)	1 ul
Volumen total	20 ul

Condiciones de reacción

60 °C	1 h
80 °C	10 min
4 °C	parada

8. Digestión enzimática:

- 25 todos los productos de ciclación fueron digeridos por Exo III, y el sistema de digestión fue

Tampón NE 10 x	1,2 ul
----------------	--------

ES 2 742 310 T3

ExoIII	1 ul
Producto de ciclación	10 ul

El sistema de digestión se depositó en una máquina de PCR y se hizo reaccionar durante 30 minutos a 37 °C. El producto digerido se purificó en una columna de purificación y se disolvió en 30 µl de tampón EB. La concentración se midió mediante Qubit, y los resultados fueron los siguientes:

5

Número de muestra	Concentración
2	4 ng/ul
3	0,3 ng/ul
4	2,33 ng/ul
5	5,74 ng/ul

Los ejemplos de la presente invención usaron una ciclación mediada por puente de ADN, la cual tiene una alta tasa de ciclación. Los resultados detallados del seguimiento y detección de cada paso durante el proceso de ciclación se muestran en la Figura 3. El carril 3 muestra el producto de ciclación después de la digestión por exonucleasa, que es el producto diana. La tasa de ciclación fue de alrededor del 20 %.

10

9. Análisis mediante PCR inversa de la zona diana utilizando cebadores consecutivos
La reacción de PCR se preparó del siguiente modo (Tabla 6)

dd H ₂ O	13 o 18 ul
Mezcla de reacción AmpliTaq Gold® 360 (2 x)	25 ul
CycEGFR18-F1	1 ul
CycEGFR18-R1	1 ul
ADN	10 o 5 ul
Volumen total	50 ul
Nota: los controles fueron P (con cebadores y sin plantilla) y N (la plantilla es ssCyc Lib no ciclada)	

15 Condiciones de la reacción de PCR (Tabla 7)

95 °C	10 min	1 ciclo
95 °C	30 s	30 ciclos
55 °C	30 s	
72 °C	30 s	
72 °C	5 min	1 ciclo

10. La segunda ronda de PCR

La segunda ronda PCR se realizó utilizando los productos de la PCR inversa como plantillas. Sistema (Tabla 8)

20

Mezcla maestra de PCR Phusion (2 x)	25 uL
P5-B1-F (10 mM)	1 uL
Cebador 2-índice 1-2 (10 uM)	1 uL
Productos de la PCR inversa	5 uL
ddH ₂ O	agregar hasta 50 uL

Programas:

98 °C	30 s	1 ciclo
98 °C	30 s	12 ciclos
65 °C	30 s	
72 °C	30 s	
72 °C	5 min	1 ciclo

Se analizaron 10 µl de producto de PCR de la segunda ronda de PCR por electroforesis en gel de agarosa al 2%, y los resultados se muestran en la Figura 4.

11. Los 40 µL restantes del producto de PCR de la segunda ronda de PCR se purificaron en una columna QIAGEN y se disolvieron en 20 µl de tampón EB para generar la biblioteca final.

12. Después del control de calidad, la biblioteca generada fue secuenciada por extremos dobles de 250 pb mediante IlluminaMiseq.

13. Cada secuencia secuenciada por extremos dobles de alto rendimiento se ensambló en una secuencia basada en regiones repetidas. Se eliminaron los conectores y se restableció la secuencia a la secuencia de plantilla original, que luego se comparó con el genoma humano (hg 19). El conjunto único de secuencias de plantillas se contó comparando las coordenadas iniciales y terminales y las secuencias de etiquetado de la secuencia de plantillas en el genoma. Usando la secuencia de plantilla única, se calculó entonces la cobertura del genoma, que puede usarse para evaluar la especificidad de la biblioteca y calcular las mutaciones de células somáticas en la región EGFR.

15 Los resultados se compararon con un kit de detección de mutaciones genéticas EGFR de AmoyDx (dirigido al mismo tejido canceroso) y una PCR digital. Los resultados de la comparación fueron:

Número de muestra	Resultados ARMS (kit AmoyDx)	Resultados de ddPCR (gotas de aceite positivas/gotas de aceite totales)	Resultados de secuenciación (plantillas positivas/plantillas totales)
LC113	19-del	19-del (2/723)	19-del c.2239_2251>C (7/1192)
LC314	WT	WT (del:0/34; 1858r:0/178; t790: 0/208)	WT (0/876)
LC320	19-del	19-del (1/161)	19-del c.2236_2250del15 (9/909)
LC2	WT	n/a	WT (0/668)
LC3	L858R	L858R (589/858)	c.2573T>G; p.L858R (1599/2510)

5 muestras analizadas por el método de la presente invención mostraron resultados altamente consistentes con las analizadas por otros métodos. En la AMRS-PCR (sistema de mutación refractario a la amplificación) la ADN polimerasa Taq carece de actividad de exonucleasa 3'-5'. En determinadas condiciones, solo se produce una amplificación eficaz cuando la última base en el extremo 3' de los cebadores de PCR es complementaria a la plantilla de ADN. Los genes mutados y los genes de tipo salvaje se pueden distinguir directamente por PCR usando cebadores adecuados dirigidos a diferentes mutaciones conocidas. Este método es principalmente para muestras de biopsia y FFPE. Se proporcionaron 5 ml de sangre periférica antes de la cirugía y muestras de FFPE después de la cirugía del mismo paciente. El tipo de mutación de las muestras de FFPE en los sitios conocidos se detectó por ARMS-PCR, y luego el ADN plasmático se detectó por PCR digital y por el método de la presente invención, respectivamente.

Ejemplo 2

La fiabilidad del presente método se verificó mediante la detección de sitios de mutación de cáncer conocidos. El método fue el mismo que el descrito en el Ejemplo 1.

ADN de línea celular tumoral: se rompieron ADN celulares que contenían la mutación híbrida c.2235_2249del15 (exón 19) por ultrasonificación, se reciclaron fragmentos de 166 ± 10 pb y se mezclaron con ADN plasmáticos de una persona normal en una cierta proporción. Se examinó la sensibilidad y la estabilidad del método.

Los resultados de detección de c.2235_2249del15 fueron los siguientes:

N.º de muestra.	Predicción	N.º de plantillas totales	N.º de plantillas Non-Del	N.º de plantillas Del	proporción
Muestra 1	Mutación 0 %	862	862	0	0,00 %
Muestra 2	Mutación 0,1	438	433	5	1,09 %

(continuación)

N.º de muestra.	Predicción	N.º de plantillas totales	N.º de plantillas Non-Del	N.º de plantillas Del	proporción
-----------------	------------	---------------------------	---------------------------	-----------------------	------------

ES 2 742 310 T3

Muestra 3	Mutación 1 %	905	843	62	6,78 %
Muestra 4	Mutación 5 %	880	786	94	10,67 %
Muestra 5	Mutación 25 %	1631	1017	614	37,65 %

LISTADO DE SECUENCIAS

5	<110> BERRY GENOMICS CO., LTD		
	<120> Un método y un kit para la detección no invasiva de mutaciones en el gen EGFR		
	<130> CIE140060PCT <160> 20		
10	<170> PatentIn versión 3.2		
	<210> 1		
	<211> 18		
	<212> ADN		
15	<213> Artificial		
	<400> 1		
	cccagcttgt ggagcctc	18	
20	<210> 2		
	<211> 23		
	<212> ADN		
	<213> Artificial		
25	<400> 2		
	gacaagaaca cagagacaag ggt	23	
	<210> 3		
	<211> 18		
30	<212> ADN		
	<213> Artificial		
	<400> 3		
35	gcagggcctc tcatggtc	18	
	<210> 4		
	<211> 19		
	<212> ADN		
	<213> Artificial		
40	<400> 4		
	cctgtgccag ggaccttac	19	
	<210> 5		
45	<211> 24		
	<212> ADN		
	<213> Artificial		
	<400> 5		
50	acgtcttctt tctctcttg teat	24	
	<210> 6		
	<211> 20		
	<212> ADN		
55	<213> Artificial		
	<400> 6		
	gtgagatggt gccacatgct	20	
60	<210> 7		
	<211> 20		
	<212> ADN		

ES 2 742 310 T3

	<213> Artificial	
5	<400> 7 gtccatggct ctgaacctca	20
	<210> 8 <211> 20 <212> ADN <213> Artificial	
10	<400> 8 ccacacagca aagcagaaac	20
	<210> 9 <211> 19 <212> ADN <213> Artificial	
15	<400> 9 cacactgacg tgcctctcc	19
	<210> 10 <211> 18 <212> ADN <213> Artificial	
20	<400> 10 cttcgcatgg tggccaga	18
	<210> 11 <211> 17 <212> ADN <213> Artificial	
25	<400> 11 cctccccgta tctcct	17
	<210> 12 <211> 19 <212> ADN <213> Artificial	
30	<400> 12 ggagataagg agccaggat	19
	<210> 13 <211> 21 <212> ADN <213> Artificial	
35	<400> 13 agcaggtct tctctgttc a	21
	<210> 14 <211> 20 <212> ADN <213> Artificial	
40	<400> 14 gagggacaga tcatcatggg	20
	<210> 15 <211> 19 <212> ADN <213> Artificial	
45		
50		
55		
60		
65		

ES 2 742 310 T3

	<400> 15		
	tttctgaca ccagggacc	19	
5	<210> 16 <211> 20 <212> ADN <213> Artificial		
10	<400> 16		
	tgacctaaag ccacctcctt	20	
15	<210> 17 <211> 381 <212> ADN <213> Ser humano		
	<400> 17		
	caagtgccgt gtcctggcac ccaagcccat gccgtggctg ctgggtcccc tgctgggcca	60	
	tgtctggcac tgctttccag catgggtgagg gctgaggtga cccttgtctc tgtgttcttg	120	
	tccccccag cttgtggagc ctcttacacc cagtggagaa gctcccaacc aagctctctt	180	
	gaggatcttg aaggaaactg aattcaaaaa gatcaaagtg ctgggctccg gtgcgttcgg	240	
	cacggtgtat aaggtaaggt ccctggcaca ggcctctggg ctgggcccga gggcctctca	300	
	tggctctggtg gggagcccag agtccttgca agctgtatat ttccatcatc tactttactc	360	
	tttgtttcac tgagtgtttg g	381	
20	<210> 18 <211> 372 <212> ADN <213> Ser humano		
25	<400> 18		
	gcaatatcag ccttaggtgc ggctccacag ccccagtgtc cctcaccttc ggggtgcatc	60	
	gctggtaaca tccaccaga tcaactgggca gcatgtggca ccatctcaca attgccagtt	120	
	aacgtcttcc ttctctctct gtcatagga ctctggatcc cagaaggtga gaaagttaa	180	
	attcccgtcg ctatcaagga attaagagaa gcaacatctc cgaaagcaa caaggaaatc	240	
	ctcgatgtga gtttctgctt tgctgtgtgg ggggccatgg ctctgaacct caggcccacc	300	
	ttttctcatg tctggcagct gctctgctct agaccctgct catctccaca tcctaaatgt	360	
30	tcactttcta tg	372	
35	<210> 19 <211> 408 <212> ADN <213> Ser humano		
	<400> 19		

ES 2 742 310 T3

ccatgagtac gtatTTTgaa actcaagatc gcattcatgc gtcttcacct ggaaggggtc 60
 catgtgcccc tccttctggc caccatgcga agccacactg acgtgcctct ccctccctcc 120
 aggaagccta cgtgatggcc agcgtggaca acccccacgt gtgccgcctg ctgggcatct 180
 gcctcacctc caccgtgcag ctcatcacgc agctcatgcc cttcggctgc ctctggact 240
 atgtccggga acacaaagac aatattggct cccagtacct gctcaactgg tgtgtgcaga 300
 tcgcaaaggt aatcagggaa gggagatacg gggaggggag ataaggagcc aggatcctca 360
 catgCGgtct gcgctcctgg gatagcaaga gtttgccatg gggatatg 408

<210> 20
 <211> 415
 <212> ADN
 <213> Ser humano

5

<400> 20

ctaacgttcg ccagccataa gtcctcgacg tggagaggct cagagcctgg catgaacatg 60
 accctgaatt cggatgcaga gcttcttccc atgatgatct gtccctcaca gcagggctctt 120
 ctctgtttca gggcatgaac tacttggagg accgtcgctt ggtgcaccgc gacctggcag 180
 ccaggaacgt actggtgaaa acaccgcagc atgtcaagat cacagatTTT gggctggcca 240
 aactgctggg tgcggaagag aaagaatacc atgcagaagg aggcaaagta aggaggtggc 300
 tttaggtcag ccagcatttt cctgacacca gggaccaggc tgccttccca ctagctgtat 360
 tgtttaacac atgcagggga ggatgctctc cagacattct gggtgagctc gcagc 415

10

REIVINDICACIONES

1. Un método para detectar de forma no invasiva mutaciones del gen EGFR en un sujeto, que comprende las siguientes etapas:
- 5 diseñar cebadores de acuerdo con exones del gen EGFR;
 extraer ADN plasmáticos en una muestra de plasma obtenida del sujeto;
 conectar los ADN plasmáticos extraídos con conectores de etiquetado;
 pre-amplificar por PCR los ADN plasmáticos conectados con conectores de etiquetado;
 10 ciclar los ADN pre-amplificados para obtener ADN ciclados;
 amplificar por PCR los ADN ciclados utilizando los cebadores diseñados; y
 realizar secuenciación de alto rendimiento del producto amplificado por PCR y analizar las mutaciones del gen EGFR.
- 15 2. El método de acuerdo con la reivindicación 1, **caracterizado por que** los cebadores diseñados son un par de cebadores que están adyacentes y están ampliados de forma regresiva.
3. El método de acuerdo con la reivindicación 2, **caracterizado por que** los cebadores del par de cebadores ampliados de forma regresiva se encuentran en el extremo 5' o 3' de los exones del gen EGFR.
- 20 4. El método de acuerdo con la reivindicación 2, **caracterizado por que** el espacio del par de cebadores ampliados de forma regresiva es 0-1/2 del total de pares de bases de los ADN.
5. El método de acuerdo con la reivindicación 3, **caracterizado por que** los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de los genes EGFR.
- 25 6. El método de acuerdo con la reivindicación 3, **caracterizado por que** el extremo 5' de los cebadores ampliados de forma regresiva contiene secuencias conectoras para la biblioteca de secuenciación de alto rendimiento.
- 30 7. El método de acuerdo con la reivindicación 3, **caracterizado por que** los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20, o exón 21 de los genes EGFR, y las secuencias de los cebadores son las siguientes:
- E18-1-3F: CCCAGCTTGTGGAGCCTC
 E18-1-3R: GACAAGAACACAGAGACAAGGGT
 E18-2-F: GCAGGGCCTCTCATGGTC
 E18-2-R: CCTGTGCCAGGGACCTTAC
 E19-1-F: ACGTCTTCCTTCTCTCTGTGCAT
 E19-1-R: GTGAGATGGTGCCACATGCT
 E19-2-F: GTCCATGGCTCTGAACCTCA
 E19-2-R: CCACACAGCAAAGCAGAAAC
 E20-1-F: CAACTGACGTGCCTCTCC
 E20-1-R: CTTCGCATGGTGGCCAGA
 E20-2-1F: CCTCCCCGTATCTCCCT
 E20-2-1R: GGAGATAAGGAGCCAGGAT
 E21-1-F: AGCAGGGTCTTCTCTGTTTCA
 E21-1-R: GAGGGACAGATCATCATGGG
 E21-2-F: TTTCTGACACCAGGGACC
 E21-2-R: TGACCTAAAGCCACCTCCT.
- 35 8. El método de acuerdo con la reivindicación 1, **caracterizado por que** la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN.
9. El método de acuerdo con la reivindicación 1, **caracterizado por que** los genes EGFR en los ADN plasmáticos tienen mutaciones de tipo inserción, supresión, sustitución o fusión génica.
- 40 10. El método de acuerdo con la reivindicación 1, **caracterizado por que** las tecnologías de secuenciación de alto rendimiento se seleccionan de Roche/454 FLX, Illumina /Hiseq/Miseq, Applied Biosystems SOLID y Life Technologies/Ion Torrent/Proton.
- 45 11. Un kit para la detección no invasiva de mutaciones del gen EGFR, que comprende:
 cebadores para regiones de pre-amplificación a analizar en los genes EGFR,

una ADN ciclasa y

cebadores para amplificar los ADN diana, en donde los cebadores para amplificar los ADN diana son cebadores diseñados de acuerdo con los exones del gen EGFR.

5 12. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** los cebadores para amplificar los ADN diana son un par de cebadores que están adyacentes y están ampliados de forma regresiva.

10 13. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** los cebadores del par de cebadores ampliados de forma regresiva se encuentran en el extremo 5' o 3' de los sitios o regiones a analizar en los genes EGFR.

15 14. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** el espacio del par de cebadores ampliado de forma regresiva es 0-1/2 del total de pares de bases de los ADN plasmáticos.

20 15. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20 o exón 21 de los genes EGFR.

16. El kit de acuerdo con la reivindicación 14 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** los cebadores ampliados de forma regresiva se dirigen al exón 18, exón 19, exón 20, o exón 21 de los genes EGFR, y las secuencias de los cebadores son las siguientes:

E18-1-3F: CCCAGCTTGTGGAGCCTC
 E18-1-3R: GACAAGAACACAGAGACAAGGGT
 E18-2-F: GCAGGGCCTCTCATGGTC
 E18-2-R: CCTGTGCCAGGGACCTTAC
 E19-1-F: ACGTCTTCCTTCTCTCTCTGTCAT
 E19-1-R: GTGAGATGGTGCCACATGCT
 E19-2-F: GTCCATGGCTCTGAACCTCA
 E19-2-R: CCACACAGCAAAGCAGAAAC
 E20-1-F: CACACTGACGTGCCTCTCC
 E20-1-R: CTTCGCATGGTGGCCAGA
 E20-2-1F: CCTCCCCGTATCTCCCT
 E20-2-1R: GGAGATAAGGAGCCAGGAT
 E21-1-F: AGCAGGGTCTTCTCTGTTTCA
 E21-1-R: GAGGGACAGATCATCATGGG
 E21-2-F: TTTCTGACACCAGGGACC
 E21-2-R: TGACCTAAAGCCACCTCCTT.

25 17. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** el extremo 5' de los cebadores ampliados de forma regresiva contiene secuencias conectoras para la biblioteca de secuenciación de alto rendimiento.

30 18. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** los conectores contienen secuencias de etiquetado.

35 19. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** los ADN plasmáticos se pre-amplifican antes de ciclarse.

20. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** la ciclación es una ciclación de ADN de cadena sencilla mediada por puente de ADN.

40 21. El kit de acuerdo con la reivindicación 11 para la detección no invasiva de mutaciones del gen EGFR, **caracterizado por que** los sitios o regiones a analizar en los genes EGFR tienen mutaciones de tipo inserción, supresión, sustitución o fusión génica.

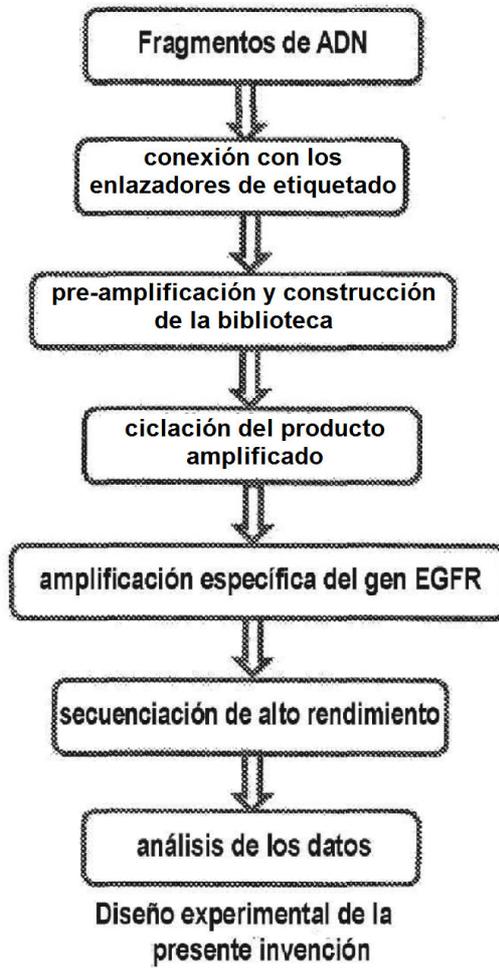


Fig. 1

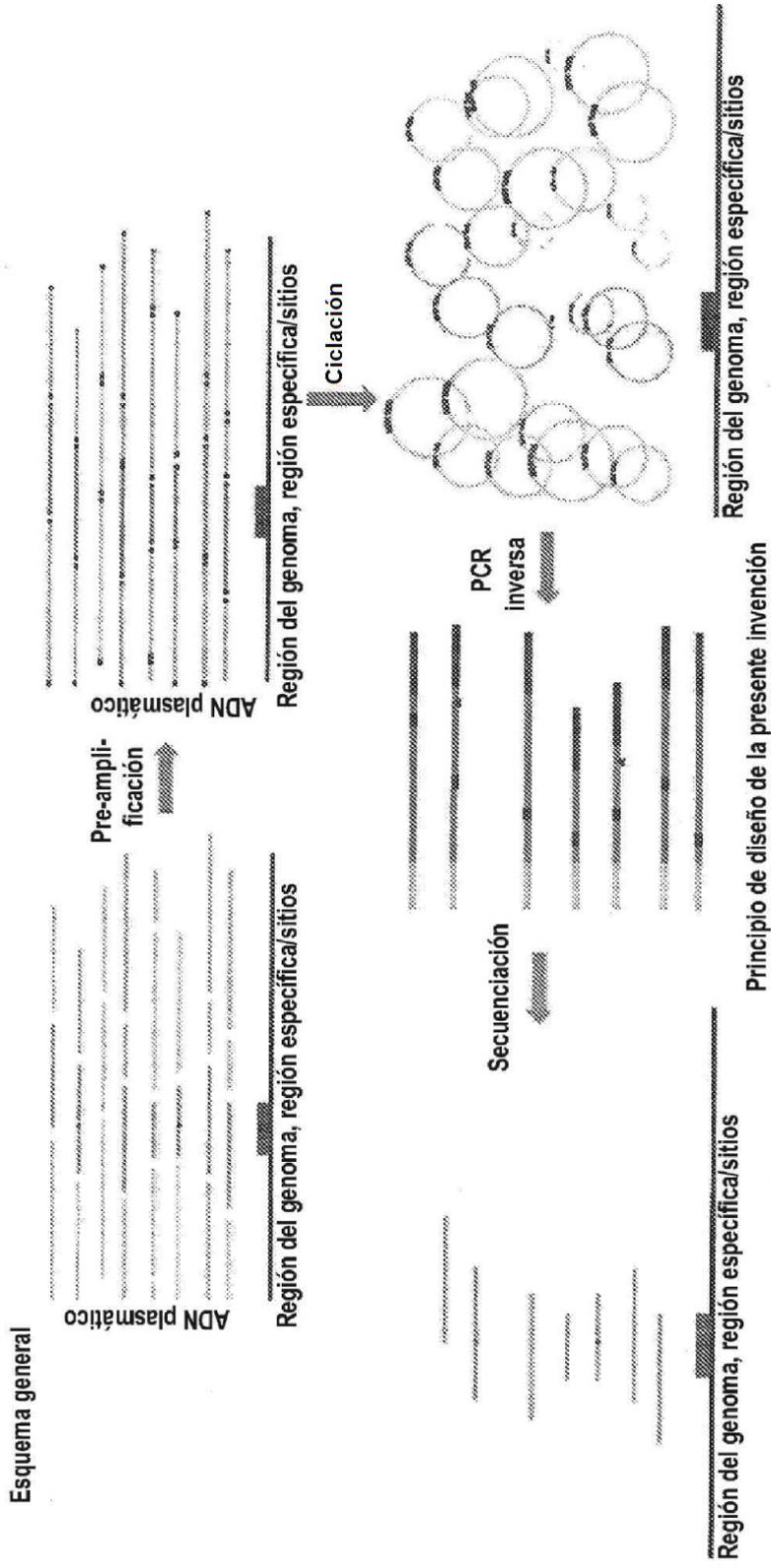
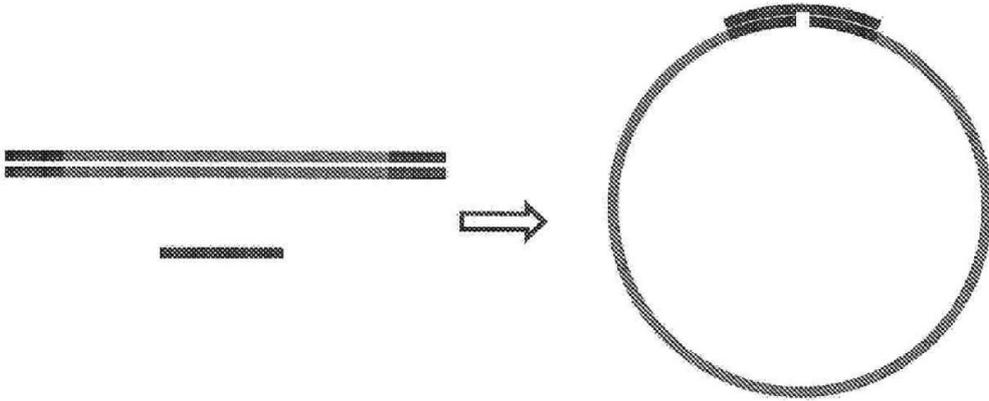


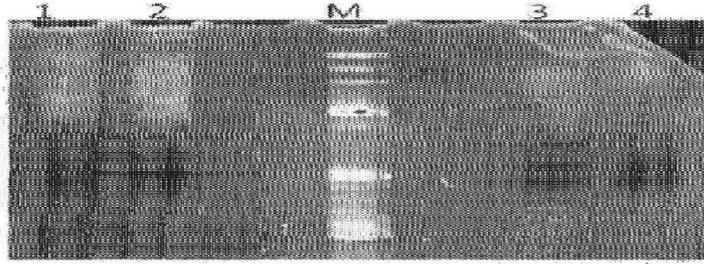
Fig. 2A

Principio de ciclación



Principio de diseño de la presente invención

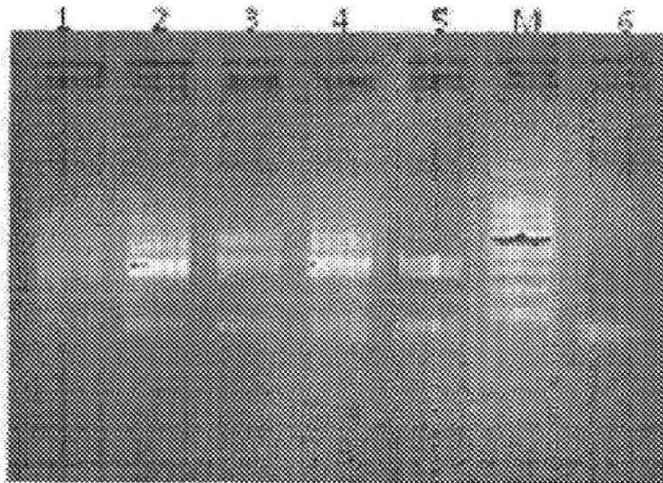
Fig. 2B



- 1: Producto de PCR antes de la ciclación
2. Producto después de la ciclación (no digerido por una exonucleasa) M: marcador
3. Producto digerido por una exonucleasa después de la ciclación (no purificado)
4. Producto purificado y digerido por una exonucleasa después de la ciclación

Tasa de ciclación de la presentación invención (gel desnaturalizante 6 %)

Fig. 3



Biblioteca final de la muestra de EGFR (gel de agarosa 2 %)

Fig. 4

1: LC2; 2: LC3; 3: LC113; 4: LC314; 5: LC320; 6: control negativo;
M: escalera de ADN de 100 pb (TaKaRa)