

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 533 767**

51 Int. Cl.:

C12Q 1/68 (2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **17.04.2006 E 08021839 (9)**

97 Fecha y número de publicación de la concesión europea: **31.12.2014 EP 2045335**

54 Título: **Métodos para el análisis de trastornos proliferativos celulares**

30 Prioridad:

15.04.2005 US 672242 P
02.05.2005 US 676997 P
08.07.2005 US 697521 P
01.08.2005 US 704860 P
17.08.2005 US 709318 P
04.10.2005 US 723602 P
30.03.2006 US 787402 P

45 Fecha de publicación y mención en BOPI de la traducción de la patente:
15.04.2015

73 Titular/es:

EPIGENOMICS AG (100.0%)
Geneststrasse 5
10829 Berlin, DE

72 Inventor/es:

DISTLER, JÜRGEN;
HILDMANN, THOMAS;
LESCHE, RALF;
LOFTON-DAY, CATHY;
MODEL, FABIAN;
SCHUSTER, MATTHIAS;
SLEDZIEWSKI, ANDREW;
SONG, XIAOLING y
TETZNER, REIMO

74 Agente/Representante:

LAZCANO GAINZA, Jesús

ES 2 533 767 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

Métodos para el análisis de trastornos proliferativos celulares

5 Campo de la invención

La presente invención se refiere a un método para detectar y/o clasificar el carcinoma colorrectal en un sujeto que comprende determinar la presencia o ausencia de metilación CpG de FOXL2 en una muestra biológica aislada de dicho sujeto en donde la hipermetilación CpG de FOXL2 es indicativa de la presencia o clase de dicho carcinoma colorrectal.

10

Antecedentes

15

Incidencia y diagnóstico de cáncer. El cáncer es la segunda causa principal de muerte de los Estados Unidos. Las tasas de mortalidad se podrían mejorar significativamente si los métodos de detección actuales se mejoraran en términos de facilidad de detección, sensibilidad y conformidad del paciente. Los métodos actuales que se recomiendan para el diagnóstico de cáncer frecuentemente son caros y no son adecuados para su aplicación como pruebas de detección a gran escala para la población.

20

El cáncer hepatocelular (HCC) es el cuarto cáncer más común en el mundo, su incidencia varía de 2.1 por 100,000 en Norteamérica a 80 por 100,000 en China. En los Estados Unidos, se estima que habrá 17,550 nuevos casos diagnosticados en 2005 y 15,420 muertes por esta enfermedad. La ecografía del hígado, alfafetoproteína y CT convencional se obtienen regularmente en la evaluación diagnóstica del HCC (cáncer hepatocelular o cáncer primario de hígado), pero frecuentemente son demasiado insensibles para detectar pequeñas lesiones multifocales y para la planificación del tratamiento.

25

En los Estados Unidos la incidencia anual de cáncer colorrectal es de aproximadamente 150,000, y 56,600 personas mueren de cáncer colorrectal cada año. El riesgo de cáncer colorrectal en la población general es de aproximadamente 5 a 6 por ciento. A pesar de los intensos esfuerzos de los últimos años en la revisión y la detección precoz del cáncer de colon, hasta hoy en día la mayoría de los casos se diagnostican en una etapa avanzada con metástasis regional o distante. Aunque las opciones terapéuticas incluyen cirugía y quimioterapia adyuvante o paliativa, la mayoría de los pacientes mueren a causa de la progresión de su cáncer en unos meses. La identificación de los cambios moleculares que subyacen en el desarrollo del cáncer de colon puede ayudar a desarrollar nuevos controles, detección, diagnóstico y opciones terapéuticas que pueden mejorar el mal pronóstico global de estos pacientes.

30

35

Las directrices actuales para la detección colorrectal según la Sociedad Americana del Cáncer utiliza una de cinco opciones diferentes para el examen en personas de riesgo promedio de 50 años de edad o más. Estas opciones incluyen: 1) análisis de sangre oculta en heces (FOBT) cada año, 2) sigmoidoscopia flexible cada cinco años, 3) FPBT anual más sigmoidoscopia flexible cada cinco años, 4) enema de bario de doble contraste (DCBE) cada cinco años o 5) colonoscopia cada diez años. A pesar de que estos procedimientos de prueba se aceptan bien por la comunidad médica, no se realiza la aplicación de la revisión generalizada para el cáncer colorrectal. La conformidad del paciente es un factor importante para el uso limitado debido a las molestias o inconvenientes que se asocian con los procedimientos. La prueba FOBT, aunque es un procedimiento no invasivo, requiere restricciones en la dieta y otras 3-5 días antes de la prueba. Los niveles de sensibilidad de esta prueba también son muy bajos para el adenocarcinoma colorrectal con una amplia variabilidad en función de la prueba. Las mediciones de sensibilidad para la detección de adenomas son aún menores ya que la mayoría de los adenomas no sangran. Por el contrario, la sensibilidad para los procedimientos más invasivos, tales como la sigmoidoscopia y colonoscopia es bastante alta debido a la visualización directa del lumen del colon. Ningún ensayo aleatorio evaluó la eficacia de estas técnicas, sin embargo, mediante la utilización de datos de estudios de casos y controles y de datos del Estudio Nacional de Pólipos (Estados Unidos) se demostró que la eliminación de los pólipos adenomatosos resulta en una reducción de 76-90% en la incidencia de CRC. La sigmoidoscopia tiene la limitación de sólo visualizar el lado izquierdo del colon dejando lesiones en el colon derecho sin ser detectados. Ambos procedimientos de alcance son caros, requieren una preparación catártica y tienen mayor riesgo de morbilidad y mortalidad. Pruebas mejoradas con mayor sensibilidad, especificidad, facilidad de uso y con disminución de los costos son claramente necesarias antes de la revisión más amplia general para el cáncer colorrectal se convierten en rutina.

40

45

50

55

La detección temprana del cáncer de color se basa generalmente en la prueba de sangre oculta en heces (FOBT) que se realiza anualmente en individuos asintomáticos. Las recomendaciones actuales que se adoptan por varias organizaciones de salud, lo que incluye a la Sociedad Americana del Cáncer, exigen la prueba de sangre oculta en heces a partir de los 50 años de edad, que se repite cada año hasta que el paciente ya no obtenga beneficio de la detección. Una FOBT positiva conduce a una examinación por colonoscopia del intestino; un procedimiento costoso e invasivo, con una tasa de complicaciones graves de una por cada 5,000 exámenes. Sólo el 12% de los pacientes con heces hemo positivas se

60

5 diagnostica con cáncer o pólipos en el momento de la colonoscopia. Varios estudios muestran que la detección por FOBT no mejora la mortalidad relacionada con el cáncer o la supervivencia general. La conformidad con la prueba de sangre oculta es pobre; menos del 20 por ciento de la población se ofrece o completa la FOBT como se recomienda. Si la FOBT se hace correctamente, el paciente recolecta una muestra fecal de tres deposiciones consecutivas. Las muestras se obtienen mientras el paciente se adhiere a directrices dietéticas y evita medicamentos que se conocen por inducir sangrado gastrointestinal oculto. En realidad, los médicos con frecuencia fallan en instruir a los pacientes adecuadamente, los pacientes con frecuencia no se adhieren al protocolo, y algunos pacientes encuentran la tarea de recoger las muestras fecales difícil o desagradable, por lo tanto, la conformidad con la prueba de sangre oculta anual es pobre. Si la comprobación de sensibilidad y especificidad se puede mejorar sobre métodos actuales, la frecuencia de las pruebas podría reducirse, la recolección de muestras consecutivas se eliminaría, las modificaciones de los horarios de medicación y de la dieta se eliminarían, y la conformidad del paciente se mejoraría. Para agravar el problema de la conformidad, la sensibilidad y especificidad de la FOBT para detectar el cáncer de colon es pobre. La pobre especificidad de la prueba conduce a la colonoscopia innecesaria, lo que añade un gasto considerable para la detección del cáncer de colon.

10 La especificidad de la FOBT se calculó del 96% en el mejor de los casos, con una sensibilidad del 43% (adenomas) y 50% (carcinoma colorrectal). La sensibilidad se puede mejorar mediante la utilización de una FOBT por inmunoensayo tal como la que se produce bajo el nombre 'InSure™', con una sensibilidad mejorada del 77% (adenomas) y 88.9% (carcinoma colorrectal).

15 *Marcadores moleculares de la enfermedad.* Los marcadores moleculares de la enfermedad ofrecen varias ventajas sobre otros tipos de marcadores, una ventaja es que incluso muestras de tamaños muy pequeños y/o muestras cuyas arquitecturas de tejido no se mantienen pueden analizarse de forma eficiente. En la última década se demostró que un número de genes se expresan diferencialmente entre carcinomas de colon y normales. Sin embargo, ningún marcador único o en combinación demostró ser suficiente para el diagnóstico de los carcinomas de colon. Se demostró recientemente que los enfoques que se basan en ARNm de alta dimensión son capaces de proporcionar un mejor medio para distinguir entre diferentes tipos de tumores y lesiones benignas y malignas. Sin embargo, su aplicación como una herramienta de diagnóstico de rutina en un ambiente clínico se ve impedida por la extrema inestabilidad del ARNm, los cambios en la expresión que se producen rápidamente tras ciertos desencadenantes (*por ejemplo*, recolección de las muestras), y, lo más importante por la gran cantidad de ARNm necesario para el análisis (Lipshutz, R. J. y otros, Nature Genetics 21:20-24, 1999; Bowtell, D. D. L. Nature genetics suppl. 21:25-32, 1999), el cual frecuentemente no puede obtenerse a partir de una biopsia de rutina.

20 Se sugirió el uso de marcadores biológicos para mejorar más aun la sensibilidad y especificidad de la FOBT, ejemplos de tales pruebas incluyen el ensayo de análisis de heces PreGen-Plus™ disponible de EXACT Sciences que tiene una sensibilidad del 20% (adenoma) y 52% (carcinoma colorrectal) y una especificidad del 95% en ambos casos. Esta prueba analiza la presencia de 23 mutaciones en el ADN que se asocian con el desarrollo de neoplasias de colon. Se conoce el uso de la metilación del ADN como marcador de cáncer de colon. Por ejemplo, Sabbioni y otros (Molecular Diagnosis 7:201-207, 2003) detectaron hipermetilación de un panel de genes que consiste en TPEF, HIC1, DAPK y MGMT en sangre periférico en 98% de los pacientes de carcinoma de colon. Sin embargo, esto proporciona una base adecuada para un ensayo comercialmente vendible, ya que la especificidad de una prueba de este tipo también debe ser suficientemente alta.

25 El modelo actual de la patogénesis colorrectal favorece una progresión paso a paso de adenomas, que incluye el desarrollo de displasia y finalmente signos de cáncer invasivo. Los cambios moleculares que subyacen a esta secuencia adenoma-carcinoma incluyen alteraciones genéticas y epigenéticas de genes supresores de tumores (APC, p53, DCC), la activación de oncogenes (K-ras) y la inactivación de genes de reparación de desajuste de ADN. Recientemente, se revelaron nuevos cambios moleculares y defectos genéticos. Por lo tanto, la activación de la vía de señalización Wnt no sólo incluye las mutaciones del gen APC, sino que también puede resultar de mutaciones en la β -catenina. Además, las alteraciones en la vía de señalización del TGF- β junto con sus transductores de señal SMAD4 y SMAD2 se relacionaron con el desarrollo de cáncer de colon.

30 A pesar de los recientes avances en la comprensión de la patogénesis de los adenomas y carcinomas de colon y de sus cambios genéticos y moleculares, los cambios genéticos y epigenéticos implicados en la aparición de la metástasis se conocen menos. Sin embargo, generalmente se acepta bien que el proceso de invasión y proteólisis de la matriz extracelular, así como también la infiltración de la membrana basal vascular implican proteínas adhesivas, tales como los miembros de la familia de los receptores de integrinas, las cadherinas, la superfamilia de las inmunoglobulinas, la proteína de unión a laminina y el receptor CD44. Además de la adhesión, el proceso de formación de metástasis también incluye la inducción y regulación de la angiogénesis (VEGF, bFGF), la inducción de la proliferación celular (EGF, HGF, IGF) y la activación de enzimas proteolíticas (MMPs, Tulips, uPAR), así como también la inhibición de la apoptosis (Bcl-2, Bcl-X). Más recientemente otros grupos compararon los cambios genéticos y moleculares en lesiones metastásicas con los cambios encontrados en cánceres colorrectales primarios. Así, Kleeff y otros informaron de la pérdida de DOC-2, un gen candidato

supresor de tumores, tanto en cáncer colorrectal primario como metastásico. Además, Zauber y otros informaron de que en su serie de 42 cánceres colorrectales las mutaciones Ki-ras en cánceres primarios fueron idénticas en todas las 42 lesiones pareadas primarias y metastásicas sincronas. Similarmente la pérdida de heterocigosidad en el locus APC fue idéntica para 39 metástasis sincronas y carcinomas pareados. Los autores concluyeron que para los genes Ki-ras y APC los cambios genéticos en las metástasis son idénticos a los del cáncer colorrectal primario. Sin embargo, otros grupos encontraron cambios genéticos y moleculares en los cánceres de colon metastásicos, que no están presentes en los cánceres primarios. Así, se informó del desarrollo de LOH del cromosoma 3p en la metástasis colorrectal. Adicionalmente, mediante la utilización de hibridación genómica comparativa se encontraron varias alteraciones en la metástasis de hígado que eran exclusivas de lesiones metastásicas (-9q, -11q, y -17q).

Metilación de la isla CpG. Además de las mutaciones la metilación aberrante de las islas CpG se demostró que conduce al silenciamiento transcripcional de ciertos genes que se relacionaron anteriormente a la patogénesis de diversos tipos de cáncer. Las islas CpG son secuencias cortas que son ricas en dinucleótidos CpG y por lo general se pueden encontrar en la región 5' de aproximadamente el 50% de todos los genes humanos. La metilación de las citosinas en estas islas conduce a la pérdida de la expresión génica y se reportó en la inactivación del cromosoma X y la impronta genómica.

Recientemente, varios grupos también analizaron la metilación de varios genes en el cáncer colorrectal y reportaron el silenciamiento transcripcional por la metilación del promotor para p16INK4, p14ARF, p15INK4b, MGMT, hMLH1, GSTP1, DAPK, CDH1, TIP-3 y APC entre otros. Así además de la inactivación mutacional de ciertos genes, la hipermetilación de estos genes también contribuye significativamente a la patogénesis de esta enfermedad.

En los últimos años varios genes que se metilan en el cáncer de colon se identificaron por MS-APPCR. Este grupo de genes, entre otros, incluye TPEF/HPP1 el cual se metila frecuentemente en los cánceres de colon y que se identificó independientemente por dos grupos diferentes con el uso del método MS-APPCR (ver, *por ejemplo*, Young J, Biden KG, Simms LA, Huggard P, Karamatic R, Eyre HJ, Sutherland GR, Herath N, Barker M, Anderson GJ, Fitzpatrick DR, Ramm GA, Jass JR, Leggett BA. HPP1: a transmembrane protein-encoding gene commonly methylated in colorectal polyps and cancers. Proc Natl Acad Sci USA 98:265-270, 2001).

Enfoque multifactorial. El diagnóstico del cáncer se basa tradicionalmente en la detección de marcadores moleculares individuales (*por ejemplo*, mutaciones genéticas, niveles elevados de PSA). Desafortunadamente, el cáncer es un estado de enfermedad en el cual los marcadores individuales típicamente fracasaron para detectar o diferenciar muchas formas de la enfermedad. Así, los ensayos que reconocen sólo un único marcador demostraron ser de valor predictivo limitado. Un aspecto fundamental de esta invención es que los diagnósticos de cáncer basados en la metilación y la detección, diagnóstico y seguimiento terapéutico de este tipo de enfermedades proporcionarán mejoras significativas sobre el estado de la técnica que utiliza análisis de un solo marcador mediante el uso de una selección de múltiples marcadores. El enfoque analítico multiplexado es particularmente bien adecuado para el diagnóstico del cáncer ya que el cáncer no es una enfermedad simple, este enfoque de "panel" multifactorial es consistente con la naturaleza heterogénea del cáncer, tanto citológicamente como clínicamente.

La clave para la exitosa implementación de un enfoque de panel para las pruebas diagnósticas basadas en la metilación es el diseño y desarrollo de paneles optimizados de marcadores que pueden caracterizar y distinguir estados de enfermedad. La presente invención describe una pluralidad de paneles de genes particularmente eficientes y únicos, el análisis de metilación de uno o una combinación de los miembros del panel permite la detección de trastornos proliferativos celulares de colon con una particularmente alta sensibilidad, especificidad y/o valor predictivo. *Desarrollo de las pruebas médicas.* Dos medidas de evaluación claves de cualquier detección médica o prueba de diagnóstico son su sensibilidad y especificidad, que miden qué tan bien la prueba se desarrolla para detectar con precisión todos los individuos afectados sin excepción, y sin incluir falsamente a personas que no tienen la enfermedad objetivo (valor predictivo). Históricamente, muchas pruebas de diagnóstico se criticaron debido a la pobre sensibilidad y especificidad.

Un resultado verdadero positivo (TP) es cuando la prueba es positiva y la afección está presente. Un resultado falso positivo (FP) es cuando la prueba es positiva pero la afección no está presente. Un resultado verdadero negativo (TN) es cuando la prueba es negativa y la afección no está presente. Un resultado falso negativo (FN) es cuando la prueba es negativa pero la afección no está presente. En este contexto: sensibilidad = $TP/(TP+FN)$; especificidad = $TN/(FP+TN)$; y valor predictivo = $TP/(TP+FP)$.

La sensibilidad es una medida de la capacidad de la prueba para detectar correctamente la enfermedad objetivo en un individuo que se está analizando. Una prueba que tiene poca sensibilidad produce una alta tasa de falsos negativos, es decir, personas que tienen la enfermedad pero que se identifican falsamente como libres de dicha enfermedad. El peligro potencial de un falso negativo es que el individuo enfermo permanecerá sin diagnosticar y sin tratamiento durante cierto período de tiempo, durante el cual la enfermedad puede progresar a una etapa más avanzada en donde los tratamientos, si

- los hay, pueden ser menos eficaces. Un ejemplo de una prueba que tiene una baja sensibilidad es una prueba de sangre a base de proteínas para el VIH. Este tipo de ensayo exhibe pobre sensibilidad porque falla en detectar la presencia del virus hasta que la enfermedad se establece bien y el virus invade el torrente sanguíneo en cantidades por debajo del estándar. En contraste, un ejemplo de una prueba que tiene alta sensibilidad es la detección de carga viral mediante la utilización de la reacción en cadena de la polimerasa (PCR). La alta sensibilidad se logra porque este tipo de prueba puede detectar cantidades muy pequeñas del virus. La alta sensibilidad es particularmente importante cuando las consecuencias de la falta de un diagnóstico son altas.
- La especificidad, por el contrario, es una medida de la capacidad de una prueba para identificar con precisión a los pacientes que están libres del estado de enfermedad. Una prueba que tiene pobre especificidad produce una alta tasa de falsos positivos, es decir, individuos que se identifican falsamente como que tienen la enfermedad. Un inconveniente de los falsos positivos es que obligan a los pacientes a someterse a tratamientos de procedimientos médicos innecesarios con riesgos inherentes, estrés emocional y financiero, y que podrían tener efectos adversos sobre la salud del paciente. Una característica de las enfermedades que hace que sea difícil desarrollar pruebas de diagnóstico con alta especificidad es que los mecanismos de la enfermedad, particularmente en el cáncer, a menudo implican una pluralidad de genes y proteínas. Adicionalmente, ciertas proteínas puede elevarse por razones no relacionadas con un estado de enfermedad. Un ejemplo de una prueba que tiene una alta especificidad es una prueba que se basa en genes que se puede detectar una mutación en la p53. La especificidad es importante cuando el costo o riesgo que se asocia con más procedimientos de diagnóstico o más intervenciones médicas es muy alto.
- WO 2006/008128 describe métodos y productos para la detección/clasificación de cáncer de mama que comprende determinar el nivel de metilación de FOXL2. Se describen además las secuencias de nucleótidos respectivas.
- US B1 6,783,933 describe un método para detectar/clasificar un trastorno proliferativo celular colorrectal mediante la determinación del estado de metilación de las islas CpG en las regiones reguladoras del gen de CACNA1G en el que la metilación de CpG es indicativo de la presencia de dicho trastorno.
- WO 01/29254 describe FOXL2 como un marcador para el síndrome de blefarofimosis.
- Necesidad pronunciada en la técnica.* Se acepta generalmente que existe una necesidad pronunciada en la técnica para mejorar el tamizaje y la detección temprana de los cánceres. Como un ejemplo, si la especificidad de detección de cáncer de colon se puede aumentar, el problema de los falsos resultados positivos de la prueba que conducen a la exploración colonoscópica innecesaria se reduciría lo que conduciría al ahorro de costes y a la mejora de la seguridad.
- En vista de la incidencia de los cánceres en general y más particularmente a las desventajas que se asocian con los métodos actuales de detección de trastornos proliferativos celulares hepatocelulares y colorrectales existe una necesidad sustancial en la técnica de mejorar los métodos para la detección temprana del cáncer, en particular de cáncer de colon, para utilizarlos adicionalmente a o como sustituto de las pruebas disponibles en la actualidad.
- El gen humano Septina 9 (que también se conoce como proteína de fusión similar a MLL-septina, proteína de fusión MSF-A similar a MLL-septina, S1pa, Eseptina, Msf, proteína similar a septina septina de Ovario/Mama (Ov/Br septina) y Septina D1) se localiza en el cromosoma 17q25 dentro del cóntigo AC068594.15.1.168501 y es miembro de la familia de genes Septina.
- Se postula que los miembros de la familia del gen Septina se asocian con múltiples funciones celulares que van desde el transporte de vesículas a la citocinesis. La interrupción de la acción de Septina 9 resulta en la división celular incompleta, ver Surka, M.C., Tsang, C.W., y Trimble, W.S. *Mol Biol Cell*, 13: 3532-45 (2002). La *Septina 9* y otras proteínas mostraron ser parejas de fusión del proto-oncogeno MLL lo que sugiere una función en la tumorigenesis, ver Osaka, M, Rowley, J.D. y Zeleznik-Le, N.J. *PNAS*, 96:6428-6433 (1999). Burrows y otros informaron de un estudio a profundidad de la expresión de las múltiples isoformas del gen Septina 9 en el cáncer de ovario y mostraron la expresión específica de tejido de diversos transcritos, ver Burrows, J. F., Chanduloy, y otros S.E.H. *Journal of Pathology*, 201:581-588 (2003).
- Un estudio reciente (técnica anterior que se publicó en fecha posterior a la prioridad) de más de 7.000 tejidos normales y tumorales indica que hay una sobre-expresión consistente de las isoformas de Septina 9 en un número de tejidos tumorales, ver Scott, M., Hyland, P.L., y otros *Oncogene*, 24: 4688-4700 (2005). Los autores especulan que el gen parece ser un gen del cáncer tipo II donde los cambios en el control del procesamiento del transcrito de ARN de diferentes productos proteicos, y los niveles de estas isoformas proteicas alteradas pueden dar respuestas a la función del gen en tumores malignos.
- La proteína MSF (factor estimulante de la migración) que se transcribe a partir del gen FN1 también se implica en la

carcinogénesis (ver WO99/31233), sin embargo hay que señalar que esta proteína no es objeto de la presente aplicación, y en la actualidad no se conoce que se asocie con el gen Septina 9/ MSF ni con los productos transcritos de los mismos.

5 A partir de las referencias que se citaron anteriormente se puede observar que los mecanismos biológicos que vinculan dicho gen a la tumorigénesis siguen sin estar claros. En la WO 200407441 se reivindica que el aumento del número de copias y la sobre-expresión del gen es un marcador de cáncer, y además proporciona medios para el diagnóstico y tratamiento del mismo de acuerdo con dicha observación. La WO 200407441 en consecuencia es la técnica anterior más próxima ya que tiene el mayor número de características en común con el método y los ácidos nucleicos de la presente invención, y porque se relaciona con el mismo campo (diagnóstico de cáncer). La principal diferencia entre la presente invención y la de la WO 200407441 es que la presente invención muestra por primera vez que la baja expresión del gen Septina 9 se asocia con el cáncer. Más particularmente esto se ilustra por medio del análisis de metilación. La correlación entre la expresión y la metilación del ADN, y los métodos para la determinación de la metilación del ADN se conocen en la técnica (ver WO 99/28498). Sin embargo, no sería obvio para aquellos con experiencia en la técnica que la baja expresión se asocie también con el desarrollo de cáncer, en particular debido a que la WO 200407441 describe la modulación de dicha expresión a niveles más bajos como una terapia potencial para el cáncer.

20 La SEQ ID NO: 24 proporciona una secuencia rica en CpG que se localiza en el cromosoma 3q23, corriente abajo del gen Factor de Transcripción L2 Forkhead (FOXL2, Factor Forkhead Pituitario, OMIM 605597). La SEQ ID NO: 24 hasta ahora no se asocia con ningún tipo de cáncer. La región codificante FOXL2 es muy conservada entre los mamíferos. La evidencia inmunohistoquímica indica que FOXL2 es una proteína nuclear que se expresa específicamente en párpados y en células foliculares ováricas fetales y de adultos. Se sugiere que FOXL2 puede desempeñar una función en la diferenciación de las células somáticas ováricas y en el desarrollo y/o mantenimiento adicional del folículo (J. Med. Genet. 39: 916-922, 2002.).

25 Más aún, las mutaciones en el gen FOXL2 se asocian con el síndrome blefarofimosis/ ptosis/epicanto inverso (BPES), un síndrome que afecta a los párpados y al ovario (Am. J. Hum. Genet. 72: 478-487, 2003.; Hum. Mutat. 24: 189-193, 2004). El gen FOXL2 hasta ahora no se asocia con el cáncer, sin embargo, otros miembros de la familia FOX se implican en la oncogénesis. El gen FOXA1 se amplifica y se sobreexpresa en el cáncer de esófago y de pulmón, el gen FOXM1 se regula positivamente en el cáncer de páncreas y en el carcinoma de células basales debido a la regulación transcripcional por la vía Sonic Hedgehog (SHH).

30 Breve descripción de la invención

35 La presente invención proporciona un método para la detección y/o clasificación del carcinoma colorrectal en un sujeto lo que comprende determinar la presencia o ausencia de metilación CpG de FOXL2 en una muestra biológica aislada de dicho sujeto en donde la hipermetilación CpG de FOXL2 es indicativa de la presencia o clase de dicho carcinoma colorrectal. En el contexto del carcinoma colorrectal los métodos de ensayo de la invención tienen una utilidad particular para el tamizaje de las poblaciones en riesgo. Los métodos de la invención tienen ventajas sobre los métodos de la técnica anterior (lo que incluye el estándar de la industria FOBT), a causa de la mejora de sensibilidad, especificidad y probablemente de conformidad del paciente.

40 Preferentemente, la sensibilidad de dicha detección es de aproximadamente 75% a aproximadamente 96%, o de aproximadamente 80% a aproximadamente 90%, o de aproximadamente 80% a aproximadamente 85%.

45 El método es novedoso ya que actualmente no existen métodos que permitan la detección del cáncer colorrectal mediante el análisis de fluidos corporales, con una sensibilidad y especificidad suficientemente altas para su uso en un ensayo comercialmente disponible y aprobado por una entidad regulatoria. Por ejemplo, los métodos actuales que se utilizan para detectar y diagnosticar el carcinoma colorrectal incluyen colonoscopia, sigmoidoscopia, y cáncer de colon de sangre oculta en heces. En comparación con estos métodos, la invención que se describe es mucho menos invasiva que la colonoscopia, y tan, sino más sensible que la sigmoidoscopia y la FOBT. El desarrollo de un ensayo de fluido corporal representa una clara ventaja técnica sobre los métodos actuales que se conocen en la técnica en que se prevé que por lo menos para la detección del carcinoma colorrectal la conformidad del paciente para una única prueba basada en fluido corporal será más alta que para el análisis por triplicado de heces que actualmente se recomienda para la FOBT.

50 En una modalidad particular el método comprende el uso del gen o secuencia genómica de FOXL2 como un marcador para la detección del cáncer colorrectal.

55 La detección, diferenciación y distinción del cáncer colorrectal se habilita por medio del análisis del estado de metilación del gen o secuencia genómica de FOXL2, y de sus promotores o elementos reguladores.

60 El método tiene utilidad para mejorar el diagnóstico, tratamiento y seguimiento del cáncer colorrectal.

5 Preferentemente, la fuente de la muestra de prueba se selecciona del grupo que consiste en células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, heces, eyaculado, orina, sangre, y combinaciones de los mismos. Con mayor preferencia, la fuente se selecciona del grupo que consiste en heces, plasma sanguíneo, suero sanguíneo, sangre completa, células de sangre aisladas, células aisladas de la sangre que se obtuvieron del sujeto.

10 Específicamente, la presente invención proporciona un método para detectar el cáncer colorrectal y para diferenciar entre trastornos proliferativos celulares de cáncer colorrectal neoplásico y benigno, que comprende: obtener una muestra biológica que comprende ácido(s) nucleico(s) genómico(s); poner en contacto el(los) ácido(s) nucleico(s), o un fragmento de estos, con un reactivo o una pluralidad de reactivos suficiente para distinguir entre secuencias de dinucleótidos CpG metilados y no metilados dentro de al menos una secuencia objetivo del ácido nucleico del sujeto, en donde la secuencia objetivo comprende, o hibrida bajo condiciones rigurosas a, una secuencia que comprende al menos 16 nucleótidos contiguos de la secuencia SEQ ID NO:24, dichos nucleótidos contiguos comprenden al menos una secuencias de
15 dinucleótidos CpG; y determinar, basado al menos en parte en dicha distinción, el estado de metilación de al menos una secuencias de dinucleótidos CpG objetivo, o un promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpGs objetivo.

20 Preferentemente, distinguir entre secuencias de dinucleótidos CpG metilados y no metilados dentro de la secuencia objetivo comprende la conversión o no conversión dependiente del estado de metilación de al menos una de dichas secuencias de dinucleótidos CpG a la secuencia de dinucleótido convertida o no convertida correspondiente dentro de la secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43, y complementos de estas.

25 Las modalidades adicionales proporcionan un método para la detección de (o distinguirla del trastorno proliferativo celular benigno), cáncer colorrectal, que comprende: obtener una muestra biológica que tiene el ADN genómico del sujeto; extraer el ADN genómico; tratar el ADN genómico, o un fragmento del mismo, con uno o más reactivos para convertir las bases citosina no metiladas en la posición 5 a uracilo u otra base que pueda detectarse como diferente de la citosina en términos de las propiedades de hibridación; poner en contacto el ADN genómico tratado, o el fragmento tratado del mismo, con una
30 enzima de amplificación y al menos dos iniciadores que comprenden, en cada caso una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43, y complementos de estas, en donde en donde el ADN o el fragmento tratado del mismo se amplifica para producir un amplificado, o no se amplifica; y determinar, basado en la presencia o ausencia de, o en una propiedad de dicho
35 amplificado, el estado de metilación o un promedio, o un valor que refleja un promedio del nivel de metilación de al menos uno, pero con mayor preferencia una pluralidad de dinucleótidos CpG de una secuencia que se selecciona del grupo que consiste en SEQ ID NO:24.

40 Preferentemente, la determinación comprende usar al menos un método que se selecciona del grupo que consiste en: i) hibridar al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43, y complementos de estas; ii) hibridar al menos una molécula de ácido nucleico, unida a una fase sólida, que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente
45 rigurosas a una secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43, y complementos de estas; iii) hibridar al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO:42, SEQ ID NO:43, y complementos de estas, y extender al menos una de dicha molécula de ácido nucleico
50 hibridada en al menos un nucleótido base; y iv) secuenciar el amplificado.

Otras modalidades proporcionan un método para el análisis (es decir, detección y/o clasificación) de trastornos proliferativos celulares, que comprende: obtener una muestra biológica que tiene el ADN genómico del sujeto; extraer el ADN genómico; poner en contacto el ADN genómico, o un fragmento del mismo, que comprende una o más secuencias que se seleccionan del grupo que consiste en SEQ ID NO:24 o una secuencia que hibrida bajo condiciones rigurosas a esta, con una o más
55 enzimas de restricción sensibles a la metilación, en donde el ADN genómico se digiere de esta manera para producir fragmentos de digestión, o no se digiere de esta manera; y determinar, basado en la presencia o ausencia de, o en una propiedad de al menos uno de dichos fragmentos, el estado de metilación de al menos una secuencia de dinucleótidos CpG de al menos una secuencia genómica que se selecciona del grupo que consiste en SEQ ID NO:24 o un promedio, o un valor

que refleja un estado de metilación promedio de una pluralidad de secuencias de secuencias CpG de dinucleótidos de las mismas. Preferentemente, el ADN genómico digerido o sin digerir se amplifica antes de dicha determinación.

Breve descripción de las figuras

5

La Figura 1 proporciona tres gráficos. Los dos gráficos de la izquierda muestran la sensibilidad del ensayo de acuerdo con la tabla 7 en muestras de carcinoma colorrectal y de sangre en el Ejemplo 2. El gráfico superior muestra una distribución binaria de dos tipos de muestras, mientras que el gráfico inferior proporciona una distribución multiclase. El Eje Y de dichos gráficos muestra la proporción de muestras analizadas con un nivel de metilación mayor que el valor cuantificado que se muestra en el eje X. El gráfico de la derecha proporciona una ROC de la detección del carcinoma colorrectal.

10

Descripción detallada de la invención

Definiciones:

15

El término "relación observado/esperado" ("relación O/E") se refiere a la frecuencia de dinucleótidos CpG dentro de una secuencia de ADN particular, y corresponde al [número de sitios de CpG / (número de bases C x número de bases G)] / longitud de banda para cada fragmento.

20

El término "isla CpG" se refiere a una región contigua de ADN genómico que satisface los criterios de (1) tener una frecuencia de dinucleótidos CpG que corresponde a una "Relación Observado/Esperado >0.6, y (2) tener un "Contenido de GC" >0.5. Las islas CpG son típicamente, pero no siempre, de entre aproximadamente 0.2 a aproximadamente 1 KB, o hasta aproximadamente 2 kb de longitud.

25

El término "estado de metilación" o "condición de metilación" se refiere a la presencia o ausencia de 5-metilcitosina ("5-mCyt") en uno o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN. Los estados de metilación en uno o más sitios de metilación CpG particulares (cada uno con dos secuencias de dinucleótidos CpG) dentro de una secuencia de ADN incluyen "no metilado", "totalmente metilado" y "hemi-metilado".

30

El término "hemi-metilación" o "hemimetilación" se refiere al estado de metilación de un ADN bicatenario en donde sólo una hebra del mismo es metilada.

35

El término 'AUC' como se utiliza en la presente descripción es una abreviatura para el área bajo una curva. En particular se refiere al área bajo una Curva Característica Operativa del Receptor (ROC). La curva ROC es un gráfico de la tasa de verdaderos positivos contra la tasa de falsos positivos para los diferentes puntos de corte posibles de una prueba de diagnóstico. Muestra el compromiso entre sensibilidad y especificidad dependiendo del punto de corte que se seleccione (cualquier aumento en la sensibilidad se acompañará por una disminución en la especificidad). El área bajo una curva ROC (AUC) es una medida de la exactitud de una prueba de diagnóstico (cuanto mayor sea el área mejor, óptimo es 1, una prueba aleatoria puede tener una curva ROC que yace en la diagonal con un área de 0.5; para la referencia: J.P. Egan. Signal Detection Theory and ROC Analysis, Academic Press, Nueva York, 1975).

40

El término "hipermetilación" se refiere al estado de metilación promedio que corresponde a un *aumento* de la presencia de 5-mCyt en un o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN de una muestra de ADN prueba, con relación a la cantidad de 5-mCyt que se encuentra en los dinucleótidos CpG correspondientes dentro de una muestra normal de ADN control.

45

El término "hipometilación" se refiere al estado de metilación promedio que corresponde a una *disminución* de la presencia de 5-mCyt en uno o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN de una muestra de ADN prueba, con relación a la cantidad de 5-mCyt que se encuentra en los dinucleótidos CpG correspondientes dentro de una muestra normal de ADN control.

50

El término "microarreglos" se refiere ampliamente tanto a "microarreglos de ADN," como a "matrices de ADN," como se reconoce en la técnica, abarca todos los soportes sólidos que se reconocen en la técnica, y abarca todos los métodos para fijar moléculas de ácido nucleico a estos o de síntesis de ácidos nucleicos en estos.

55

Los "parámetros genéticos" son mutaciones y polimorfismos de genes y secuencias que se requieren adicionalmente para su regulación. Se designan como mutaciones, en particular, las inserciones, deleciones, mutaciones puntuales, inversiones y polimorfismos y se prefieren particularmente los SNPs (polimorfismos de un solo nucleótido).

60

Los "Parámetros epigenéticos" son, en particular, metilaciones de citosina. Los parámetros epigenéticos adicionales

incluyen, por ejemplo, la acetilación de histonas que, sin embargo, no puede analizarse directamente mediante la utilización del método que se describe pero que, a su vez, se correlaciona con la metilación del ADN.

5 El término "reactivo de bisulfito" se refiere a un reactivo que comprende bisulfito, disulfito, sulfito de hidrógeno o combinaciones de estos, útiles como se describe en la presente descripción para distinguir entre secuencias de dinucleótidos CpGs metiladas y no metiladas.

10 El término "Ensayo de metilación" se refiere a cualquier ensayo para determinar el estado de metilación de una o más secuencias de dinucleótidos CpGs dentro de una secuencia de ADN.

15 El término "MS.AP-PCR" (Reacción en Cadena de la Polimerasa Arbitrariamente Iniciada Sensible a la Metilación) se refiere a la tecnología que se reconoce en la técnica que permite un tamizaje global del genoma mediante la utilización de iniciadores ricos en CG para enfocarse en las regiones con mayor probabilidad de contener dinucleótidos CpG, y que se describen en Gonzalzo y otros, Cancer Research 57:594-599, 1997.

El término "MethyLight™" se refiere a la técnica de PCR que se basa en fluorescencia en tiempo real que se reconoce en la técnica que se describe por Eads y otros, Cancer Res. 59:2302-2306, 1999.

20 El término ensayo "HeavyMethyl™", en la modalidad de este que se implementa en la presente descripción, se refiere a un ensayo, en donde las sondas *que bloquean* la metilación específica (que se refieren también en la presente descripción como *bloqueadores*) revisten las posiciones CpG entre, o revestidas por los iniciadores de la amplificación lo que permite la amplificación selectiva específica de la metilación de una muestra de ácido nucleico.

25 El término ensayo "HeavyMethyl™ MethyLight™", en la modalidad de este que se implementa en la presente descripción, se refiere a un ensayo HeavyMethyl™ MethyLight™, que es una variación del ensayo MethyLight™, en donde el ensayo MethyLight™ se combina con las sondas *que bloquean* la metilación específica que reviste las posiciones CpG entre los iniciadores de la amplificación.

30 El término "Ms-SNuPE" (Extensión del Iniciador de un Solo Nucleótido Sensible a Metilación) se refiere al ensayo que se reconoce en la técnica que se describe por Gonzalzo & Jones, Nucleic Acids Res. 25:2529-2531, 1997.

35 El término "MSP" (PCR específica de Metilación) se refiere al ensayo de metilación que se reconoce en la técnica que se describe por Herman y otros Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996, y por la patente de Estados Unidos núm. 5,786,146.

El término "COBRA" (Análisis de Restricción de Bisulfito Combinado) se refiere al ensayo de metilación que se reconoce en la técnica que se describe por Xiong & Laird, Nucleic Acids Res. 25:2532-2534, 1997.

40 El término "MCA" (amplificación de la Isla CpG metilada) se refiere al ensayo de metilación que se reconoce en la técnica que se describe por Toyota y otros, Cancer Res. 59:2307-12, 1999, y en WO 00/26401A1.

El término "hibridación" debe entenderse como un enlace de un oligonucleótido con una secuencia complementaria a lo largo de las líneas de apareamientos de bases de Watson-Crick en el ADN muestra, formando una estructura bicatenaria.

45 "Condiciones rigurosas de hibridación," tal como se definen en la presente descripción, implican hibridar a 68 °C en solución de SSC 5x/Denhardt's 5x/SDS 1.0%, y lavar en SSC 0.2x /SDS 0.1% a temperatura ambiente, o implican el equivalente de estas reconocidas en la técnica (por ejemplo, condiciones en las que una hibridación se lleva a cabo a 60°C en amortiguador SSC 2.5 x, seguido por varias etapas de lavado a 37 °C en una concentración baja de amortiguador, y permanece estable). Condiciones moderadamente rigurosas, como se definen en la presente descripción, implican incluir el lavado en SSC 3X a 42 °C, o lo equivalente de esto que se reconoce en la técnica. Los parámetros de concentración de sal y la temperatura se pueden variar para alcanzar el nivel óptimo de identidad entre la sonda y el ácido nucleico objetivo. La orientación sobre dichas condiciones está disponible en la técnica, por ejemplo, en Sambrook y otros, 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Nueva York.; y Ausubel y otros (eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, Nueva York) en la unidad 2.10.

55 Los términos "Enzimas de restricción metilación-específica" o "enzimas de restricción metilación-sensibles" se deben tomar en el sentido de una enzima que digiere selectivamente un ácido nucleico en dependencia del estado de metilación de su sitio de reconocimiento. En el caso de tales enzimas de restricción que cortan específicamente si el sitio de reconocimiento no está metilado o hemimetilado, el corte no se llevará a cabo, o con una eficiencia significativamente reducida, si el sitio de reconocimiento está metilado. En el caso de tales enzimas de restricción que cortan específicamente si el sitio de

60

reconocimiento está metilado, el corte no se llevará a cabo, o con una eficiencia significativamente reducida si el sitio de reconocimiento no está metilado. Se prefieren las enzimas de restricción específicas de metilación, la secuencia de reconocimiento de las cuales contiene un dinucleótido CG (por ejemplo cgcg o cccggg). Se prefieren más aun para algunas modalidades enzimas de restricción que no cortan si la citosina en este dinucleótido se metila en el átomo de carbono C5.

5

Las "Enzimas de restricción específicas de la no-metilación" o "enzimas de restricción sensibles a la no-metilación" son enzimas de restricción que cortan una secuencia de ácidos nucleicos independientemente del estado de metilación con una eficiencia casi idéntica. También se les llama "enzimas de restricción metilación-inespecíficas".

10

El término "gen" debe entenderse como que incluye a todas las variantes transcritas de los mismos (por ejemplo, el término "Septina 9" incluirá por ejemplo su transcrito truncado Q9HC74) y todo promotor y elementos reguladores del mismo. Además, como se conoce una pluralidad de SNPs dentro de dicho gen el término debe entenderse como que incluye todas las variantes de secuencia del mismo.

15

El término "pre-canceroso" o "pre-neoplásico" y equivalentes de los mismos deben entenderse que incluyen cualquier trastorno de la proliferación celular que experimente una transformación maligna. Ejemplos de tales condiciones incluyen, en el contexto de los trastornos proliferativos celulares colorrectales, trastornos proliferativos celulares con un alto grado de displasia y las siguientes clases de adenomas:

20

Nivel 1: penetración de las glándulas malignas a través de la mucosa muscular en la submucosa, dentro de la cabeza del pólipo;

Nivel 2: la misma invasión submucosa, pero presente en la unión de la cabeza al tallo;

Nivel 3: invasión del tallo; y

Nivel 4: invasión de la base del tallo en la conexión a la pared del colon (este nivel corresponde a la etapa Dukes A).

25

Descripción general:

30

La presente invención proporciona un método para detectar y/o clasificar el cáncer colorrectal en un sujeto que comprende determinar la metilación CpG del gen o secuencia genómica de FOXL2 en una muestra biológica que se aísla de dicho sujeto en donde la metilación CpG es indicativa de la presencia o clase de dicho trastorno. Dicho marcador puede utilizarse para el diagnóstico del cáncer colorrectal, lo que incluye la detección temprana durante las etapas pre-cancerosas de la enfermedad, y además para la diferenciación del cáncer colorrectal benigno del neoplásico.

35

El marcador de la presente invención es particularmente eficiente en la detección o distinción entre el cáncer celular colorrectal, proporcionando de esta manera medios mejorados para la detección temprana, la clasificación y el tratamiento de dicho trastorno.

40

En una primera modalidad adicional la presente invención se basa en el análisis del estado de metilación CpG del gen o secuencia genómica de FOXL2. Se prefiere además que la secuencia de dicho gen esté de acuerdo con la TABLA 1.

45

La modificación del ADN con bisulfito es una herramienta que se reconoce en la técnica que se utiliza para evaluar la condición de metilación CpG. La 5-metilcitosina es la modificación de base covalente más frecuente en el ADN de las células eucariotas. Esta desempeña una función, por ejemplo, en la regulación de la transcripción, en la impronta genética, y en la tumorigénesis. Por lo tanto, la identificación de la 5-metilcitosina como un componente de la información genética es de gran interés. Sin embargo, las posiciones 5-metilcitosina pueden no identificarse por secuenciación, porque la 5-metilcitosina tiene el mismo comportamiento de apareamiento de bases que la citosina. Además, la información epigenética que se porta por la 5-metilcitosina se pierde por completo durante, por ejemplo, la amplificación por PCR.

50

El método que se utiliza más frecuentemente para analizar ADN para la presencia de 5-metilcitosina se basa en la reacción específica de bisulfito con citosina por lo que tras la subsiguiente hidrólisis alcalina, la citosina se convierte en uracilo, que se corresponde con timina en su comportamiento de apareamiento de bases. Significativamente, sin embargo, la 5-metilcitosina permanece sin modificarse bajo estas condiciones.

55

Por consiguiente, el ADN original se convierte de tal manera que la metilcitosina, que originalmente no podía distinguirse de la citosina por su comportamiento de hibridación, ahora puede detectarse ya que sólo la citosina permanece mediante la utilización de técnicas de biología molecular estándares, que se reconocen en la técnica, por ejemplo, por amplificación e hibridación, o por secuenciación. Todas estas técnicas se basan en las propiedades diferenciales de apareamiento de bases, que pueden ahora explotarse plenamente.

60

La técnica anterior, en términos de sensibilidad, se define por un método que comprende encerrar el ADN que se analiza en

una matriz de agarosa, lo que evita así la difusión y renaturalización del ADN (el bisulfito sólo reacciona con el ADN monocatenario), y sustituir todas las etapas de precipitación y purificación con diálisis rápida (Olek A, y otros, A modified and improved method for bisulfite based cytosine methylation analysis, *Nucleic Acids Res.* 24:5064-6, 1996). Así, es posible analizar las células individuales para la condición de metilación, lo que ilustra la utilidad y sensibilidad del método. Una visión general de los métodos que se reconocen en la técnica para la detección de 5-metilcitosina se proporciona por Rein, T., y otros, *Nucleic Acids Res.*, 26:2255, 1998.

La técnica de bisulfito, salvo pocas excepciones (por ejemplo., Zeschnick M, y otros, *Eur J Hum Genet.* 5:94-98, 1997), actualmente sólo se utiliza en la investigación. En todos los casos, los fragmentos cortos, específicos de un gen que se conoce se amplifican después de un tratamiento con bisulfito, y, o bien completamente secuenciados (Olek & Walter, *Nat Genet.* 1997 17:275-6, 1997), sometidos a una o más reacciones de extensión del iniciador (Gonzalzo & Jones, *Nucleic Acids Res.*, 25:2529-31, 1997; WO 95/00669; patente de los Estados Unidos núm. 6,251,594) para analizar las posiciones de citosina individuales, o tratado por digestión enzimática (Xiong & Laird, *Nucleic Acids Res.*, 25:2532-4, 1997). La detección por hibridación también se describió en la técnica (Olek y otros, WO 99/28498). Además, se describió el uso de la técnica de bisulfito para la detección de metilación con respecto a los genes individuales (Grigg & Clark, *Bioessays*, 16:431-6, 1994; Zeschnick M, y otros, *Hum Mol Genet.*, 6:387-95, 1997; Feil R, y otros, *Nucleic Acids Res.*, 22:695-, 1994; Martin V, y otros, *Gene*, 157:261-4, 1995; WO 97/46705 y WO 95/15373).

La presente invención proporciona el uso de la técnica de bisulfito, en combinación con uno o más ensayos de metilación, para la determinación del estado de metilación de las secuencias de dinucleótidos CpG dentro de la secuencia que se selecciona de la SEQ ID NO:24. Los dinucleótidos CpG genómicos pueden ser metilados o no metilados (lo que se conoce alternativamente como metilado ascendente y descendentemente respectivamente). Sin embargo los métodos de la presente invención son adecuados para el análisis de las muestra biológicas de naturaleza heterogénea, por ejemplo, una baja concentración de células tumorales dentro de un fondo de sangre o heces. En consecuencia, al analizar el estado de metilación de una posición de CpG dentro de una muestra tal la persona experta en la técnica puede utilizar un ensayo cuantitativo para determinar el nivel (por ejemplo, el por ciento, fracción, relación, proporción o grado) de metilación en una posición particular CpG en oposición a un estado de metilación. En consecuencia el término estatus de metilación o estado de metilación debe entenderse además en el sentido de un valor que refleja el grado de metilación en una posición CpG. A menos que se indique específicamente los términos "hipermetilado" o "no metilado" debe entenderse que significan un nivel de metilación por encima del de un punto de corte especificado, en donde dicha línea de corte puede ser un valor que representa la media o la mediana del nivel de metilación de una población dada, o es preferentemente un nivel de corte optimizado. El punto "de corte" también se denomina en la presente descripción como "umbral". En el contexto de la presente invención los términos "metilado", "hipermetilado" o "metilado ascendente" debe entenderse que incluyen un nivel de metilación por encima del punto de corte que sea cero (0) % (o sus equivalentes) metilación para todas las posiciones CpG dentro y asociada con (por ejemplo, en regiones reguladoras o promotoras) el gen o secuencia genómica de FOXL2.

De acuerdo con la presente invención, la determinación del estado de metilación de las secuencias de dinucleótidos CpG dentro de la SEQ ID NO: 24 tiene utilidad tanto en el diagnóstico como en la caracterización del cáncer colorrectal. Procedimientos de ensayo de metilación. Varios procedimientos de ensayos de metilación se conocen en la técnica, y pueden usarse junto con la presente invención. Estos ensayos permiten la determinación del estado de metilación de uno o una pluralidad de dinucleótidos CpG (por ejemplo, islas CpG) dentro de una secuencia de ADN. Tales ensayos implican, entre otras técnicas, la secuenciación de ADN del ADN tratado con bisulfito, PCR (para la amplificación específica de secuencia), análisis de transferencia Southern, y el uso de enzimas de restricción sensibles a la metilación.

Por ejemplo, la secuenciación genómica se simplifica para el análisis de los patrones de metilación del ADN y la distribución de 5-metilcitosina mediante la utilización del tratamiento con bisulfito (Frommer y otros, *Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). Además, se usa la digestión con enzimas de restricción de los productos de PCR amplificados a partir del ADN convertido con bisulfito, por ejemplo, el método descrito por Sadri & Hornsby (*Nucl. Acids Res.* 24:5058-5059, 1996), o COBRA (análisis de restricción de bisulfito combinado) (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997).

COBRA. El análisis *COBRA*[™] es un ensayo de metilación cuantitativo útil para determinar los niveles de metilación de ADN en los loci de gen específico en cantidades pequeñas de ADN genómico (Xiong & Laird, *Nucleic Acids Res.* 25:2532-2534, 1997). Brevemente, la digestión con enzimas de restricción se utiliza para revelar las diferencias en la secuencia dependiente de la metilación en los productos de PCR de ADN tratado con bisulfito sódico.

Las diferencias en la secuencia dependiente de la metilación se introducen primero en el ADN genómico por tratamiento con bisulfito estándar de acuerdo con el procedimiento descrito por Frommer y otros (*Proc. Natl. Acad. Sci. USA* 89:1827-1831, 1992). La amplificación por PCR del ADN convertido con bisulfito se realiza después mediante la utilización de iniciadores específicos para las islas CpG de interés, seguido de la digestión con endonucleasas de restricción, electroforesis en gel, y detección con la utilización sondas de hibridación marcadas y específicas. Los niveles de metilación en la muestra de ADN

original se representan por las cantidades relativas de producto de PCR digerido y no digerido de una manera linealmente cuantitativa a través de un amplio espectro de niveles de metilación de ADN. Además, esta técnica se puede aplicar fácilmente al ADN que se obtiene a partir de las muestras de tejidos microdisecadas embebidas en parafina.

5 Los reactivos típicos (*por ejemplo*, como pueden encontrarse en un estuche típico a base de COBRAT™) para el análisis COBRA™ pueden incluir, pero sin limitarse a: iniciadores de PCR para el gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); enzima de restricción y amortiguador adecuado; y oligonucleótido del gen de hibridación; oligonucleótido de control de hibridación; estuche de marcaje de quinasa para sonda de oligonucleótido; y nucleótidos marcados. Además, los reactivos de conversión con bisulfito pueden incluir: amortiguador de desnaturalización de ADN; amortiguador sulfonación, reactivos o estuches de recuperación de ADN (*por ejemplo*, precipitación, ultrafiltración, columna de afinidad); amortiguador de desulfonación; y componentes de recuperación de ADN.

15 Preferentemente, los ensayos tales como "MethylLight™" (una técnica de PCR en tiempo real basada en fluorescencia) (Eads y otros, Cancer Res. 59:2302-2306, 1999), reacciones de Ms-SNuPE (extensión del iniciador de un solo nucleótido sensible a metilación) (Gonzalzo & Jones, Nucleic Acids Res. 25:2529-2531, 1997), PCR específica de metilación ("MSP"; Herman y otros, Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; patente de Estados Unidos núm. 5,786,146), y amplificación de la Isla CpG metilada ("MCA"; Toyota y otros, Cancer Res. 59:2307-12, 1999) se usan solos o en combinación con otros de estos métodos.

20 La técnica del ensayo "HeavyMethyl™", es un método cuantitativo para evaluar las diferencias de metilación basadas en la amplificación específica de metilación del ADN tratado con bisulfito. Las sondas *que bloquean* la metilación específica (que se refieren también en la presente descripción como *bloqueadores*) revisten las posiciones CpG entre, o revestidas por los iniciadores de la amplificación permiten la amplificación selectiva específica de la metilación de una muestra de ácido nucleico.

25 El término ensayo "HeavyMethyl™ MethylLight™", en la modalidad de este que se implementa en la presente descripción, se refiere a un ensayo HeavyMethyl™ MethylLight™, que es una variación del ensayo MethylLight™, en donde el ensayo MethylLight™ se combina con las sondas *que bloquean* la metilación específica que reviste las posiciones CpG entre los iniciadores de la amplificación.

30 Los reactivos típicos (*por ejemplo*, tal como se pueden encontrar en un estuche típico basado en MethylLight™) para el análisis HeavyMethyl™ pueden incluir, pero sin limitarse a: iniciadores de PCR para genes específicos (o secuencia de ADN tratada con bisulfito o isla CpG); oligonucleótidos bloqueadores; amortiguadores y desoxinucleótidos de PCR optimizados; y Taq polimerasa.

35 *MSP*. La MSP (PCR específica de la metilación) permite evaluar virtualmente la condición de metilación de cualquier grupo de sitios CpG dentro de una isla CpG, independiente del uso de enzimas de restricción sensibles a la metilación (Herman y otros Proc. Natl. Acad. Sci. USA 93:9821-9826, 1996; patente de Estados Unidos núm. 5,786,146). En resumen, el ADN se modifica por bisulfito sódico convirtiendo todas las citosinas no metiladas, pero no las metiladas en uracilo, y posteriormente se amplifica con iniciadores específicos para ADN metilado frente al no metilado. La MSP requiere solamente pequeñas cantidades de ADN, es sensible a 0.1% de alelos metilados de un locus determinado de la isla CpG, y se puede realizar en el ADN que se extrae de muestras incluidas en parafina. Los reactivos típicos (*por ejemplo*, como se pueden encontrar en un típico estuche que se basa en MSP) para el análisis de MSP pueden incluir, pero sin limitarse a: iniciadores de PCR metilados y no metilados para el gen específico (o secuencia de ADN tratada con bisulfito o isla CpG), amortiguadores y desoxinucleótidos de PCR optimizados, y sondas específicas.

40 *MethylLight™*. El ensayo MethylLight™ es un ensayo de metilación cuantitativa de alto rendimiento que utiliza tecnología de PCR en tiempo real basada en fluorescencia (TaqMan®) que no requiere de manipulaciones adicionales después de la etapa de PCR (Eads y otros, Cancer Res. 59:2302-2306, 1999). Brevemente, el proceso MethylLight™ comienza con una muestra mixta de ADN genómico que se convierte, con una reacción de bisulfito sódico, en una mezcla mixta de diferencias en la secuencia dependientes de la metilación de acuerdo con procedimientos estándar (el proceso de bisulfito convierte los residuos de citosina no metiladas a uracilo). La PCR basada en fluorescencia se lleva a cabo después en una reacción "sesgada" (con iniciadores de PCR que solapan los dinucleótidos CpG). La discriminación de secuencia puede ocurrir a nivel del proceso de amplificación y a nivel del proceso de la detección de fluorescencia.

50 El ensayo MethylLight™ puede utilizarse como una prueba cuantitativa de los patrones de metilación en la muestra de ADN genómico, en donde la discriminación de secuencia ocurre a nivel de la hibridación de la sonda. En esta versión cuantitativa, la reacción de PCR se proporciona para la amplificación específica de la metilación en presencia de una sonda fluorescente que solapa un sitio potencial de metilación en particular. Se proporciona un control no-sesgado para la cantidad de ADN de entrada por una reacción en la que ni los iniciadores, ni la sonda revisten ninguno de los dinucleótidos CpG.

Alternativamente, una prueba cualitativa para la metilación genómica se logra mediante el sondeo de la mezcla de PCR sesgada, ya sea con oligonucleótidos control que no "revisten" sitios de metilación que se conozcan (una versión basada en fluorescencia de las técnicas HeavyMethyl™ y MSP), o con oligonucleótidos oligonucleótidos que revisten los posibles sitios de metilación.

5

El proceso MethyLight™ puede utilizarse con cualquier sonda adecuada por ejemplo "TaqMan®", Lightcycler® etc.... Por ejemplo, el ADN genómico bicatenario se trata con bisulfito de sodio y se somete a uno de los dos conjuntos de reacciones de PCR utilizando sondas TaqMan®; por ejemplo, con los iniciadores de MSP y/o oligonucleótidos bloqueadores HeavyMethyl y la sonda TaqMan®. La sonda TaqMan® se marca doble con las moléculas fluorescentes "reportero" e "inactivador", y se diseña para ser específica de una región GC de contenido relativamente alto tal que se funde a aproximadamente a 10 °C de temperatura más alta en el ciclo de PCR que los iniciadores directo o inverso. Esto permite que la sonda TaqMan® permanezca completamente hibridada durante la etapa de apareamiento/extensión del PCR. Debido a que la Taq polimerasa sintetiza enzimáticamente una nueva cadena durante el PCR, eventualmente se logrará la sonda TaqMan® apareada. La actividad endonucleasa 5' a 3' de la Taq polimerasa desplazará después la sonda TaqMan® mediante la digestión para liberar la molécula reportera fluorescente para la detección cuantitativa de su señal ahora inactivada mediante la utilización de un sistema de detección de fluorescencia en tiempo real.

10

15

Los reactivos típicos (*por ejemplo*, tal como se pueden encontrar en un estuche típico basado en MethyLight™) para el análisis de MethyLight™ puede incluir, pero sin limitarse a: iniciadores de PCR para el gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); sondas TaqMan® o Lightcycler®; amortiguadores y desoxinucleótidos de PCR optimizados; y Taq polimerasa.

20

El ensayo QM™ (metilación cuantitativa) es una prueba cuantitativa alternativa para los patrones de metilación en muestras de ADN genómico, en donde la discriminación de secuencia ocurre al nivel de la hibridación de la sonda. En esta versión cuantitativa, la reacción de PCR se proporciona para la amplificación no-sesgada en presencia de una sonda fluorescente que solapa un sitio potencial de metilación en particular. Se proporciona un control no-sesgado para la cantidad de ADN de entrada por una reacción en la que ni los iniciadores, ni la sonda revisten ninguno de los dinucleótidos CpG. Alternativamente, una prueba cualitativa para la metilación genómica se logra mediante el sondeo de la mezcla de PCR sesgada, ya sea con oligonucleótidos control que no "revisten" sitios de metilación que se conozcan (una versión basada en fluorescencia de las técnicas HeavyMethyl™ y MSP), o con oligonucleótidos oligonucleótidos que revisten los posibles sitios de metilación.

25

30

El proceso QM™ puede utilizarse con cualquier sonda adecuada por ejemplo "TaqMan®", Lightcycler® etc... en el proceso de amplificación. Por ejemplo, el ADN genómico bicatenario se trata con bisulfito de sodio y se somete a iniciadores no sesgados y a la sonda TaqMan®. La sonda TaqMan® se marca doble con las moléculas fluorescentes "reportero" e "inactivador", y se diseña para ser específica de una región GC de contenido relativamente alto tal que se funde a aproximadamente a 10 °C de temperatura más alta en el ciclo de PCR que los iniciadores directo o reverso. Esto permite que la sonda TaqMan® permanezca completamente hibridada durante la etapa de extensión de la hibridación PCR. Debido a que la Taq polimerasa sintetiza enzimáticamente una nueva cadena durante el PCR, eventualmente se logrará la sonda TaqMan® apareada. La actividad endonucleasa 5' a 3' de la Taq polimerasa desplazará después la sonda TaqMan® mediante la digestión para liberar la molécula reportera fluorescente para la detección cuantitativa de su señal ahora inactivada mediante la utilización de un sistema de detección de fluorescencia en tiempo real. Los reactivos típicos (*por ejemplo*, tal como se pueden encontrar en un estuche típico basado en QM™) para el análisis de QM™ puede incluir, pero sin limitarse a: iniciadores de PCR para el gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); sondas TaqMan® o Lightcycler®; amortiguadores y desoxinucleótidos de PCR optimizados; y Taq polimerasa.

35

40

45

Ms-SNuPE. La técnica Ms-SNuPE™ es un método cuantitativo para evaluar diferencias de metilación en sitios CpG específicos basados en el tratamiento del ADN con bisulfito, seguido por la extensión del iniciador de un solo nucleótido (Gonzalzo & Jones, Nucleic Acids Res. 25:2529-2531, 1997). En resumen, el ADN genómico se hace reaccionar con bisulfito sódico para convertir la citosina no metilada a uracilo, dejando la 5-metilcitosina sin cambios. La amplificación de la secuencia objetivo que se desea se realiza después mediante la utilización de iniciadores de PCR específicos para ADN convertido con bisulfito, y el producto resultante se aísla y se utiliza como molde para el análisis de metilación en el(os) sitio(s) CpG de interés. Pequeñas cantidades de ADN se pueden analizar (*por ejemplo*, secciones de patología microdisecadas), y esto evita la utilización de enzimas de restricción para determinar la condición de metilación en los sitios CpG.

50

55

Los reactivos típicos (*por ejemplo*, como pueden encontrarse en un estuche típico basado en Ms-SNuPE™) para el análisis de Ms-SNuPE™ pueden incluir, pero sin limitarse a: iniciadores de PCR para el gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); amortiguadores y desoxinucleótidos de PCR optimizados; estuche de ™ extracción en gel; iniciadores control positivo; iniciadores Ms-SNuPE para gen específico; amortiguador de reacción (para la reacción Ms-

60

SNUPE); y nucleótidos marcados. Además, los reactivos de conversión con bisulfito pueden incluir: amortiguador de desnaturalización de ADN; amortiguador sulfonación; reactivos o estuche de recuperación de ADN (por ejemplo, precipitación, ultrafiltración, columna de afinidad); amortiguador desulfonación; y componentes de recuperación de ADN.

5 Se determinó que la secuencia genómica de acuerdo con la SEQ ID NO:24, y variantes tratadas de origen no natural de estas de acuerdo con las SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43 tienen una nueva utilidad para la detección temprana, clasificación y/o tratamiento del cáncer colorrectal.

10 El ADN genómico se puede aislar mediante cualquier medio estándar en la técnica, lo que incluye la utilización de estuches disponibles comercialmente. En resumen, en donde el ADN de interés se encapsula mediante una membrana celular la muestra biológica se debe romper y lisar mediante medios enzimáticos, químicos o mecánicos. La solución de ADN se puede clarificar de las proteínas y otros contaminantes, *por ejemplo*, mediante digestión con proteinasa K. El ADN genómico se recupera después de la solución. Esto puede llevarse a cabo por medio de una variedad de métodos que incluyen precipitación salina, extracción orgánica o unión de ADN a un soporte de fase sólida. La elección del método se afectará por
15 varios factores, que incluyen el tiempo, costo y cantidad requerida de ADN. Todos los tipos de muestras clínicas que comprenden materia neoplásica o materia pre-neoplásica son adecuados para usar en el presente método, y se prefiere las líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, heces, efluente colónico, orina, plasma sanguíneo, suero sanguíneo, sangre completa, células de sangre aisladas, células aisladas de la sangre y combinaciones de estos. Los fluidos corporales son la fuente que se prefiere de ADN; y particularmente se prefiere
20 plasma sanguíneo, suero sanguíneo, sangre completa, células de sangre aisladas y células aisladas de la sangre.

La muestra de ADN genómico se trata después con al menos un reactivo, o serie de reactivos que distinguen entre dinucleótidos CpG metilados y no metilados dentro de al menos una región objetivo del ADN genómico, en donde la región
25 objetivo comprende, o hibrida bajo condiciones rigurosas a una secuencia de al menos 16 nucleótidos contiguos de al menos una secuencia que se selecciona del grupo que consiste en la SEQ ID NO:24, en donde dichos nucleótidos contiguos comprenden al menos una secuencias de dinucleótidos CpG.

Se prefiere particularmente que dicho reactivo convierta las bases de citosina no metiladas en la posición 5' en uracilo, timina, u otra base que es diferente a la citosina en términos de comportamiento de hibridación. Sin embargo en una
30 modalidad alternativa dicho reactivo puede ser una enzima de restricción sensible a la metilación.

La muestra de ADN genómico se trata de manera tal que las bases de citosina que están sin metilar en la posición 5' se convierten a uracilo, timina, u otra base que es muy diferente a la citosina en términos de comportamiento de hibridación. Se prefiere que este tratamiento se lleve a cabo con bisulfito (sulfito de hidrógeno, disulfito) y una posterior hidrólisis alcalina.
35 Un tratamiento de este tipo resulta en la conversión de la SEQ ID NO:24 a SEQ ID NO:30, SEQ ID NO:31, en donde dichos dinucleótidos CpG son metilados o de SEQ ID NO:42 a SEQ ID NO:43, en donde dichos dinucleótidos CpG son no metilados.

A continuación se analiza el ADN que se trató con el fin de determinar el estado de metilación de las secuencias génicas objetivo (al menos un gen o secuencia genómica que se selecciona del grupo que consiste en FOXL2 antes del tratamiento). Se prefiere particularmente que la región objetivo comprenda, o hibride bajo condiciones rigurosas a por lo menos 16 nucleótidos contiguos del gen o secuencia genómica de FOXL2. Se prefiere que la secuencia de dicho gen SEQ ID NO: 24 se analice. El método de análisis se puede seleccionar a partir de los que se conocen en la técnica, que incluyen los que se listan en la presente descripción. Particularmente se prefieren MethyLight™, MSP y el uso de oligonucleótidos bloqueadores (HeavyMethyl™) como se describe en la presente descripción. Se prefiere además que cualquier oligonucleótido que se utilice en este tipo de análisis (lo que incluye iniciadores, oligonucleótidos de bloqueo y sondas de detección) deben ser complementarios de manera inversa, idénticos, o hibridar en condiciones rigurosas o altamente rigurosas con un segmento de al menos 16 pares de bases de largo de las secuencias base de una o más de SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID
40 NO: 42, SEQ ID NO: 43, y secuencias complementarias de las mismas.

La metilación aberrante, más específicamente la hipermetilación del gen o secuencia genómica de FOXL2 (lo que incluye a su promotor y/o regiones reguladoras) se asocia con la presencia de carcinomas colorrectales. En consecuencia en donde una muestra biológica presente dentro cualquier grado de metilación, dicha muestra se debe determinar como neoplásica.
50

La neoplasia de colon se define en la presente descripción como todos los cánceres de colon y adenomas de más de 1 cm., o subconjuntos de los mismos. Los negativos pueden definirse como individuos sanos.
55

En una modalidad el método describe el uso del gen o secuencia genómica de FOXL2, (o del promotor y/o regiones reguladoras del mismo) como un marcador para la diferenciación, la detección y distinción del cáncer de colon.

Particularmente se prefiere el uso de microarreglos. El proceso de análisis de microarreglos se puede dividir en dos partes principales. La primera es la inmovilización de las secuencias génicas que se conocen en portaobjetos de vidrio u otro soporte sólido seguido por la hibridación del ADNc marcado con fluorescencia (que comprende las secuencias que se interrogan) con los genes que se conocen inmovilizados en el portaobjetos de vidrio (u otra fase sólida). Después de la hibridación, los arreglos se digitalizan mediante un escáner de arreglo micro fluorescente. El análisis de la intensidad de fluorescencia relativa de diferentes genes proporciona una medida de las diferencias en la expresión génica.

Los arreglos de ADN pueden generarse mediante la inmovilización de oligonucleótidos que se presintetizaron en portaobjetos de vidrio preparados u otras superficies sólidas. En este caso, las secuencias de genes representativos se fabrican y preparan mediante la utilización de la síntesis estándar de oligonucleótidos y métodos de purificación. Estas secuencias de genes sintetizados son complementarias a los ARN transcritos de los genes de interés (en este caso los genes o secuencias genómicas se seleccionan de entre el grupo de FOXL2) y tienden a ser secuencias más cortas en el intervalo de 25-70 nucleótidos. Alternativamente, los oligos inmovilizados pueden sintetizarse químicamente in situ sobre la superficie del portaobjetos. La síntesis de oligonucleótidos in situ implica la adición consecutiva de los nucleótidos adecuados a los puntos en el microarreglo, los puntos que no reciben un nucleótido se protegen durante cada etapa del proceso mediante la utilización máscaras físicas o virtuales. Preferentemente dichos ácidos nucleicos sintetizados son ácidos nucleicos bloqueados.

En experimentos de perfiles de expresión de microarreglos, los moldes de ARN que se utilizan son representativos del perfil de transcripción de las células o tejidos en estudio. El ARN se aisló por primera vez de las poblaciones de células o tejidos que se comparan. Cada muestra de ARN se utiliza después como un molde para generar el ADNc marcado con fluorescencia a través de una reacción de transcripción inversa. El marcaje fluorescente del ADNc puede realizarse por tanto marcaje directo como por métodos de marcaje indirecto. Durante el marcaje directo, los nucleótidos modificados con fluorescencia (*por ejemplo*, Cy[®]3- o Cy[®]5-dCTP) se incorporan directamente en el ADNc durante la transcripción inversa. Alternativamente, el marcaje indirecto se puede lograr mediante la incorporación de nucleótidos aminoalil-modificados durante la síntesis del ADNc y conjugar después un colorante N-hidroxisuccinimida (NHS)-éster al ADNc aminoalil-modificado después que se completa la reacción de transcripción inversa. Alternativamente, la sonda puede no marcarse, pero puede ser detectable por la unión específica con un ligando que se marca, ya sea directa o indirectamente. Los marcadores adecuados y los métodos para marcar ligandos (y sondas) se conocen en la técnica, e incluyen, por ejemplo, marcadores radiactivos que se pueden incorporar por métodos que se conocen (por ejemplo, desplazamiento de mella o tratamiento con quinasa). Otros marcadores adecuados incluyen, pero sin limitarse a biotina, grupos fluorescentes, grupos quimioluminiscentes (por ejemplo, dioxetanos, particularmente dioxetanos activados), enzimas, anticuerpos, y similares.

Para realizar el análisis de la expresión génica diferencial, el ADNc que se genera a partir de diferentes muestras de ARN se marca con Cy[®]3. El ADNc resultante marcado se purifica para eliminar nucleótidos no incorporados, colorante libre y ARN residual. Después de la purificación, las muestras de ADNc marcadas se hibridan al microarreglo. La rigurosidad de la hibridación se determina por una serie de factores durante la hibridación y durante el procedimiento de lavado, incluyendo la temperatura, fuerza iónica, longitud de tiempo y concentración de formamida. Estos factores se describen en, por ejemplo, Sambrook y otros (Molecular Cloning: A Laboratory Manual, 2da ed., 1989). El microarreglo se escanea post-hibridación usando un escáner fluorescente para microarreglos. La intensidad de fluorescencia de cada punto indica el nivel de expresión del gen analizado; manchas brillantes corresponden a genes fuertemente expresados, mientras que las manchas oscuras indican expresión débil.

Una vez que se obtienen las imágenes, se deben analizar los datos sin procesar. En primer lugar, la fluorescencia de fondo se debe sustraer de la fluorescencia de cada punto. Los datos se normalizan a una secuencia control, tales como ácidos nucleicos que se añaden exógenamente (preferentemente ARN o ADN), o un panel de genes de mantenimiento para dar cuenta de cualquier hibridación no específica, imperfecciones del arreglo o variabilidad en el arreglo de la configuración, marcaje del ADNc, hibridación o lavado. La normalización de datos permite que se comparen los resultados de múltiples arreglos.

Modalidades particulares de la presente invención proporcionan una nueva aplicación del análisis de los niveles de metilación y/o patrones dentro de dichas secuencias que permite una detección precisa, la caracterización y/o el tratamiento del cáncer colorrectal. La detección temprana del cáncer se relaciona directamente con el pronóstico de la enfermedad, y el método descrito de ese modo permite al médico y paciente tomar más y mejores decisiones del tratamiento informado.

Mejoras adicionales

La presente invención proporciona nuevos usos para la secuencia genómica SEQ ID NO:24.

Un objetivo de la invención comprende el análisis del estado de metilación de uno o más dinucleótidos CpG dentro de la secuencia de la SEQ ID NO: 24 y secuencias complementarias a ella.

5 También se describen ácidos nucleicos tratados, derivados de la SEQ ID NO:24 genómica, en donde el tratamiento es adecuado para convertir al menos una base de citosina no metilada de la secuencia de ADN genómico a uracilo u otra base que se detecta diferente a la citosina en términos de hibridación. Las secuencias genómicas en cuestión pueden comprender uno o más posiciones CpG metiladas consecutivas. Dicho tratamiento comprende preferentemente el uso de un reactivo que se selecciona del grupo que consiste en bisulfito, sulfito de hidrógeno, disulfito, y combinaciones de estos. Se describe un ácido nucleico modificado de origen no natural que comprende una secuencia de al menos 16 bases de nucleótidos contiguos de longitud de una secuencia que se selecciona del grupo que consiste en SEQ ID NOS: 30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43. Dicho ácido nucleico es al menos 50, 100, 150, 200, 250 o 500 pares de base de longitud de un segmento de la secuencia de ácido nucleico descrita en SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43. Se prefiere particularmente una molécula de ácido nucleico que no es idéntica o complementaria a toda o una parte de la secuencias SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43, pero no la SEQ ID NO:24 u otro ADN de origen natural.

Se prefiere que dicha secuencia comprenda, al menos, un dinucleótido CpG, TpA o CpA y secuencias complementarias a estas. Las secuencias de SEQ ID NO: 30 a SEQ ID NO:31, SEQ ID NO:42 a SEQ ID NO:43 proporcionan versiones modificadas que son de origen no natural del ácido nucleico de acuerdo con la SEQ ID NO:24, en donde la modificación de cada secuencia genómica resulta en la síntesis de un ácido nucleico que tiene una secuencia que es única y distinta de dicha secuencia genómica como sigue. Para cada ADN genómico de cadena sentido, por ejemplo, SEQ ID NO:24, se describen cuatro versiones convertidas. Una primera versión donde "C" se convierte a "T", pero "CpG" sigue siendo "CpG" (*es decir*, corresponde al caso donde, para la secuencia genómica, todos los residuos "C" de secuencias de dinucleótidos CpG se metilan y son así no convertidas); una segunda versión describe el complemento de la secuencia de ADN genómico descrito (es decir, cadena antisentido), en donde "C" se convierte a "T", pero "CpG" sigue siendo "CpG" (*es decir*, corresponde al caso donde, para todos los residuos "C" de secuencias de dinucleótidos CpG se metilan y así no se convierten). Las secuencias convertida no metiladas de SEQ ID NO:24 corresponden a las SEQ ID NOS: 30 a SEQ ID NO:31. Se proporciona una tercera versión de cada secuencia genómica, en donde "C" se convierte a "I" para todos los residuos "C", que incluyen las secuencias de dinucleótidos "CpG" (*es decir*, corresponde al caso donde, para las secuencias genómicas, todos los residuos "C" de secuencias de dinucleótidos CpG son no metilados); una versión químicamente convertida final de cada secuencia, describe el complemento de la secuencia de ADN genómico descrita (por ejemplo cadena antisentido), en donde "C" convierte a "T" para todos los residuos "C", que incluyen las secuencias de dinucleótidos "CpG" (*es decir*, corresponde al caso donde, para el complemento (cadena antisentido) de cada secuencia genómica, todos los residuos "C" de secuencias de dinucleótidos CpG son no metiladas). Las secuencias convertidas 'metiladas descendientemente' de SEQ ID NO:24 corresponde a las SEQ ID NOS:42 a SEQ ID NO:43.

Significativamente, en otro tiempo, las secuencias y moléculas de ácido nucleico de acuerdo con las SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43, no estaban implicadas en relacionadas con la detección, clasificación o tratamiento de trastornos proliferativos celulares.

Se describen además oligonucleótidos u oligómeros adecuados para utilizar en los métodos para detectar el estado de metilación de la citosina dentro del ADN genómico o tratado (que se modifica químicamente), de acuerdo con la SEQ ID NO:24, SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43. Dichos ácidos nucleicos de oligonucleótido u oligómero proporciona nuevos medios diagnósticos. Dicho oligonucleótido u oligómero que comprende una secuencia de ácido nucleico que tiene una longitud de al menos nueve (9) nucleótidos que es idéntica a, hibrida, en condiciones moderadamente rigurosas o rigurosas (como se define en la presente descripción anteriormente), a una secuencia ácido nucleico tratada de acuerdo con las SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43, y/o secuencias complementarias a estas, o a una secuencia genómica de acuerdo con las SEQ ID NOS:24 y/o secuencias complementarias a estas.

Se describen moléculas de ácido nucleico (por ejemplo, oligonucleótidos y moléculas de ácido nucleico péptido (PNA) (PNA-oligómeros)) que se hibridan en condiciones de hibridación moderadamente rigurosas y/o rigurosas a toda o una porción de la secuencia que se selecciona del grupo que consiste en SEQ ID NO:24, SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43, o a los complementos de estas. Se describe particularmente una molécula de ácido nucleico que hibrida en condiciones de hibridación moderadamente rigurosas y/o rigurosas a toda o una porción de la secuencias SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43, pero no a la SEQ ID NO:24 u otro ADN genómico humano.

La porción idéntica o de hibridación de los ácidos nucleicos de hibridación es típicamente de al menos 9, 16, 20, 25, 30 o 35 nucleótidos de longitud.

Los ácidos nucleicos de hibridación del tipo descrito en la presente se pueden usar, por ejemplo, como un iniciador (por ejemplo un iniciador de PCR), o un diagnóstico y/o una sonda o iniciador pronóstico. Preferentemente, la hibridación de la sonda de oligonucleótido a una muestra de ácido nucleico se lleva a cabo en condiciones rigurosas y la sonda es 100% idéntica a la secuencia objetivo. La estabilidad del ácido nucleico dúplex o híbrido se expresa como la temperatura de fusión o T_m , que es la temperatura a la que una sonda se disocia de un ADN objetivo. Esta temperatura de fusión se utiliza para definir las condiciones de rigurosidad que se requieren.

Para las secuencias objetivo que se relacionan y sustancialmente idénticas a la secuencia correspondiente de SEQ ID NO:24 (tales como variantes alélicas y SNPs), en lugar de idénticas, es útil establecer primero la temperatura más baja a la que sólo se produce la hibridación homóloga con una concentración de sal particular (por ejemplo, SSC o SSPE). Después, suponiendo que el 1 % de desajustes resulta en una disminución de 1 °C en la T_m , la temperatura del lavado final en la reacción de hibridación se reduce como consecuencia (por ejemplo, si se buscan las secuencias que tienen > 95% de identidad con la sonda, la temperatura de lavado final se disminuye por 5 °C). En la práctica, el cambio en la T_m puede ser entre 0.5 °C y 1.5 °C por 1% de desajuste.

Los ejemplos de oligonucleótidos descritos de longitud X (en nucleótidos), como se indicó por las posiciones de polinucleótido con referencia a, por ejemplo, SEQ ID NO:24, incluyen los correspondientes a los conjuntos (conjuntos sentido y antisentido) de oligonucleótidos de solapamiento de longitud X consecutivos, donde los oligonucleótidos dentro de cada conjunto de solapamiento consecutivo (correspondiente a un valor X dado) se definen como el conjunto finito de oligonucleótidos Z a partir de las posiciones de nucleótidos:

n para $(n + (X-1))$;

donde $n=1, 2, 3, \dots (Y (X-1))$;

donde Y es igual a la longitud (nucleótidos o pares de bases) de la SEQ ID NO: 24 (2804);

donde X es igual a la longitud común (en nucleótidos) de cada oligonucleótido en el conjunto (por ejemplo, $X=20$ para un conjunto de 20-mers de solapamiento consecutivo); y

donde el número (Z) de oligómeros de solapamiento consecutivos de longitud X para una determinada SEQ ID NO de longitud Y es igual a $Y-(X-1)$. Por ejemplo $Z = 2804-19 = 2785$ para cualquiera de los conjuntos sentido o antisentido de SEQ ID NO: 24, donde $X=20$.

Preferentemente, el conjunto se limita a los oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.

Los ejemplos de oligonucleótidos de 20-mer incluyen el siguiente conjunto de 2785 oligómeros (y el conjunto antisentido complementario a este), que se indica por las posiciones de polinucleótido con referencia a la SEQ ID NO: 24:

1-20, 2-21, 3-22, 4-23, 5-24, y 2785-2804.

Preferentemente, el conjunto se limita a los oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.

Del mismo modo, los ejemplos descritos de oligonucleótidos de 25-mer incluyen el siguiente conjunto de 2780 oligómeros (y el conjunto antisentido complementario a este), que se indica por las posiciones de polinucleótido con referencia a la SEQ ID NO:24:

1-25, 2-26, 3-27, 4-28, 5-29, y 2780-2804.

Preferentemente, el conjunto se limita a los oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.

Se describen, para cada una de la SEQ ID NO:24, SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43 (sentido y antisentido), múltiples conjuntos de solapamiento consecutivos de oligonucleótidos u oligonucleótidos modificados de longitud X, donde, por ejemplo, $X= 9, 10, 17, 20, 22, 23, 25, 27, 30$ o 35 nucleótidos.

Los oligonucleótidos u oligómeros como los descritos constituyen herramientas eficaces útil para determinar parámetros genéticos y epigenéticos de la secuencia genómica SEQ ID NO:24. Los conjuntos que se prefieren de tales oligonucleótidos u oligonucleótidos modificados de longitud X son aquellos conjuntos de solapamiento consecutivos de oligómeros correspondientes a las SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43 (y a los complementos de estos). Preferentemente, dichos oligómeros comprenden al menos un dinucleótido CpG, TpG o CpA.

Particularmente los oligonucleótidos u oligómeros descritos son en los que la citosina de las secuencias de dinucleótido CpG (o del correspondiente dinucleótido convertido TpG o CpA) está dentro del tercio medio del oligonucleótido; es decir, donde

el oligonucleótido es, por ejemplo, 13 bases de longitud, el dinucleótido CpG, TpG o CpA se coloca dentro del quinto al noveno nucleótido desde el extremo 5'.

5 Los oligonucleótidos que se describen en la presente también se pueden modificar por el enlace químico del oligonucleótido a una o más porciones o conjugados para mejorar la actividad, estabilidad o detección del oligonucleótido. Tales porciones o conjugados incluyen cromóforos, fluoróforos, lípidos tales como colesterol, ácido cólico, tioéter, cadenas alifáticas, fosfolípidos, poliaminas, polietilenglicol (PEG), porciones de palmitilo, y otros como se escribe en, por ejemplo, las Patentes de Estados Unidos Números 5,514,758, 5,565,552, 5,567,810, 5,574,142, 5,585,481, 5,587,371, 5,597,696 y 5,958,773. Las sondas pueden existir también en forma de un PNA (ácido nucleico peptídico) que tiene propiedades de apareamiento que se prefieren particularmente. Así, el oligonucleótido puede incluir otros grupos adjuntos tales como péptidos, y puede incluir agentes de escisión desencadenantes de la hibridación (Krol y otros, BioTechniques 6:958-976, 1988) o agentes intercalantes (Zon, Pharm. Res. 5:539-549, 1988). Para este fin, el oligonucleótido se puede conjugar a otra molécula, por ejemplo, un cromóforo, fluoróforo, péptido, agente de entrecruzamiento desencadenante de la hibridación, agente de transporte, agente de escisión desencadenante de la hibridación etc.

15 El oligonucleótido puede comprender también al menos un azúcar modificada que se reconoce en la técnica y/o porción de base, o puede comprender una cadena principal modificada o un enlace internucleosídico no natural.

20 Los oligonucleótidos u oligómeros como los descritos se usan típicamente en 'conjuntos,' que contienen al menos un oligómero para el análisis de cada uno de los dinucleótidos CpG de la secuencia genómica SEQ ID NO:24 y las secuencias complementarias a estas, o a los dinucleótidos CpG, TpG o CpA correspondientes dentro de una secuencia de los ácidos nucleicos tratados de acuerdo con las SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43 y secuencias complementarias a estas. Sin embargo, se prevé que para los factores económicos u otros puede ser preferible analizar una selección limitada de los dinucleótidos CpG dentro de dichas secuencias, y el contenido del conjunto de oligonucleótidos se altera como consecuencia.

30 Por lo tanto, en modalidades particulares, se describe un conjunto de al menos dos (2) (oligonucleótidos y/o oligómeros-PNA) útiles para detectar el estado de metilación de citosina en ADN genómico tratado (SEQ ID NO.:30 a SEQ ID NO:31, SEQ ID NO:42 a SEQ ID NO:43), o en el ADN genómico (SEQ ID NO:24 y las secuencias complementarias a estas). Estas sondas permiten el diagnóstico, clasificación y/o terapia de parámetros genéticos y epigenéticos de trastornos proliferativos celulares del hígado y/o colorrectales. El conjunto de oligómeros pueden usarse también para detectar polimorfismos de nucleótido (SNPs) en ADN genómico tratado (SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NO:42 a SEQ ID NO:43), o en el ADN genómico (SEQ ID NO:24 y las secuencias complementarias a estas).

35 En modalidades que se prefieren, al menos uno, y con mayor preferencia todos los miembros de un conjunto de oligonucleótidos se unen a una fase sólida.

40 En modalidades adicionales, la presente invención describe un conjunto de al menos dos (2) oligonucleótidos que se usan como oligonucleótidos 'iniciadores' para amplificar secuencias de ADN de una de SEQ ID NO:24, SEQ ID NOS:30 a SEQ ID NO:31, SEQ ID NOS:42 a SEQ ID NO:43, y las secuencias complementarias a estas, o segmentos de estos.

45 Se prevé que los oligonucleótidos pueden constituir toda o parte de una "matriz" o "matriz de ADN" (es decir, un arreglo de diferentes oligonucleótidos y/o PNA-oligómeros unidos a una fase sólida). Se puede caracterizar por ejemplo una matriz de ese tipo de diferentes secuencias de oligonucleótido- y/o PNA-oligómero en que se dispone sobre la fase sólida en forma de una red rectangular o hexagonal. La superficie de fase sólida puede estar compuesta de silicio, vidrio, poliestireno, aluminio, acero, hierro, cobre, níquel, plata, u oro. La nitrocelulosa así como plásticos tales como nailon, que pueden existir en forma de gránulos o también como matrices de resina, también se pueden utilizar. Una visión general de la técnica anterior en la fabricación de matriz de oligómero pueden obtenerse de una edición especial de Nature Genetics (Nature Genetics Supplement, volumen 21, enero 1999, y de la bibliografía citada en ella). Sondas marcadas con fluorescencia se usan frecuentemente para el escaneo de matrices de ADN inmovilizadas. La fijación simple de colorantes Cy3 y Cy5 al 5'-OH de la sonda específica es particularmente adecuada para los marcadores de fluorescencia. La detección de la fluorescencia de las sondas hibridadas se puede llevar cabo, por ejemplo, a través de un microscopio confocal. Los colorantes Cy3 y Cy5, además de muchos otros, están disponibles comercialmente.

55 También se prevé que los oligonucleótidos o secuencias particulares de estos, pueden constituir toda o parte de una "matriz virtual" en donde se utilizan los oligonucleótidos o secuencias particulares de estos, por ejemplo, como "especificadores" como parte de, o en combinación con una población diversa de sondas marcadas únicas para analizar una mezcla compleja de analitos. Tal método se describe, por ejemplo, en US 2003/0013091 (Estados Unidos serie número 09/898,743, que se publicó el 16 de enero de 2003). En tales métodos, se generan suficientes marcadores de manera que cada ácido nucleico

en la mezcla compleja (es decir, cada analito) se puede unir de forma única por un marcador único y así detecta (cada marcador se cuenta directamente, resultando en una lectura digital de cada especie molecular en la mezcla).

Se prefiere particularmente que los oligómeros como se describe se utilicen para al menos uno de: detección de; detección y diferenciación entre dos o más subclases de; diagnóstico de; pronóstico de; tratamiento de; seguimiento de; y tratamiento y seguimiento de trastornos celulares proliferativos colorrectales y/o del hígado. Esto es posible mediante el uso de dichos conjuntos para la detección o la detección y diferenciación de una o más de las siguientes clases de tejidos: carcinoma colorrectal, adenoma de colon, tejido del colon inflamatorio, adenomas de colon con displasia grado 2 de menos de 1 cm, adenomas de colon con displasia grado 3 mayores de 1 cm, tejido de colon normal, tejido sano que no pertenece al colon y tejido de cáncer que no pertenece al colon.

Se prefieren particularmente aquellos conjuntos de oligómeros de acuerdo con los Ejemplos.

En la modalidad que más prefiere del método, se determina la presencia o ausencia de cáncer colorrectal o la diferenciación del mismo a partir de trastornos benignos. Esto se logra mediante el análisis del estado de metilación de una secuencia objetivo que comprende al menos una posición CpG comprendiendo dicha secuencia, o hibridándose bajo condiciones rigurosas con al menos 16 nucleótidos contiguos de la secuencia de la SEQ ID NO: 24 y complementos de la misma.

En una modalidad preferida, dicho método comprende las siguientes etapas: En la *primera etapa*, se obtiene una muestra de tejido que se analiza. La fuente puede ser cualquier fuente adecuada, tal como líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, heces, efluente colónico, orina, plasma sanguíneo, suero sanguíneo, sangre completa, células de sangre aisladas, células aisladas de la sangre y todas las combinaciones posibles de estos. Se prefiere que dichas fuentes de ADN sean heces o fluidos corporales que se seleccionan del grupo que consiste en efluente colónico, orina, plasma sanguíneo, suero sanguíneo, sangre completa, células de sangre aisladas, células aisladas de la sangre.

El ADN genómico se aísla después de la muestra. El ADN genómico se puede aislar mediante cualquier medio estándar en la técnica, lo que incluye el uso de estuches disponibles comercialmente. Brevemente, en donde el ADN de interés se encapsula mediante una membrana celular la muestra biológica se debe romper y lisar mediante medios enzimáticos, químicos o mecánicos. La disolución de ADN se puede clarificar de las proteínas y otros contaminantes por ejemplo mediante digestión con proteinasa K. El ADN genómico se recupera después de la disolución. Esto puede llevarse a cabo por medio de una variedad de métodos que incluyen precipitación salina, extracción orgánica o unión de ADN a un soporte de fase sólida. La elección del método se afectará por varios factores, que incluyen el tiempo, costo y cantidad requerida de ADN.

En donde la muestra de ADN no se encierra en una membrana (por ejemplo, ADN circulante a partir de una muestra de sangre) pueden emplearse los métodos estándar en la técnica para el aislamiento y/o purificación de ADN. Tales métodos incluyen el uso de un reactivo degenerador de proteínas por ejemplo, sal caotrópica por ejemplo hidrocloreto de guanidina o urea; o un detergente por ejemplo dodecil sulfato de sodio (SDS), bromuro de cianógeno. Los métodos alternativos incluyen pero no se limitan a la precipitación con etanol o precipitación con propanol, concentración al vacío entre otros por medio de una centrífuga. La persona experta en la técnica también puede hacer uso de dispositivos tales como dispositivos de filtro, por ejemplo, ultrafiltración, membranas o superficies de sílice, partículas magnéticas, partículas de poliestireno, superficies de poliestireno, superficies cargadas positivamente, y membranas cargadas positivamente, membranas cargadas, superficies cargadas, membranas cargadas conmutadas, superficies cargadas conmutadas.

En la *segunda etapa* del método la muestra de ADN genómico se trata después de manera tal que las bases de citosina que están sin metilar en la posición 5' se convierten a uracilo, timina, u otra base que es muy diferente a la citosina en términos de comportamiento de hibridación. Esto se entenderá como "pretratamiento" o "tratamiento" en la presente descripción.

Esto se logra preferentemente por medio de tratamiento con un reactivo de bisulfito. El término "reactivo de bisulfito" se refiere a un reactivo que comprende bisulfito, disulfito, sulfito de hidrógeno o combinaciones de estos, útiles como se describe en la presente descripción para distinguir entre secuencias de dinucleótidos CpGs metiladas y no metiladas. Los métodos de dicho tratamiento se conocen en la técnica (por ejemplo, PCT/EP2004/011715, que se incorpora por referencia en su totalidad). Se prefiere que el tratamiento con bisulfito se lleve a cabo en presencia de disolventes desnaturalizantes tales como, pero sin limitarse al n-alkilenglicol, particularmente dietilenglicol dimetil glicol éter (DME), o en presencia de derivados de dioxano o dioxano. En una modalidad preferida, los disolventes desnaturalizantes se utilizan en concentraciones de entre 1% y 35% (v/v). También se prefiere que la reacción de bisulfito se lleve a cabo en presencia de sequestradores tales como, pero sin limitarse a derivados de cromano, *por ejemplo*, ácido 6-hidroxi-2, 5,7,8, -tetrametilcromano 2-carboxílico o ácido trihidroxibenzoico y derivados del mismo, *por ejemplo*, ácido Gálico (ver: PCT/EP2004/011715 que se incorpora por referencia en su totalidad). La conversión de bisulfito se lleva a cabo

preferentemente a una temperatura de reacción entre 30 °C y 70 °C, de manera que la temperatura se aumenta a más de 85 °C durante períodos cortos de tiempo durante la reacción (ver: PCT/EP2004/011715 que se incorpora por referencia en su totalidad). El ADN que se trata con bisulfito se purifica preferentemente antes de la cuantificación. Esto puede llevarse a cabo por cualquier medio que se conozca en la técnica, tal como pero sin limitarse a la ultrafiltración, que se lleva a cabo preferentemente por medio de columnas MicroconTM (que se fabrican por MilliporeTM). La purificación se lleva a cabo según un protocolo del fabricante que se modificó (ver: PCT/EP2004/011715 que se incorpora por referencia en su totalidad).

En la *tercera etapa* del método, fragmentos del ADN tratado se amplifican, mediante la utilización de conjuntos de iniciadores de oligonucleótidos de acuerdo con la presente invención, y una enzima de amplificación. La amplificación de varios segmentos de ADN puede llevarse a cabo simultáneamente en uno y el mismo recipiente de reacción. Típicamente, la amplificación se lleva a cabo mediante la utilización de una reacción en cadena de la polimerasa (PCR). Preferentemente, dichos amplificadores son de 100 a 2,000 pares de bases de longitud. El conjunto de oligonucleótidos iniciadores incluye al menos dos oligonucleótidos cuyas secuencias son cada una complementarias de manera inversa, idénticas, o se hibridan en condiciones rigurosas o muy rigurosas a al menos un segmento de 16 pares de bases de longitud de una secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43 y secuencias complementarias de las mismas.

En una modalidad alternativa del método, el estado de metilación de las posiciones CpG que se preseleccionaron dentro de la SEQ ID NO: 24 se puede detectar mediante el uso de oligonucleótidos iniciadores metilación-específicos. Esta técnica (MSP) se describe en la Patente de Estados Unidos con núm. 6,265,171 a Herman. El uso de iniciadores específicos del estado de metilación para la amplificación de ADN tratado con bisulfito permite la diferenciación entre ácidos nucleicos metilados y no metilados. Los pares de iniciadores MSP contienen al menos un iniciador que se hibrida a un dinucleótido CpG tratado con bisulfito. Por lo tanto, la secuencia de dichos iniciadores comprende al menos un dinucleótido CpG. Los iniciadores MSP específicos para el ADN no metilado contiene una "T" en la posición de la posición C en el CpG. Preferentemente, por lo tanto, se requiere que la secuencia base de dichos iniciadores comprenda una secuencia que tiene una longitud de al menos de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de las SEQ ID NO: 30, SEQ ID NO:31, SEQ ID NO: 42, SEQ ID NO:43 y secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG. Una modalidad preferida adicional del método comprende el uso de oligonucleótidos bloqueadores (el ensayo HeavyMethylTM). El uso de tales oligonucleótidos bloqueadores se ha descrito por Yu y otros, BioTechniques 23:714-720, 1997. Los oligonucleótidos de la sonda de bloqueo se hibridan con el ácido nucleico tratado con bisulfito concurrentemente con los iniciadores de PCR. La amplificación por PCR del ácido nucleico se termina en la posición 5' de la sonda de bloqueo, tal que la amplificación de un ácido nucleico se suprime donde está presente la secuencia complementaria a la sonda de bloqueo. Las sondas se pueden diseñar para hibridar con el ácido nucleico que se trata con bisulfito de manera específica al estado de metilación. Por ejemplo, para la detección de ácidos nucleicos metilados dentro de una población de ácidos nucleicos no metilados, la supresión de la amplificación de ácidos nucleicos que no se metilan en la posición en cuestión se lleva a cabo mediante el uso de sondas de bloqueo que comprenden un 'CpA' o "TpA" en la posición en cuestión, a diferencia de un "CpG" si se desea la supresión de la amplificación de ácidos nucleicos metilados.

Para los métodos de PCR que utilizan oligonucleótidos bloqueadores, la ruptura eficaz de la amplificación mediada por polimerasa requiere que los oligonucleótidos bloqueadores no se alarguen por la polimerasa. Preferentemente, esto se logra a través del uso de bloqueadores que son 3'-desoxi oligonucleótidos, u oligonucleótidos derivatizados en la posición 3' con un grupo diferente a un grupo "libre" hidroxilo. Por ejemplo, los oligonucleótidos 3'-O-acetilo son representativos de una clase que se prefiere de molécula bloqueadora.

Además, se debe impedir la descomposición mediada por polimerasa de los oligonucleótidos bloqueadores. Preferentemente, tal impedimento comprende ya sea el uso de una polimerasa que carece de actividad 5'-3' exonucleasa, o el uso de oligonucleótidos bloqueadores modificados que tienen, por ejemplo, puentes tioato en el extremo 5' de estos que hacen resistente a las nucleasa a la molécula bloqueadora. Las aplicaciones particulares pueden no requerir tales modificaciones en el 5' del bloqueador. Por ejemplo, si los sitios de unión al bloqueador e iniciador se solapan, impidiendo de ese modo la unión del iniciador ((*por ejemplo*, con exceso de bloqueador), se impedirá sustancialmente la degradación del oligonucleótido bloqueador. Esto es porque la polimerasa no extenderá el iniciador hacia, y a través del (en la dirección 5'-3') bloqueador— un proceso que resulta normalmente en la degradación del oligonucleótido bloqueador hibridado.

Una modalidad de bloqueador/PCR que se prefiere particularmente para los propósitos de la presente invención y tal como se aplica en la presente descripción, comprende el uso de oligómeros de ácido nucleico peptídico (PNA) como oligonucleótidos bloqueadores. Tales oligómeros bloqueadores de PNA son idealmente adecuados, porque no son ni descompuestos ni extendidos por la polimerasa.

Preferentemente, por lo tanto, se requiere que la secuencia base de dichos oligonucleótidos de bloqueo comprenda una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratado de acuerdo con una de SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43, y secuencias complementarias a estas, en donde la secuencia base de dichos oligonucleótidos comprende al menos uno de los dinucleótidos CpG, TpG o CpA.

Los fragmentos que se obtienen por medio de la amplificación pueden llevar un marcador directo o indirectamente detectable. Se prefieren los marcadores en la forma de marcadores de fluorescencia, radionucleidos o fragmentos de moléculas separables que tienen una masa típica que se puede detectar en un espectrómetro de masas. Donde dichos marcadores son marcadores de masa, se prefiere que los amplificados marcados tengan una sola carga neta positiva o negativa, permitiendo una mejor detectabilidad en el espectrómetro de masas. La detección puede llevarse a cabo y visualizarse por medio de, *por ejemplo*, espectrometría de masas de desorción/ionización láser asistida por matriz (MALDI) o usando la espectrometría de masas con electrospray (ESI).

La Espectrometría de Masas de Desorción/ionización Láser Asistida por Matriz (MALDI-TOF) es un desarrollo muy eficaz para el análisis de biomoléculas (Karas & Hillenkamp, Anal Chem., 60:2299-301, 1988). Un analito se embebe en una matriz absorbente de luz. La matriz se evapora por un pulso corto de láser que transporta así la molécula de analito en la fase de vapor de manera no fragmentada. El analito se ioniza por colisiones con moléculas de la matriz. Un voltaje que se aplica acelera los iones en un tubo de vuelo en campo libre. Debido a sus diferentes masas, los iones se aceleran a diferentes velocidades. Los iones más pequeños alcanzan el detector más pronto que los más grandes. La espectrometría MALDI-TOF es muy adecuada para el análisis de péptidos y proteínas. El análisis de ácidos nucleicos es algo más difícil (Gut & Beck, Current Innovations and Future Trends, 1:147-57, 1995). La sensibilidad con respecto a los análisis de ácidos nucleicos es aproximadamente 100 veces menor que el de los péptidos, y disminuye de manera desproporcionada con el aumento de tamaño del fragmento. Más aun, para los ácidos nucleicos que tienen una cadena principal múltiple cargada negativamente, el proceso de ionización a través de la matriz es considerablemente menos eficaz. En la espectrometría MALDI-TOF, la selección de la matriz desempeña un papel eminentemente importante. Para la desorción de los péptidos, varias matrices muy eficaces se encontraron que producen una cristalización muy fina. Existen ahora varias matrices de respuesta para el ADN, sin embargo, no se redujo la diferencia de sensibilidad entre los péptidos y ácidos nucleicos. Esta diferencia en sensibilidad se puede reducir, sin embargo, modificando químicamente el ADN de manera tal que se vuelva más similar a un péptido. Por ejemplo, los ácidos nucleicos fosforotioato, en los que los fosfatos habituales de la cadena principal se sustituyen con tiofosfatos, se pueden convertir en un ADN de carga neutra usando la química de alquilación sencilla (Gut & Beck, Nucleic Acids Res. 23: 1367-73, 1995). El acoplamiento de una etiqueta de carga a este ADN modificado resulta en un aumento de la sensibilidad de MALDI-TOF al mismo nivel que el que se encontró para los péptidos. Una ventaja adicional del etiquetado de carga es el aumento de la estabilidad del análisis contra las impurezas, que hace considerablemente más difícil la detección de sustratos no modificados.

En la *cuarta etapa* del método, los amplificados que se obtienen durante la tercera etapa del método se analizan para determinar la condición de metilación de los dinucleótidos CpG antes del tratamiento.

En modalidades donde se obtuvieron los amplificados por medio de la amplificación MSP, la presencia o ausencia de un amplificado es en sí misma indicativa del estado de metilación de las posiciones de CpG cubiertas por el iniciador, de acuerdo con las secuencias de bases de dicho iniciador.

Los amplificados que se obtienen por medio de PCR específica de metilación y estándar pueden analizar adicionalmente por medio de métodos basados en la hibridación, tales como, pero sin limitarse a las tecnologías de matriz como tecnologías basadas en sonda así como por medio de técnicas tales como secuenciación y extensión dirigida por molde.

En una modalidad del método, los amplificados sintetizados en la etapa tres posteriormente se hibridan con una matriz o un conjunto de oligonucleótidos y/o sondas de PNA. En este contexto, la hibridación tiene lugar de la siguiente manera: el conjunto de sondas que se utilizan durante la hibridación se compone preferentemente de al menos 2 oligonucleótidos u oligómeros PNA; en el proceso, los amplificados sirven como sondas que se hibridan con oligonucleótidos previamente unidos a una fase sólida; los fragmentos no hibridados se eliminan posteriormente; dichos oligonucleótidos contienen al menos una secuencia base que tiene una longitud de al menos 9 nucleótidos que es complementaria inversa o idéntica a un segmento de las secuencias bases especificadas en la presente lista de secuencias; y el segmento comprende al menos un dinucleótido CpG, TpG o CpA. La porción de hibridación de los ácidos nucleicos de hibridación es típicamente de al menos 9, 15, 20, 25, 30 o 35 nucleótidos de longitud. Sin embargo, las moléculas más largas tienen utilidad inventiva, y se encuentran así dentro del alcance de la presente invención.

En una modalidad preferida, dicho dinucleótido está presente en el tercio central del oligómero. Por ejemplo, en donde el

oligómero comprende un dinucleótido CpG, dicho dinucleótido es preferentemente del quinto al noveno nucleótido del extremo 5' de un 13-mer. Un oligonucleótido existe para el análisis de cada dinucleótido CpG dentro de una secuencia que se selecciona de la SEQ ID NO:24, y las posiciones equivalentes dentro de una secuencia que se selecciona del grupo que consiste en la SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO:43.

Dichos oligonucleótidos también pueden estar presentes en forma de ácidos nucleicos peptídicos. A continuación se eliminan los amplificadores no hibridados. Los amplificadores hibridados se detectan. En este contexto, se prefiere que los marcadores que se unen a los amplificadores sean identificables en cada posición de la fase sólida en la que se encuentra una secuencia de oligonucleótidos.

En aún una modalidad adicional del método, la condición de metilación genómica de las posiciones de CpG se pueden determinar por medio de sondas de oligonucleótidos (como se detalló anteriormente) que se hibridan con el ADN tratado con bisulfito concurrentemente con los iniciadores de amplificación de PCR (en donde dichos iniciadores pueden ser ya sea de metilación específica o estándar).

Una modalidad particularmente preferida de este método es el uso de PCR Cuantitativo en Tiempo Real basado en la fluorescencia (Heid y otros, Genome Res. 6:986-994, 1996; *ver también* la patente de Estados Unidos núm. 6,331,393) que emplea una sonda de oligonucleótido fluorescente marcada dual (TaqMan™ PCR, usando un Sistema de Detección de Secuencia ABI Prism 7700, Perkin Elmer Applied Biosystems, Foster City, California). La reacción de PCR TaqMan™ emplea el uso de un oligonucleótido de interrogación no extendible, llamado sonda TaqMan™, el cual, en modalidades que se prefieren, se designa para hibridar con una secuencia rica en CpG que se sitúa entre los iniciadores de amplificación directo y reverso. La sonda TaqMan™ comprende además una "porción reportera" fluorescente y una "porción inactivadora" unida covalentemente a las porciones enlazadoras (*por ejemplo*, fosforamiditas) unidas a los nucleótidos del oligonucleótido TaqMan™. Para el análisis de metilación dentro de los ácidos nucleicos posterior al tratamiento bisulfito, se requiere que la sonda sea de metilación específica, como se describe en la patente de Estados Unidos núm. 6,331,393, conocido también como ensayo MethyLight™. Las variaciones sobre la metodología de detección TaqMan™ que también son adecuadas para usar con la invención descrita incluyen el uso de la tecnología de doble sonda (Lightcycler™) o iniciadores de amplificación fluorescente (tecnología Sunrise™). Ambas técnicas se pueden adaptar de manera adecuada para el uso con ADN tratado con bisulfito, y además para el análisis de metilación dentro de los dinucleótidos CpG.

En una modalidad preferida adicional del método, la *cuarta etapa* del método comprende el uso de extensión de oligonucleótido dirigido por plantilla, tal como MS-SNuPE como se describe en Gonzalzo & Jones, Nucleic Acids Res. 25:2529-2531, 1997.

En aún una modalidad adicional del método, la *cuarta etapa* del método comprende la secuenciación y análisis posterior de la secuencia del amplificado generado en la *tercera etapa* del método (Sanger F., y otros, Proc Natl Acad Sci USA 74:5463-5467, 1977).

Mejor modo

En la modalidad más preferida del método los ácidos nucleicos genómicos se aíslan y se tratan de acuerdo con las tres primeras etapas del método que se describió anteriormente, a saber:

- a) obtener, de un sujeto, una muestra biológica que tenga el ADN genómico del sujeto;
- b) extraer o de cualquier otra forma aislar el ADN genómico;
- c) tratar el ADN genómico de b), o un fragmento de este, con uno o más reactivos para convertir las bases de citosina que no son metiladas en la posición 5 de esta a uracilo o a otra base que se detecta diferente a la citosina en términos de propiedades de hibridación; y en donde
- d) amplificar después que el tratamiento en c) se lleva a cabo en una manera específica de metilación, específicamente mediante el uso de iniciadores específicos de metilación u oligonucleótidos de bloqueo, y adicionalmente en donde
- e) la detección de los amplificadores se lleva a cabo por medio de una sonda en tiempo real, como se describe anteriormente. Preferentemente, donde la amplificación posterior de d) se lleva a cabo por medio de iniciadores específicos de metilación, como se describió anteriormente, dichos iniciadores específicos de metilación comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácido nucleico tratada de acuerdo con una de una secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43 y las secuencias complementarias a estas, en donde la secuencia base de dichos oligómeros comprende al menos un dinucleótido CpG.

Etapa e) del método, específicamente la detección de los amplificadores específicos indicativos del estado de metilación de una o más posiciones de CpG de la SEQ ID NO:24 se lleva a cabo por medio de métodos de detección en tiempo real como se describió anteriormente.

5 Las modalidades adicionales de la invención proporcionan un método para el análisis del estado de metilación del ADN genómico de acuerdo con la invención (SEQ ID NO: 24 y complementos de la misma) sin la necesidad de la conversión de bisulfito. Los métodos se conocen en la técnica en donde un reactivo enzima de restricción sensible a la metilación, o una serie de reactivos enzimas de restricción que comprende reactivos enzimas de restricción sensibles a la metilación que distinguen entre dinucleótidos CpG metilados y no metilados dentro de una región objetivo se utilizan en la determinación de la metilación, por ejemplo, pero sin limitarse a DMH.

10 En la *primera etapa* de tales modalidades adicionales, la muestra de ADN genómico se aísla de tejido o de fuentes celulares. El ADN genómico se puede aislar mediante cualquier medio estándar en la técnica, lo que incluye el uso de estuches disponibles comercialmente. Brevemente, en donde el ADN de interés se encapsula mediante una membrana celular la muestra biológica se debe romper y lisar mediante medios enzimáticos, químicos o mecánicos. La solución de ADN se puede clarificar de las proteínas y otros contaminantes, por, *por ejemplo*, mediante digestión con proteinasa K. El ADN genómico se recupera después de la solución. Esto puede llevarse a cabo por medio de una variedad de métodos que incluyen precipitación salina, extracción orgánica o unión de ADN a un soporte de fase sólida. La elección del método se afectará por varios factores, que incluyen el tiempo, costo y cantidad requerida de ADN. Todos los tipos de muestra que comprenden materia neoplásica o potencialmente neoplásica son adecuados para usar en el presente método, se prefieren las líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, heces, efluente colónico, orina, plasma sanguíneo, suero sanguíneo, sangre completa, células de sangre aisladas, células aisladas de la sangre y combinaciones de estos. Los fluidos corporales son la fuente preferida de ADN; y particularmente se prefiere plasma sanguíneo, suero sanguíneo, sangre completa, células de sangre aisladas y células aisladas de la sangre.

25 Una vez que los ácidos nucleicos se extraen, el ADN genómico de cadena doble se utiliza en el análisis.

30 En una modalidad preferida, el ADN puede escindirse antes del tratamiento con enzimas de restricción sensibles a metilación. Tales métodos se conocen en la técnica y pueden incluir tanto medios físicos como enzimáticos. Particularmente se prefiere la utilización de una o una pluralidad de enzimas de restricción que no son sensibles a la metilación, y cuyos sitios de reconocimiento son ricos en AT y no comprenden dinucleótidos CG. La utilización de tales enzimas permite la conservación de las islas CpG y regiones ricas en CpG en el ADN fragmentado. Las enzimas de restricción que no son específicas de metilación se seleccionan preferentemente del grupo que consiste en MseI, BfaI, Csp6I, Tru1I, Tvu1I, Tru9I, Tvu9I, MaeI y XspI. Particularmente se prefiere la utilización de dos o tres de tales enzimas. Particularmente se prefiere la utilización de una combinación de MseI, BfaI y Csp6I.

40 El ADN fragmentado puede entonces ligarse a los oligonucleótidos adaptadores con el fin de facilitar la posterior amplificación enzimática. La ligación de oligonucleótidos para los extremos romos y cohesivos de fragmentos de ADN se conoce en la técnica, y se lleva a cabo por medio de la desfosforilación de los extremos (por ejemplo, mediante la utilización de la fosfatasa alcalina de camarones o terneros) y la posterior ligación mediante la utilización de enzimas ligasa (por ejemplo T4 DNA ligasa) en presencia de dATPs. Los oligonucleótidos adaptadores son típicamente de al menos 18 pares de bases de longitud.

45 En la tercera etapa, el ADN (o fragmentos del mismo) se digiere con una o más enzimas de restricción sensibles a la metilación. La digestión se lleva a cabo tal que la hidrólisis del ADN en el sitio de restricción es informativa del estado de metilación de un dinucleótido CpG específico del gen o secuencia genómica FOXL2.

50 Preferentemente, la enzima de restricción específica de metilación se selecciona del grupo que consiste en Bsi E1, Hga I, HinP1, Hpy99I, Ava I, Bce AI, Bsa HI, BsiI, BstUI, BshI236I, AccII, BstFNI, McrBC, Glal, Mvnl, HpaII (HapII), HhaI, Acil, SmaI, HinP1I, HpyCH4IV, EagI y mezclas de dos o más de las enzimas anteriores. Se prefiere una mezcla que contenga las enzimas de restricción BstUI, HpaII, HpyCH4IV y HinP1I.

55 En la cuarta etapa, que es opcional pero que es una modalidad preferida, se amplifican los fragmentos de restricción. Esto se lleva a cabo preferentemente mediante la utilización de una reacción en cadena de la polimerasa, y dichos amplificadores pueden llevar marcadores detectables adecuados como se discutió anteriormente, es decir marcadores fluoróforos, radionucleidos y marcadores de masa. Se prefiere particularmente la amplificación por medio de una enzima de amplificación y al menos dos iniciadores que comprenden, en cada caso una secuencia contigua de al menos 16 nucleótidos de longitud complementaria a, o que hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia que se selecciona del grupo que consiste en SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43, y complementos de

las mismas. Preferentemente dicha secuencia contigua es al menos de 16, 20 o 25 nucleótidos de longitud. En una modalidad alternativa dichos iniciadores pueden ser complementarios a los adaptadores ligados a los fragmentos.

En la quinta etapa se detectan los amplificados. La detección puede ser por cualquier medio estándar en la técnica, por ejemplo, pero sin limitarse a, análisis de electroforesis en gel, análisis de hibridación, incorporación de etiquetas detectables dentro de los productos de PCR, análisis de matriz de ADN, análisis MALDI o ESI. Preferentemente dicha detección se lleva a cabo por hibridación a al menos un ácido nucleico o ácido nucleico peptídico que comprende en cada caso una secuencia contigua de al menos 16 nucleótidos de longitud que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a la SEQ ID NO: 24, y complementos de la misma. Preferentemente dicha secuencia contigua es de al menos 16, 20 o 25 nucleótidos de longitud.

Posterior a la determinación del estado de metilación o nivel de los ácidos nucleicos genómicos la presencia, la ausencia o clase de cáncer colorrectal se deduce basado en el estado de metilación o nivel de al menos una secuencia de dinucleótidos CpG de SEQ ID NO:24, o un promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG de SEQ ID NO:24, en donde metilación se asocia con cáncer colorrectal. En donde dicha metilación se determina por medios cuantitativos el punto de corte para determinar dicha presencia de metilación es preferentemente cero (es decir, en donde una muestra exhibe cualquier grado de metilación se determina como que tiene un estado metilado en la posición CpG que se analiza). No obstante, se prevé que la persona experta en la técnica puede desear ajustar dicho valor de corte con el fin de proporcionar un ensayo de una especificidad o sensibilidad que se prefiera particularmente. En consecuencia dicho valor de corte puede aumentarse (aumentando por lo tanto la Especificación), dicho valor de corte puede estar dentro de un intervalo que se selecciona del grupo que consiste en 0%-5%, 5%-10%, 10%-15%, 15%-20%, 20%-30% y 30%-50%. Los cortes particularmente preferidos son 10%, 15%, 25%, y 30%.

Ensayos de Diagnóstico y Pronóstico de trastornos proliferativos celulares

La presente invención permite el diagnóstico de los eventos que son desfavorables para los pacientes o individuos en los que importantes parámetros genéticos y/o epigenéticos dentro del gen o secuencia genómica de FOXL2 pueden utilizarse como marcadores. Dichos parámetros que se obtienen por medio de la presente invención pueden compararse con otro conjunto de parámetros genéticos y/o epigenéticos, sirviendo las diferencias como base para un diagnóstico y/o pronóstico de eventos que son desventajosos para pacientes o individuos.

Más específicamente la presente invención permite el tamizaje de las poblaciones en riesgo para la detección temprana de carcinomas colorrectales. Los carcinomas colorrectales celulares neoplásicos presentan una menor metilación (es decir, una disminución de la expresión) en el gen o secuencia genómica de FOXL2, en oposición a dichos trastornos benignos que no la tienen de esa manera.

Específicamente, la presente invención proporciona ensayos de diagnóstico y clasificación de cáncer que se basan en la medición de la expresión diferencial (preferentemente metilación) de una o más secuencias de dinucleótidos CpG de la SEQ ID NO: 24 que comprenden tal secuencia de dinucleótido CpG. Típicamente, tales ensayos implican la obtención de una muestra de un sujeto, la realización de un ensayo para medir la expresión de al menos un gen o secuencia genómica que se selecciona del grupo que consiste en FOXL2 mediante la determinación del estado de metilación de la SEQ ID NO: 24 derivada de la muestra, en relación con una muestra de control, o un estándar que se conoce y haciendo un diagnóstico basado en ella.

Ejemplos

Ejemplo 1

En el siguiente ejemplo las secuencias que se enumeran a continuación se analizaron por medio de MSP y/o ensayo HeavyMethyl. Los ensayos se diseñaron para ejecutarse en la plataforma Lightcycler (Roche Diagnostics), pero otros instrumentos que se utilizan comúnmente en la técnica son también adecuados.

Los amplificados de MSP se detectan por medio de sondas de detección marcadas fluorescentes estilo Taqman, los amplificados HeavyMethyl se detectan por medio de sondas duales de estilo Lightcycler.

Región genómica de interés:

SEQ ID NO.: 24
 Tipo de ensayo: HeavyMethyl
 Iniciadores:

ES 2 533 767 T3

5 SEQ ID NO.: 254
SEQ ID NO.: 255
Bloqueadores:
SEQ ID NO.: 256
Sondas:
SEQ ID NO: 257 (fluo marcada)
SEQ ID NO: 258 (Red640 marcada)

10 Programa del ciclo térmico:
Desnaturalización a 95°C

95 °C 10min

15 55 ciclos:

Desnaturalización a 95°C	10 seg.	(20°C/s)
Hibridación 56°C	30 seg.	(20°C/s)
Extensión 72°C	10 seg.	(20°C/s)

20

25 Fusión

95 °C	10seg.	20
35 °C	20seg.	20
95 °C	0seg.	0,1

30

35 Región genómica de interés:
SEQ ID NO.: 24
Tipo de ensayo: HeavyMethyl
Iniciadores:
SEQ ID NO.: 264
SEQ ID NO.: 265

40 Bloqueadores:
SEQ ID NO.: 266
Sondas:
SEQ ID NO: 267 (fluo marcada)
SEQ ID NO: 268 (red640 marcada)

45 Programa del ciclo térmico:
Desnaturalización a 95 °C

50 95 °C 10min

55 55 ciclos.

Desnaturalización a 95°C	10 seg.	(20°C/s)
Hibridación 56°C	30 seg.	(20°C/s)
Extensión 72°C	10 seg.	(20°C/s)

5

Fusión

10	95 °C	10seg.	20
	35 °C	20seg.	20
	95 °C	0seg.	0,1

15

Ejemplo 2

20 El siguiente análisis se realizó con el fin de seleccionar los paneles que se prefieren adecuados para la detección y/o diagnóstico del carcinoma colorrectal basados en el análisis de metilación del ADN en sangre completa.

25 Se analizó el rendimiento de cada marcador mediante la utilización de una plataforma de ensayo (Lightcycler) y ensayos en tiempo real (MSP y/o HeavyMethyl) como sería adecuado para su uso en un entorno de laboratorio clínico o de referencia. El rendimiento de cada marcador se probó independientemente en tejido de carcinoma colorrectal y sangre completa, con el fin de proporcionar una indicación de la exactitud de cada marcador.

Los paneles se seleccionaron del grupo de marcadores que consiste en:

30 SEQ ID NO: 24

Cada marcador se analizó por medio de al menos un ensayo específico de metilación, a saber MSP y/o HeavyMethyl, como se muestra en la Tabla 2.

35 Un ensayo más (no específico de la metilación), que se refiere de aquí en adelante como el ensayo C3 se realizó con el fin de cuantificar la cantidad total de ADN en cada muestra. El ensayo C3 es un ensayo de ADN de bisulfito que detecta el ADN total independientemente del estado de metilación. Se utilizaron los siguientes iniciadores y sondas:

40 Iniciador: GGAGTGGAGGAAATTGAGAT SEQ ID NO: 62
 Iniciador: CCACACAACAAATACTCAAAAC SEQ ID NO: 63
 Sonda.: TGGGTGTTTGTAAATTTTTGTTTGTGTTAGGTT SEQ ID NO: 64

45 Cada ensayo se realizó por duplicado en el carcinoma colorrectal, tejido adyacente normal y/o muestras de sangre completa como se muestra en la Tabla 3.

La extracción de ADN se realizó mediante la utilización de estuches disponibles en el mercado, y la conversión de bisulfito se llevó a cabo con modificaciones menores de acuerdo con el método que se describió en Olek y otros (1996).

50 Todos los ensayos (C3 y específicos de metilación) se realizaron mediante la utilización de la plataforma Lightcycler.

Interpretación de datos

Cálculo de la concentración de ADN.

55 Los Cp (valores del punto de cruce) y las curvas de intensidad según los cálculos del software del instrumento Lightcycler se utilizaron para determinar la concentración de ADN. La concentración de ADN se calculó por referencia del valor de CP de cada pocillo a una curva estándar tanto para los ensayos metilación como para el ensayo C3.

60 Réplicas de la muestra

En la mayoría de los casos cada ensayo se ejecutó dos veces por muestra, lo que resultó en múltiples mediciones por muestra. Para cada muestra se calcula una puntuación como sigue:

1. Calcular la relación $v1/v2$ para todos los pares de muestras
2. Si ambos están por debajo de un umbral de 0.1 ng, la relación se establece en =, si uno es = y el otro está por encima del umbral, ajuste la relación a 100
3. Para cada muestra de ensayo cuya relación es superior a 2.5 no se analiza más
4. Para muestras que no tengan exactamente dos réplicas la media se toma sin tomar ninguna puntuación

10 Porcentaje de metilación

Todas las muestras que medían menos de 1 ng de ADN mediante la utilización del ensayo C3 no se consideraron más. Para cada muestra se calculó el porcentaje de metilación detectado como la concentración medida de ADN cuantificada mediante la utilización de los ensayos de metilación sobre la concentración de ADN en la muestra según se cuantificó por el ensayo C3.

La detección de metilación se determinó a tres niveles umbrales diferentes, ver las tablas), así como a todos los nivel de metilación (es decir, cualquier muestra en donde se detectó metilación se consideró positiva).

20 La sensibilidad de cada ensayo se determinó a partir de la tasa de detección positiva de muestras de carcinoma colorrectal, en donde la sensibilidad se determinó como % de muestras en donde se detectó positivamente metilación (es decir, positivos verdaderos).

25 La especificidad de cada ensayo se determinó a partir de la tasa de detección negativa de muestras de sangre completa (es decir, tasa de detección de negativos verdaderos) en donde los falsos positivos se descontaron a partir del número total de muestras que se analizaron.

Resultados

30 La proporción de las muestras que se analizaron con metilación medida dentro de varios umbrales por ensayos individuales se muestran en las Tablas 4 (tejidos de carcinoma colorrectal) y 5 (sangre completa).

35 La Figura 1 muestra la gráfica de distribución binaria (parte superior izquierda de la figura) y cuando procede la gráfica de distribución multiclase (lado inferior izquierdo de la figura) de la proporción (eje Y) de tejido de carcinoma colorrectal y sangre completa, (y en algunos casos de tejidos adyacentes normales) muestras con un nivel de metilación medido por encima de un punto de corte que se especificó (eje X). En el lado derecho de cada figura se encuentra un gráfico ROC de sensibilidad contra especificidad. La curva ROC es un gráfico de la tasa de verdaderos positivos contra la tasa de falsos positivos para los diferentes puntos de corte posibles de una prueba de diagnóstico. Se muestra la compensación entre la sensibilidad y la especificidad dependiendo del punto de corte que se seleccionó (cualquier aumento en la sensibilidad se acompañará por una disminución en la especificidad). El área bajo una curva ROC (AUC) es una medida de la exactitud de una prueba de diagnóstico (cuanto mayor sea el área mejor, óptimo es 1, una prueba puede tener una curva ROC que yace en la diagonal con un área de 0.5; para referencia: J.P. Egan. Signal Detection Theory and ROC Analysis, Academic Press, Nueva York, 1975). El AUC de cada gráfico ROC y el valor p de Wilcoxon se muestran en la Tabla 7.

45 Etapa

Un análisis más detallado de los resultados del carcinoma colorrectal de acuerdo con la etapa del carcinoma se muestra en la Tabla 6. En dicha tabla se muestra para todas las etapas del CRC la sensibilidad del marcador que se basa en dos umbrales de metilación diferentes (>10% y > 20%). Para la mayoría de los marcadores, la sensibilidad es uniforme a través de todas las etapas del CRC por lo que estos marcadores serían adecuados para la detección de todas las etapas de CRC en una prueba de detección o monitoreo. Parece existir una tendencia por una mayor sensibilidad en los cánceres de la Etapa II.

55 Ejemplo 3

Se analizó el rendimiento del marcador mediante la utilización de una plataforma de ensayo (Lightcycler) y de ensayos en tiempo real (MSP y/o HeavyMethyl) como sería adecuado para su uso en un entorno de laboratorio clínico o de referencia. El rendimiento de cada marcador se probó independientemente en el tejido colorrectal (tejido adyacente normal), el tejido de carcinoma colorrectal y sangre completa, con el fin de proporcionar una indicación de la exactitud del marcador.

El conjunto de muestras 1 se analizó mediante la utilización de los siguientes ensayos que se detallan en la Tabla 2:

SEQ ID NO: 24 (ensayo 5)

5

El conjunto de muestras 2 se analizó mediante la utilización de los siguientes ensayos como se detallan en la Tabla 2:

SEQ ID NO: 24 (Ensayo 5b)

10

como se detalla en la Tabla 10.

Extracción de ADN y tratamiento con bisulfito

15

El ADN se aisló de todas las muestras por medio del método Magna Pure (Roche) de acuerdo con las instrucciones del fabricante. El eluato resultante de la purificación se convirtió a continuación de acuerdo con la siguiente reacción de bisulfito.

20

El eluato se mezcló con 354 µl de disolución de bisulfito (5.89 mol/l) y 146 µl de dioxano que contiene un secuestrante de radicales (ácido 6-hidroxi-2, 5,7,8-tetrametilcromano 2-carboxílico, 98.6 mg en 2.5 ml de dioxano). La mezcla de reacción se desnaturizó durante 3 min a 99 °C y posteriormente se incubó al siguiente programa de temperatura por un total de 7 h min 50 °C; un pico térmico (99.9 °C) durante 3 min; 1.5 h 50 °C; un pico térmico (99 °C) durante 3 min; 3 h 50 °C. La mezcla de reacción se purificó posteriormente por ultrafiltración mediante la utilización de una columna Millipore Microcon™. La purificación se llevó a cabo esencialmente de acuerdo con las instrucciones del fabricante. Con este propósito, la mezcla de reacción se mezcló con 300 µl de agua, se cargó en la membrana de ultrafiltración, se centrifugó durante 15 min y posteriormente se lavó con tampón TE 1x. El ADN permanece en la membrana en este tratamiento. Después se lleva a cabo la desulfonación. Con este propósito, se añadieron 0.2 mol/l de NaOH y se incubó durante 10 min. Después se llevó a cabo una centrifugación (10 min), seguida por una etapa de lavado con tampón TE 1 x. Después de esto, el ADN se eluyó. Con este propósito, la membrana se mezcló durante 10 minutos con 75 µl de tampón TE 1 x caliente (50 °C). La membrana se volcó de acuerdo con las instrucciones del fabricante. Posteriormente se realizó una centrifugación repetida, con la cual el ADN se removió de la membrana. 10 µl de eluato se utilizaron para el ensayo de PCR Lightcycler en tiempo real.

30

Disoluciones de reacción y condiciones de los ciclos térmicos

SEQ ID NO: 24 ensayo 5B (ensayo HeavyMethyl)

35

Solución de reacción:

Agua

40

MgCl2	3,00	MM *
Iniciador directo	0,30	µM
Iniciador inverso	0,30	µM
bloqueador	4,00	µM
detect. sondas fluo	0,15	µM
detect. sondas rojas	0,15	µM
1a+1b mezcla de reactivos	1,00	x

50

Condiciones de ciclo térmico:

55

desnat a 95 °C

60

ES 2 533 767 T3

5	95 °C	10min		
	55 ciclos:			
	desnat a 95°C	10 seg.	(20°C/s)	
	Hibridación 58°C	30 seg.	(20°C/s)	detección
	Extensión 72°C	10 seg.	(20°C/s)	
10	Fusión			
	95 °C	10seg.	20	
	35 °C	20seg.	20	
15	95 °C	0seg.	0,1	

20 SEQ ID NO: 24 ensayo 5 (ensayo HeavyMethyl)

Solución de reacción:

25	agua			
	MgCl2	3,00	mM (amortiguador incluye mM)	
	Iniciador directo	0,30	µM	
30	Iniciador inverso	0,30	µM	
	bloqueador	4,00	µM	
	Sonda Lightcycler	0,15	µM	
	Sonda Lightcycler	0,15	µM	
35	1a+1b mezcla de reactivos	1,00	x	

40 Condiciones del ciclo térmico:

desnat a 95 °C

45	95 °C	10min		
	55 ciclos:			
	desnat a 95°C	10 seg.	(20°C/s)	
	Hibridación 58°C	30 seg.	(20°C/s)	detección
50	Extensión 72°C	10 seg.	(20°C/s)	
	Fusión			
	95 °C	10seg.	20	
	35 °C	20seg.	20	
55	95 °C	0seg.	0,1	

60

Interpretación de datos

5 Cálculo de la concentración de ADN.

Los Cp (valores de puntos de cruce) que se calcularon por el programa del instrumento Lightcycler se utilizaron para determinar la concentración de ADN. La concentración de ADN se calculó por referencia del valor CP de cada pocillo a una curva estándar tanto para los ensayos de metilación como para el ensayo C3.

10 En la mayoría de los casos cada ensayo se ejecutó dos veces por muestra, lo que resulta en múltiples mediciones por muestra.

Porcentaje de metilación

15 Todas las muestras que medían menos de 4 ng de ADN mediante la utilización del ensayo C3 no se consideraron más. Para cada muestra se calculó el porcentaje de metilación detectado como la concentración que se midió de ADN cuantificada mediante la utilización los ensayos de metilación sobre la concentración de ADN en la muestra según lo que se cuantificó por el ensayo C3.

20 La detección de metilación se determinó en múltiples niveles de umbrales diferentes, ver las tablas), así como también a todos los niveles de metilación (es decir, cualquier muestra en donde se detectó metilación se consideró positiva).

25 La sensibilidad de cada ensayo se determinó a partir de la tasa de detección positiva de las muestras de carcinoma colorrectal, en donde la sensibilidad se determinó como el % de muestras den donde la metilación se detectó positivamente (es decir, positivos verdaderos).

30 La especificidad de cada ensayo se determinó a partir de la tasa de detección negativa de las muestras de sangre completa (es decir, tasa de detección de verdaderos negativos) en donde los falsos positivos se descontaron a partir del número total de muestras que se analizaron.

Tabla 1: Secuencia genómicas de acuerdo con la lista de secuencias

SEQ ID NO:	Ubicación en la base de datos* Ensembl	Ubicación genómica en la base de datos* Ensembl	Transcrito(s)* asociado(s)	Secuencia convertida por bisulfito metilada (sentido)	Secuencia convertida por bisulfito metilada (antisentido)	Secuencia convertida por bisulfito no metilada (sentido)	Secuencia convertida por bisulfito metilada (antisentido)
24	AC092947.12.1.72 207 58709 a 24 60723 (+)	3 140138862 a 140140876 (+)	FOXL2	30	31	42	43

*Base de datos Ensembl

Tabla 2

5	SEQ ID NO genómica:	Ensayo	iniciador	iniciador	bloqueador	Sonda	Sonda
10	SEQ ID NO: 24 (ensayo 5)	HM	Ccaaaacctaa acttacaac (SEQ ID NO: 365)	Ggaaattgag gggtaa (SEQ ID NO: 366)	Tacaacaccac caacaaaccca aaaacacaa (SEQ ID NO: 367)	GTtAATTGC GGGCGAtC GA (SEQ ID NO: 368)	CGtCGtAG CGGGTGG G (SEQ ID NO: 369)

Tabla 3: Secuencias analizadas de acuerdo con el Ejemplo 2

Ensayo	Total de no	Carcinoma	Tejido	Sangre
	muestras	colorrectal	adyacente normal	
SEQ ID NO: 24 (ensayo 5)	106	79	0	27

Tabla 4: Proporción de las muestras de carcinoma colorrectal con metilación dentro de varios umbrales

Ensayo	por encima de 0.01	por encima de 0.1	por encima de 0.3	por encima de 0.5
SEQ ID NO: 24 (ensayo 5)	0,911	0,557	0,152	0,076

Tabla 5: Proporción de muestras de sangre completa con metilación dentro de varios umbrales

Ensayo	por encima de 0.0001	por encima de 0.001	por encima de 0.011	por encima de 0.1
SEQ ID NO: 24 (ensayo 5)	0,074	0	0	0

Tabla 6: Proporción de muestras de carcinoma colorrectal dentro de varios umbrales de metilación de acuerdo con la etapa de la enfermedad

Ensayo	Etapa I		Etapa II		Etapa III		Etapa IV	
	>10%	>20%	>10%	>20%	>10%	>20%	>10%	>20%
SEQ ID NO: 24 (ensayo 5 HM)	38.5	23.1	90	60	53.8	23.1	57.1	42.9

Tabla 7: Diferenciación entre muestras de sangre y de carcinoma colorrectal como se ilustra en la Figura 1.*

Figura	Ensayo	AUC de ROC	sensibilidad/especificidad	valor P de Wilcoxon
1	SEQ ID NO: 24 (ensayo 5)	0.98 (0.93,1)	0.96/0.93	0
* se muestran los intervalos de confianza en corchetes				

Tabla 8: Conjunto de muestras 1 de acuerdo con el Ejemplo 3.

Tipo de muestra	Sexo	Edad	Etapa	T	N	M	Localización
CRC	F	39	III	4	1	0	sigmoide
CRC	F	65	III	3	2	0	íleo-ciego
CRC	M	58	IV				recto
CRC	M	63	III	3	1	0	recto
CRC	M	71	II				ascendente
CRC	F	69	I	2	0	0	ciego
CRC	F	54	III	3	2	0	ciego
CRC	M	44	IV				
CRC	F	75	IV				transverso
CRC	F	60	II				recto
CRC	M	76	I				descendente
CRC	M	69	IV				sigmoide
CRC	M	73	I	1	0	0	recto
CRC	M		II	3	0	0	ascendente
CRC	M	62	III	3	1		
CRC	F	49	IV				ascendente
CRC	F	58	III	3	1	X	ascendente
CRC	M	42	IV	3	0	1	
CRC	M	64	I	2	0	0	sigmoide
CRC	F	64	III				recto

ES 2 533 767 T3

5

10

15

20

25

30

35

40

45

50

55

CRC	F	70	III	3	1	0	íleon terminal
CRC	M	67					
CRC	M	80	III	3	1	0	rectosigmoide
CRC	F	72	IV				sigmoide
CRC	M		III				recto
CRC	M	56	I	2	0	0	sigmoide
CRC	M	72	III	2	1	0	recto
CRC	M	45	IV	4	2	1	ciego
CRC	F		II	3	0	0	
CRC	M	74	III	3	1	0	rectosigmoide
CRC	F	75	III	4	2	0	pared ciega
CRC	M		III	3	1	0	
CRC	M		I	2	0	0	ascendente
CRC	F	74	I	2	0	0	ciego
CRC	M	62	I	2	0	0	rectosigmoide
CRC	F	60	II	3	0	0	recto
CRC	F	80	II				ascendente
CRC	F	70	III	4	2	0	recto
CRC	M		III	3	1	0	
CRC	F	75	III	3	1	0	ascendente
CRC	F	49	IV	4	X	1	recto
CRC	F	47	I				ano
CRC	M	81	IV			1	
CRC	F	89	III	3	1	0	recto
CRC	M	85	III	3	1	0	ciego
CRC	M	52	III	2	1	0	
CRC	M	75	II				sigmoide
CRC	M						
CRC	F	71					
CRC	M		III				recto
CRC	M	61		3	x	0	descendente
CRC	F	56	unk				sigmoide
CRC	F	68	IV	3	2	1	sigmoide
CRC	F	65	III	3	2	0	íleo-ciego

60

ES 2 533 767 T3

5

10

15

20

25

30

35

40

45

50

55

CRC	M	88	II	3	0	0	flexura
CRC	F	72	III				ciego
CRC	M	61	IV	3	2	1	recto
CRC	M		III	3	2		
CRC	M	52	II	3	0	0	transverso
CRC	M	66	IV	2	0	1	recto
CRC	M	64	III				ascendente
CRC	F	65	II	3	0	0	
CRC	M	61	IV	3	2	1	sigmoide
CRC	M	64	III	3	1	0	ascendente
CRC	M	76	0	0			sigmoide
CRC	M	64	I	2	0	0	ascendente
CRC	M	56	I	2	0	0	transverso
CRC	F	67	II	3	0	0	sigmoide
CRC	M		II	3	0	0	ascendente
CRC	M	66	III	4	1	0	
CRC	M		II	3	0	0	
CRC	F		III				
CRC	F	65	I	2	0	X	recto
CRC	M		II	3	0	0	
CRC	M	40	I				FAP
CRC	M	77	I	2	0	0	rectosigmoide
CRC	M	65	III	4	2	0	descendente
CRC	M	68	IV				sigmoide
CRC	M	67	II				recto
CRC	M		unk				recto
CRC	F	63		3	X	0	
CRC	M	68	unk				descendente
CRC	F	53	III	3	1	0	ascendente
CRC	M		II	3	0	0	
CRC	M	68	I	2	0	0	recto
CRC	M	84	III				recto
CRC	F	53	I	1	0	0	descendente
CRC	M	72	III	4	1	0	

60

ES 2 533 767 T3

5	CRC	F	69	I	1	0	0	sigmoide
	CRC	M		II	3	0	0	descendente
	CRC	M		II	3	0	0	ciego
10	Sangre normal	F	62	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	62	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	44	n.a.	n.a.	n.a.	n.a.	n.a.
15	Sangre normal	F	57	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	51	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	66	n.a.	n.a.	n.a.	n.a.	n.a.
20	Sangre normal	M	65	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	55	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	70	n.a.	n.a.	n.a.	n.a.	n.a.
25	Sangre normal	M	40	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	42	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	68	n.a.	n.a.	n.a.	n.a.	n.a.
30	Sangre normal	F	67	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	53	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F		n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	50	n.a.	n.a.	n.a.	n.a.	n.a.
35	Sangre normal	M	50	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	51	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	56	n.a.	n.a.	n.a.	n.a.	n.a.
40	Sangre normal	M	58	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	67	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	55	n.a.	n.a.	n.a.	n.a.	n.a.
45	Sangre normal	M	62	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	M	66	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	56	n.a.	n.a.	n.a.	n.a.	n.a.
50	Sangre normal	M	56	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre normal	F	69	n.a.	n.a.	n.a.	n.a.	n.a.

55

60

5 Tabla 9 Conjunto de muestras 2 de acuerdo con el Ejemplo 3.

Tipo de muestra	Sexo	Edad	Etapa	T	N	M	Localización
CRC	F	49	IV				ascendente
CRC	F	72	IV				sigmoide
CRC	M	69	IV				sigmoide
CRC	F	58	III	3	1	X	ascendente
CRC	F	60	II				recto
CRC	F	74	I	2	0	0	ciego
CRC	F	70	III	3	1	0	íleon terminal
CRC	F	69	I	2	0	0	ciego
CRC	F	39	III	4	1	0	sigmoide
CRC	M	56	I	2	0	0	sigmoide
CRC	F		II	3	0	0	
CRC	M	64	I	2	0	0	sigmoide
CRC	M	45	IV	4	2	1	ciego
CRC	F	54	III	3	2	0	ciego
CRC	M	42	IV	3	0	1	
CRC	M	73	I	1	0	0	recto
CRC	M	62	III	3	1		
CRC	M		I	2	0	0	ascendente
CRC	F	75	III	3	1	0	ascendente
CRC	M	74	III	3	1	0	rectosigmoide
CRC	F	68	IV	3	2	1	sigmoide
CRC	F	75	IV				transverso
CRC	M	85	III	3	1	0	ciego
CRC	M	80	III	3	1	0	rectosigmoide
CRC	M	66	III	4	1	0	
CRC	F	70	III	4	2	0	recto
CRC	F	89	III	3	1	0	recto

ES 2 533 767 T3

5

10

15

20

25

30

35

40

45

50

55

60

CRC	M	67					
CRC	F	67	II	3	0	0	sigmoide
CRC	M	66	IV	2	0	1	recto
CRC	F	56	unk				sigmoide
CRC	M	72	III	2	1	0	recto
CRC	F	80	II				ascendente
CRC	M	75	II				sigmoide
CRC	F	49	IV	4	X	1	recto
CRC	M		III				recto
CRC	F	60	II	3	0	0	recto
CRC	M	62	I	2	0	0	rectosigmoide
CRC	M	88	II	3	0	0	flexura
CRC	M	61	IV	3	2	1	sigmoide
CRC	M	61		3	X	0	descendente
CRC	F	64	III				recto
CRC	M		III				recto
CRC	M	52	II	3	0	0	transverso
CRC	F	71					
CRC	M	81	IV			1	
CRC	F	65	III	3	2	0	íleo-ciego
CRC	M						
CRC	F	65	II	3	0	0	
CRC	F	72	III				ciego
CRC	M	61	IV	3	2	1	recto
CRC	M	52	III	2	1	0	
CRC	M		II	3	0	0	
CRC	F	47	I				ano
CRC	M		II	3	0	0	ascendente
CRC	M	64	III	3	1	0	ascendente
CRC	M	64	I	2	0	0	ascendente
CRC	M	76	0	0			sigmoide
CRC	M	56	I	2	0	0	transverso
CRC	M	65	III	4	2	0	descendente
CRC	M	40	I				FAP
CRC	F	53	I	1	0	0	descendente
CRC	M		II	3	0	0	

ES 2 533 767 T3

5	CRC	M		III	3	2		
	CRC	M		unk				recto
	CRC	M	68	I	2	0	0	recto
10	CRC	F	63		3	X	0	
	CRC	F		III				
	CRC	M	67	II				recto
15	CRC	F	65	I	2	0	X	recto
	CRC	M	64	III				ascendente
	CRC	M	68	IV				sigmoide
20	CRC	M		II	3	0	0	
	CRC	M	72	III	4	1	0	
	CRC	M	77	I	2	0	0	rectosigmoide
25	CRC	F	53	III	3	1	0	ascendente
	CRC	F	69	I	1	0	0	sigmoide
	CRC	M	84	III				recto
30	CRC	M		II	3	0	0	descendente
	CRC	M	68	unk				descendente
	CRC	M		II	3	0	0	ciego
35	Sangre Normal	M	55	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	M	62	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	F	57	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	F	62	n.a.	n.a.	n.a.	n.a.	n.a.
40	Sangre Normal	M	65	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	F		n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	F	44	n.a.	n.a.	n.a.	n.a.	n.a.
45	Sangre Normal	F	68	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	F	70	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	M	58	n.a.	n.a.	n.a.	n.a.	n.a.
50	Sangre Normal	M	62	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	F	53	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	F	42	n.a.	n.a.	n.a.	n.a.	n.a.
55	Sangre Normal	F	51	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	M	66	n.a.	n.a.	n.a.	n.a.	n.a.
	Sangre Normal	M	51	n.a.	n.a.	n.a.	n.a.	n.a.

60

5

10

15

20

Sangre Normal	M	40	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	M	56	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	F	56	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	F	50	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	M	50	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	F	67	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	M	67	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	M	55	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	M	66	n.a.	n.a.	n.a.	n.a.	n.a.
Sangre Normal	M	56	n.a.	n.a.	n.a.	n.a.	n.a.

25

Tabla 10: Ensayos de acuerdo con el Ejemplo 3

30

35

40

SEQ ID NO genómica:	Ensayo	Iniciador	Iniciador	Bloqueador	Sonda	Sonda
SEQ ID NO: 24 (ensayo 5)	HM	<u>ccaaaacctaa</u> <u>acttacaac</u> (SEQ ID NO: <u>102</u>)	<u>lctaaataacaa</u> <u>aatacctccatt</u> (SEQ ID NO: <u>110</u>)	<u>Tacaacaccac</u> <u>caacaaccca</u> <u>aaaacacaa</u> (SEQ ID NO: <u>104</u>)	<u>GTtAATTGC</u> <u>GGGCGAtC</u> GA (SEQ ID NO: <u>105</u>)	<u>CGtCGtAGC</u> <u>GGGTGGG</u> (SEQ ID NO: <u>106</u>)

45

LISTADO DE SECUENCIAS

<110> Epigenomics AG

5 <120> MÉTODOS PARA EL ANÁLISIS DE TRASTORNOS PROLIFERATIVOS CELULARES

<130> 536-22 EPT1

<140> EP 08 021 839.9

10 <141> 2006-04-17

<150> US 60/672,242

<151> 2005-04-15

15 <150> US 60/676,997

<151> 2005-05-02

<150> US 60/697,521

<151> 2005-07-08

20

<150> US 60/704,860

<151> 2005-08-01

<150> US 60/709,318

25 <151> 2005-08-17

<150> US 60/723,602

<151> 2005-10-04

30 <150> US 60/787,402

<151> 2006-03-30

<160> 369

35 <170> PatentIn versión 3.5

<210> 1

<400> 1

40 000

<210> 2

<400> 2

45 000

<210> 3

<400> 3

50 000

<210> 4

<400> 4

55 000

<210> 5

<400> 5

60 000

<210> 6
5 <400> 6
000
<210> 7
10 <400> 7
000
<210> 8
15 <400> 8
000
<210> 9
20 <400> 9
000
<210> 10
25 <400> 10
000
<210> 11
30 <400> 11
000
<210> 12
35 <400> 12
000
<210> 13
40 <400> 13
000
<210> 14
45 <400> 14
000
<210> 15
50 <400> 15
000
<210> 16
55 <400> 16
000
<210> 17
60 <400> 17
000

<210> 18
<400> 18
000
5 <210> 19
<400> 19
000
10 <210> 20
<400> 20
000
15 <210> 21
<400> 21
000
20 <210> 22
<400> 22
000
25 <210> 23
<400> 23
000
30 <210> 24
<211> 2804
<212> ADN.
<213> Homo sapiens
35 <400> 24

40

45

50

55

60

ES 2 533 767 T3

ccaaggcccc cgggacctga ggtcgcctgg ggcgctaaac agacccatct atgagcacat 60
 5 tcggccaaca cacccaagtt caacagtggg cgggatcttt ggctggcctc gatctctctc 120
 acagtgtgtg caaacgggta cccattagat tctttttgct tgtttcctca tctgtaaact 180
 10 ttgaaaggag agtatttccc ggggtgattc tgaaaatcat gccagaactg aaaaggcctg 240
 tgtaaatgcg aggtgtaatg atgacttatt agaagcagcc agccctgtac gtgcctgtag 300
 gatagatgct agaacgttct acaaatgtgt gcacatacac atcacgtatg tgtgaagggtg 360
 15 cacacctgtc cctagacact tgacaaatgc acacacgtat gcgtgtacct ctggctgaga 420
 gcaaaggggc taatacaggc ctatacagg acatcatcct gcagtcttgt gcgtgcacag 480
 20 tacacacgcg agcacgactg tctgtgcctg gagacaccgt atggcaccgg aggagctgct 540
 cagccgtcgg tgccacctgg gaagtagggt ttcgggtgct gctctggagc gcgggaaagt 600
 ccggatccgg gtctgctggt cagcgcctcg gcgccgccgg gctcctttat ctctcatccg 660
 25 ccggccgggtg agcggctcca gtttcaagac ccccgatcct gccggtgagg ggaggagcgg 720
 agaaggagtg cgctccgggc ccaaggaggc gcagggcccc actccccagg tctgtcgtcc 780
 30 ttcctttatc tcctcaggtc acacccccac ttaaataaa gaagggtgcc tgactcattc 840
 aggtcagggg agtcggcccc cgggagcctc ctgcctcac aagccatggt gaagggaggc 900
 35 gaagccaggg tttgccacgg cctgggagaa aatgcggctg cagggctctt ttgggctgca 960
 gcgctggccc ggggccccaa gactgttaag gtgtgtgtgg gaggcgcgca gtgtggctac 1020
 40 cgaagagccc tccatcaccc caccggcagg ccgccgggct cgtcctctc tcatgcttcc 1080
 aagggttctg agctgcgcag cactcgcac acccagaagt gtctgtggag aaaacgacct 1140
 caggtgtaag cccaaggctg ctgcctgcaa aggcaatgcc gtggagactg ggttccacag 1200
 45 cgacctgggt ttttaagtca ccgaaagcta cagggcaggg tcacaaacac tctcccgcct 1260
 tctctgcaga accgtgcggc tgacaggaga gcgttgcgca gaaaattcga ggccgggcgc 1320
 50 tggaggagtc tccgcccggc ggagaaggca cagaggcgcc cctgagaggc gcagctggaa 1380
 caggcgatgc acgggttcgg atctggccgg ccaaccgagc ccagcggctg cgaaaaccag 1440
 agtcgccaca gaggttgagc acgattccat ttctggggat cgggtccggt ggggagccac 1500
 55 tgtccctaac gcctccagag actttaagta ggacaatata atgctggtag gagcaagggtg 1560

60

ES 2 533 767 T3

caaggaattt agcggcagaa gcctttcggg ggggtgggga aaggcagatc cgtgcctcgt 1620
 ctgagcctgg ggggtggggac agccccctggc cttgtagccc ctgttccggg gatcagctgg 1680
 5 gcccaccta cccccctaca tggggttgaa aggggccaat gggacggccc cgccgcccct 1740
 ccctcgggat tccacagggc cctgaccggg cctttatctc tgcacgggtca gcagtcgcgg 1800
 10 gggcttcagg aaggaggaca aaggcccggg gtcaccgagg gcggggggctc ggtttcctgg 1860
 ctctctgctc acactcacag cctttagcgg ttgttggggg aagattttta aaaatatgtg 1920
 tcgaatttcc ttttttctc ttctagaaac aaacaacaa aaaaggcaaa aggcgaattc 1980
 15 cccttcactc ctcgtccata gagattaaag tttcctggga tcctgccctt ttttttctt 2040
 tgactgccta gaaatacttg ttctcccttg tacatagagg aaaatgcggg gaaaggtttt 2100
 20 ttaaaactcg gtttcatact attattatta ctaaggacaa ccgggcaggc tgagggtcca 2160
 acgtggatga tccgagttgg cctcgcggcg gggctctgca gccactgccc tgtgcgctca 2220
 25 gcacctctgg gggcgatcag ggccccctgc cttccgcccg ccgcccggca gtcgagagca 2280
 ccctgtgccc agactggccg actcattctc ccccgaattt tgtttagagc tggcaagggg 2340
 30 gacttagctc gcgccccaaag acctgggctt gcagcgccgc caacaggccc ggggacacga 2400
 ggcgctccag gccggggtct tcccggctgc tggccccctc cgtccccac ccgctggcgg 2460
 cgctcggtc gcccgcaatt gaccacaacc gcttctctgc tttgccccctc aggtttcccg 2520
 35 tttctccaca aaggcctagg ggagcctcgc ccacaggctg accctgcaac ctctggcccc 2580
 gtggctacct tctgcttttc tgaaaaagaa aaggaaaaaa aaaaaaaaaa agaaaaaatc 2640
 40 aaccagtcga gggaggcgcg tgaggactgg aggcgccagc cggacagcct atgagtaggt 2700
 ccctgggtcg gtgccgcttc gcgggtcagc acggcctttc tcagagaaaa cctcctaaac 2760
 45 gtgtgaagac cgctttgggg gaagcgagag ggaggttggg ggag 2804

50
 55 <210> 25
 <400> 25
 000
 60 <210> 26

<400> 26
 000
 5 <210> 27
 <400> 27
 000
 10 <210> 28
 <400> 28
 000
 15 <210> 29
 <400> 29
 000
 20 <210> 30
 <211> 2804
 <212> ADN.
 <213> Secuencia artificial
 25 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 <400> 30
 30 ttaaggttcg tccgatttga ggtcgtttgg ggcgttaaata agatttattt atgagtatat 60
 tcggttaata tatttaagtt taatagtggg cgggattttt ggttggtttc gatttttttt 120
 atagtgtgtg taaacgggta tttattagat tttttttggt tgttttttta tttgtaaatt 180
 35 ttgaaaggag agtatttttc ggggtgattt tgaaaattat gttagaattg aaaaggtttg 240
 tgtaaatgcg aggtgtaatg atgatttatt agaagtagtt agttttgtac gtgtttgtag 300
 40 gatagatggt agaacgtttt ataaatgtgt gtatatatat attacgtatg tgtgaagggtg 360
 tatatttggt tttagatatt tgataaatgt atatacgtat gcgtgtattt ttggttgaga 420
 45 gtaaaggggt taatataggg tttatatagg atattatttt gtagttttgt gcgtgtatag 480
 tatatacgcg agtacgattg tttgtgtttg gagatatcgt atggatcgg aggagttggt 540
 tagtcgtcgg tgttatttgg gaagtagggt ttcgggtggt gttttggagc gcgggaaagt 600
 50 tcggattcgg gtttgttggt tagcgtttcg gcgtcgtcgg gtttttttat tttttattcg 660
 tcggtcgggt agcggtttta gttttaagat tttcgatttt gtcggtgagg ggagggagcg 720
 55 agaaggagtg cgtttcgggt ttaagagggt gtagggtttt attttttagg tttgtcgttt 780
 60

tttttttatt ttttaggtt atatttttat ttaaataaa gaagggtgtt tgatttattt 840
 5 aggttaggga agtcggtttt cgggagtttt ttgtttttat aagttatggt gaagggaggc 900
 gaagttaggg tttgttacgg tttgggagaa aatgcggttg tagggttttt ttgggttgta 960
 10 gcgttggttc ggggttttaa gattgttaag gtgtgtgtgg gaggcgcgta gtgtggttat 1020
 cgaagagttt tttattatcg tatcggtagg tcgtcgggtt cgtttttttt ttatgttttt 1080
 aagggttttg agttgcgtag tattcgtatt atttagaagt gtttgtggag aaaacgattt 1140
 15 taggtgtaag ttttaaggttg ttgtttgtaa aggtaatgtc gtggagattg ggttttatag 1200
 cgatttgggt ttttaagtta tcgaaagtta taggtaggg ttataaatat tttttcgttt 1260
 20 tttttgtaga atcgtgcggt tgataggaga gcgttcgta gaaaattcga ggtcggcgt 1320
 tggaggagt ttcgcggttc ggagaaggta tagaggcgtt tttgagaggc gtagttggaa 1380
 taggcgatgt acgggttcgg atttggtcgg ttaatcgagt ttagcggtcg cgaaaattag 1440
 25 agtcgttata gaggttgagt acgattttat ttttgggat cgggttcggt ggggagttat 1500
 tgtttttaac gtttttagag attttaagta ggataatata atgttgtag gagtaagggtg 1560
 30 taaggaattt agcggtagaa gtttttcggg ggggtgggga aaggtagatt cgtgtttcgt 1620
 ttgagtttg ggggtgggat agtttttgg tttgtagttt ttgtttcggg gattagtgg 1680
 gttatttaa ttttttata tggggttgaa aggggttaat gggacggtt cgtcgttttt 1740
 35 ttttcgggat tttatagggt tttgattcgg tttttattt tgtacggtta gtagtcgagg 1800
 gggtttagg aaggaggata aaggttcgg gttatcggg gcgggggttc ggttttttg 1860
 40 tttttgttt atatttatag ttttagcgg ttgttggggg aagattttta aaaatatgtg 1920
 tcgaattttt ttttttttt ttttagaaat aaataataa aaaaggtaaa aggcaattt 1980
 45 tttttattt ttcgtttata gagattaaag ttttttggga ttttgtttt ttttttttt 2040
 tgattgtta gaaatattg ttttttttg tatatagagg aaaatgcggg gaaaggttt 2100
 50 ttaaattcg gttttatatt attattatta ttaaggataa tcggtaggt tgaggttta 2160
 acgtggatga ttcgagttg tttcgcgctg gggttttgta gttattgttt tgtcgttta 2220
 gtatttttg gggcgattag ggtttttcg ttttcgttc tcgttcggtta gtcgagagta 2280
 55 ttttgtttt agattggtc atttatttt tttcgaattt tgttagagt tggttaaggg 2340

60

ES 2 533 767 T3

gatttagttc gcgttttaag atttgggttt gtagcgtcgt taataggttc ggggatacga 2400
 5 ggcgtttttag gtcgggggttt ttcggttgt tggttttttt cgttttttat tcgttggcgg 2460
 cgtttcggtc gttcgttaatt gatttaattc gttttttgcg tttgtttttt aggttttttcg 2520
 10 tttttttata aaggtttagg ggagtttcgt ttataggttg attttgtaat ttttggttcg 2580
 gtggttattt tttgtttttt tgaaaaagaa aaggaaaaaa aaaaaaaaaa agaaaaaatt 2640
 aattagtcga gggaggcgcg tgaggattgg aggcgttagt cggatagttt atgagtaggt 2700
 15 ttttgggtcg gtgtcgtttc gcgggttagt acggtttttt ttagagaaaa ttttttaaac 2760
 gtgtgaagat cgttttgggg gaagcgagag ggaggttggg ggag 2804

20

<210> 31
 <211> 2804
 <212> ADN.
 25 <213> Secuencia artificial

<220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

30

<400> 31

ttttttaat tttttttcg ttttttttaa agcggttttt atacgttttag gaggtttttt 60
 ttgagaaagg tcgtgtgat tcgcgaagcg gtatcgattt agggatttat ttataggttg 120
 35 ttcggttggc gtttttagtt tttacgcgtt tttttcgatt ggttgatttt tttttttttt 180
 tttttttttt tttttttttt tttagaaaag tagaaggtag ttatcgggtt agaggttgta 240
 40 gggttagttt gtggcgagg ttttttagg tttttgtgga gaaacgggaa atttgagggg 300
 taaacgtagg aagcgggttg ggtaattgc gggcgatcga ggcgtcgtta gcgggtgggg 360
 45 agcgagaggg gttagtagtc gggaagattt cggtttgagc cgtttcgtgt tttcgggttt 420
 gttggcggcg ttgtaagttt aggttttggg gcgcgagtta agtttttttt gttagtttta 480
 aataaaattc gggggagaat gagtcggtta gtttgggtat aggggtgttt cgattgtcgg 540
 50 gcggcgggcg gaagcgtagg ggttttgatc gtttttagag gtgttgagcg tatagggtag 600
 tggttgtaga gtttcggcgc gaggttaatt cggattattt acgttgggat tttagtttgt 660
 55 tcggttgttt ttagtaataa taatagtatg aaatcgagtt ttaaaaaatt ttttttcgta 720
 ttttttttta tgtataaggg agaataagta ttttttaggta gttaaagaaa aaaaaagggt 780

60

ES 2 533 767 T3

5
 10
 15
 20
 25
 30
 35
 40
 45
 50
 55
 60

aggattttag gaaatnttaa tttttatgga cgaggagtga aggggaattc gttttttgtt 840
 ttttttgttt gtttgttttt agaagagaaa aaaaggaaat tcgatataata tttttaaaaa 900
 ttttttttta ataatcgta aaggttgtga gtgtgagtag agagttagga aatcgagttt 960
 tcgttcgcgg tgatatcggg tttttgtttt tttttttgaa gttttcgcga ttgttgatcg 1020
 tgtagagata aaggtcgggt tagggttttg tggaaatttc agggaggggc ggccggggtcg 1080
 ttttattggt tttttttaat tttatgtagg ggggttaggt gggtttagtt gattttcgga 1140
 ataggggta taaggtagg ggttgttttt attttaggt ttagacgagg tacggatttg 1200
 ttttttttta ttttttcgaa aggtttttgt cgtaaattt tttgtatttt gtttttatta 1260
 gtattatatt gttttattta aagtttttgg aggcgtagg gatagtgggt ttttatcgga 1320
 ttcgattttt agaaatggaa tcgtgttaa tttttgtggc gattttgggt ttcgcgatcg 1380
 ttgggttcgg ttggtcgggt agattcgaat tcgtgtatcg tttgttttag ttgcgttttt 1440
 taggggcgtt tttgtgtttt tttcgggtcg cggagatttt ttagcgttc ggtttcgaat 1500
 tttttgcgta acgttttttt gtagtcgta cggttttgta gagaaggcgg gagagtgttt 1560
 gtgattttgt tttgtagttt tcggtgattt aaaaatttag gtcgttgtgg aatttagttt 1620
 ttacggatatt gttttttag gtagtagttt tgggtttata tttgaggtcg tttttttat 1680
 agatattttt gggtagtcg agtggtgcgt agtttagaat ttttgaagt atgagaggag 1740
 gacgagttcg gcggtttgtc ggtgcggtga tggagggttt ttcggtagtt atattgcgcg 1800
 ttttttatat atattttaat agttttgggg tttcgggta gcgtttagt ttaaagagat 1860
 tttgtagtcg ttttttttt taggtcgtgg taaattttgg tttcgttttt ttttattatg 1920
 gtttgtgagg gtaggaggtt ttcgggggtc gatttttttg atttgaatga gttaggtatt 1980
 tttttttatt ttaagtgggg gtgtgatttg aggagataaa ggaaggacga tagatttggg 2040
 gagtgggggt ttgtagtttt ttgggttcgg agcgtatttt ttttcgtttt tttttttat 2100
 cggtaggatc gggggttttg aaattggagt cgtttatcgg tcggcggatg agagataaag 2160
 gagttcggcg gcgtcgaagc gttgattagt agattcggat tcggattttt tcgcgtttta 2220
 gagtagtatt cgaaatttta ttttttaggt ggtatcgacg gttgagtagt ttttcgggtg 2280
 ttatacgggtg ttttttaggta tagatagtcg tgttcgcgtg tgtattgtgt acgtataaga 2340

ES 2 533 767 T3

	ttgtaggatg atgttttgta tgagtttgt attagtttt ttgttttag ttagaggat	2400
	acgtatacgt gtgtgtattt gtttaagtgt tagggatagg tgtgtattt tatatatac	2460
5	tgatgtgtat gtgtatata ttgtagaacg ttttagtatt tttttatag gtacgtatag	2520
	ggttggttgt ttttaataag ttattattat atttcgtatt tatataggtt ttttagttt	2580
10	tggtatgatt tttagaatta tttcgggaaa tttttttt ttaaagtta tagatgagga	2640
	aataagtaaa aagaatttaa tgggtattcg tttgtatata ttgtgagaga gatcgaggtt	2700
15	agttaaagat ttcgtttatt gttgaattg ggtgtgtgg tcgaatgtgt ttatagatgg	2760
	gtttgttag cgttttaggc gattttaggt tcggcgggtt ttgg	2804

20
<210> 32

25
<400> 32
000

<210> 33

30
<400> 33
000

<210> 34

35
<400> 34
000

<210> 35

40
<400> 35
000

<210> 36

45
<400> 36
000

<210> 37

50
<400> 37
000

<210> 38

55
<400> 38
000

<210> 39

60
<400> 39
000

ES 2 533 767 T3

<210> 40
 <400> 40
 5 000
 <210> 41
 <400> 41
 10 000
 <210> 42
 <211> 2804
 <212> ADN.
 15 <213> Secuencia artificial
 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 20 <400> 42

ttaaggtttg ttggatttga ggttgtttgg ggtgttaaata agatttattt atgagtatat 60
 ttggttaata tattaagtt taatagtggg tgggattttt ggttggtttt gatttttttt 120
 atagtgtgtg taaatgggta tttattagat tttttttggt tgttttttta tttgtaaatt 180
 ttgaaaggag agtatttttt ggggtgattt tgaaaattat gttagaattg aaaaggtttg 240
 tgtaaagtgt aggtgtaatg atgatttatt agaagtagtt agttttgtat gtgtttgtag 300
 gatagatggt agaatgtttt ataaatgtgt gtatatatat attatgtatg tgtgaagggtg 360
 35 tatatttgtt tttagatatt tgataaatgt atatatgtat gtgtgtattt ttggttgaga 420
 gtaaaggggt taatataggg tttatatagg atattatttt gtagttttgt gtgtgtatag 480
 40 tatatatgtg agtatgattg tttgtgtttg gagatattgt atggtattgg aggagttggt 540
 tagttgttgg tgttatttgg gaagtagggt tttgggtggt gttttggagt gtgggaaagt 600
 ttggatttgg gtttgttgggt tagtgttttg gtgttgttgg gtttttttat tttttatttg 660
 45 ttggttgggt agtggtttta gttttaagat ttttgatttt gttggtgagg ggaggagtg 720
 agaaggagtg tgttttgggt ttaagagggt gtagggtttt attttttagg tttgttgttt 780
 50 tttttttatt ttttttaggt atatttttat ttaaaataaa gaagggtggt tgatttattt 840

55
 60

aggttaggga agttggtttt tgggagtttt ttgtttttat aagttatggt gaagggaggt 900
 5 gaagttaggg tttgttatgg tttgggagaa aatgtggttg tagggttttt ttgggttgta 960
 gtgttggttt ggggttttaa gattgttaag gtgtgtgtgg gaggtgtgta gtgtggttat 1020
 10 tgaagagttt tttattattg tattggtagg ttgttgggtt tgtttttttt ttatgttttt 1080
 aagggttttg agttgtgtag tatttgtatt atttagaagt gtttgtggag aaaatgattt 1140
 taggtgtaag ttttaaggtt ttgtttgtaa aggtaatgtt gtggagattg ggttttatag 1200
 15 tgatttgggt ttttaagtta ttgaaagtta tagggtaggg ttataaatat ttttttgttt 1260
 tttttgtaga attgtgtggt tgataggaga gtgttgtgta gaaaatttga ggttgggtgt 1320
 20 tggaggagtt tttgtggttt ggagaaggta tagagggtt tttgagaggt gtagttggaa 1380
 taggtgatgt atgggtttgg atttggttgg ttaattgagt ttagtgggtg tgaaaattag 1440
 agttgttata gaggttgagt atgattttat ttttgggat tgggtttggt ggggagttat 1500
 25 tgtttttaat gtttttagag attttaagta ggataatata atgttggtag gagtaagggt 1560
 taaggaattt agtggtagaa gttttttggg ggggtgggga aaggtagatt tgtgttttgt 1620
 30 ttgagtttg ggggtgggat agtttttgggt tttgtagttt ttgttttggg gattagttgg 1680
 gtttatttaa tttttttata tggggttgaa aggggttaat gggatggttt tgtttgtttt 1740
 tttttgggat tttatagggt tttgatttgg tttttatttt tgtatggta gtagttgtgg 1800
 35 gggttttagg aaggaggata aaggtttgggt gttattgtgg gtgggggttt ggttttttgg 1860
 tttttgttt atatttatag tttttagtggt ttgttggggg aagattttta aaaatatgtg 1920
 40 ttgaattttt tttttttttt ttttagaaat aaataaataa aaaaggtaaa aggtgaattt 1980
 ttttttattt tttgtttata gagattaaag ttttttggga ttttgttttt tttttttttt 2040
 45 tgattgttta gaaatatttg tttttttttg tatatagagg aaaatgtggg gaaaggtttt 2100
 ttaaaatttg gttttatatt attattatta ttaaggataa ttgggtaggt tgaggtttta 2160
 50 atgtggatga tttgagttgg ttttgtgttg gggttttgta gttattgttt tgtgtgttta 2220
 gtatttttgg ggggtgattag ggtttttgtg tttttgttg ttgtttggta gttgagagta 2280
 ttttgtgttt agattggttg atttattttt ttttgaattt tgttttagagt tggttaagggg 2340
 55 gatttagttt gtgttttaag atttgggttt gtagtgttgt taataggttt ggggatatga 2400

60

ES 2 533 767 T3

5 gggtgttttag gttggggttt ttttggttgt tggttttttt tgttttttat ttgtttggtg 2460
 tgttttggtt gtttgaatt gatttaattt gttttttgtg tttgtttttt aggttttttg 2520
 ttttttata aaggtttagg ggagttttgt ttataggttg attttgtaat ttttggtttg 2580
 10 gtggttattt tttgtttttt tgaaaaagaa aaggaaaaaa aaaaaaaaaa agaaaaaatt 2640
 aattagttga gggagggtg tgaggattgg aggtgttagt tggatagttt atgagtaggt 2700
 ttttgggttg gtgttttttt gtgggttagt atggtttttt ttagagaaaa ttttttaaat 2760
 15 gtgtgaagat tgttttgggg gaagtgagag ggaggttgga ggag 2804

20 <210> 43
 <211> 2804
 <212> ADN.
 <213> Secuencia artificial

25 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

 <400> 43
 30 tttttttaat ttttttttg ttttttttaa agtggttttt atatgttttag gaggtttttt 60
 ttgagaaagg ttgtgttgat ttgtgaagtg gtattgattt agggatttat ttataggttg 120
 tttggttggt gtttttagtt tttatgtgtt ttttttgatt ggttgatttt tttttttttt 180
 35 tttttttttt tttttttttt tttagaaaag tagaaggtag ttattgggtt agaggttgta 240
 gggttagttt gtgggtgagg ttttttagg tttttgtgga gaaatgggaa atttgagggg 300
 40 taaagttagg aagtgggttg ggttaattgt gggtgattga ggtgttgta gtgggtgggg 360
 agtgagagg gttagtagtt gggaagattt tggttggag tgttttgtgt ttttgggttt 420
 gttgggtggt ttgtaagttt aggttttggg gtgtgagtta agtttttttt gttagtttta 480
 45 aataaaattt gggggagaat gagttggta gtttgggtat aggggtgttt tgattgttgg 540
 gtggtgggtg gaagttagg ggttttgatt gtttttagag gtgttgagtg tatagggtag 600
 50 tggttgtaga gttttgggtg gaggttaatt tggattattt atgttgggat tttagtttgt 660
 ttggttgttt ttagtaataa taatagtatg aaattgagtt ttaaaaaatt tttttttgta 720
 55 ttttttttta tgtataaggg agaataagta tttttaggta gttaaagaaa aaaaaagggt 780
 aggatttttag gaaattttta tttttatgga tgaggagtga aggggaattt gttttttggt 840

60

ttttttgttt gtttgttttt agaagagaaa aaaaggaaat ttgatata tttttaaaaa 900
 ttttttttta ataattgtta aaggttgtga gtgtgagtag agagttagga aattgagttt 960
 5 ttgtttggg tgatattggg tttttgtttt ttttttgaa gtttttgtga ttgttgattg 1020
 tgtagagata aaggttgggt tagggttttg tggaattttg agggaggggt ggtggggttg 1080
 10 ttttattggt ttttttaat tttatgtagg ggggttaggt gggtttagtt gatttttggg 1140
 ataggggtta taaggtagg ggttgtttt attttaggt ttagatgagg tatggattg 1200
 ttttttttta ttttttgaa aggtttttgt tgttaaattt tttgtattt gttttatta 1260
 15 gtattatatt gttttattta aagttttgg aggtttagg gatagtggtt ttttattgga 1320
 tttgattttt agaaatggaa ttgtgttaa tttttgtgtt gattttggtt tttgtgattg 1380
 20 ttgggtttgg ttggttggtt agatttgaat ttgtgtattg tttgttttag ttgtgtttt 1440
 taggggtgtt tttgtgtttt ttttgggtt tgagatttt tttagtgtt ggttttgaat 1500
 25 tttttgtgta atgtttttt gttagtgtg tggttttga gagaagggtg gagagtgtt 1560
 gtgattttgt tttgtagttt ttggtgattt aaaaatttag gttgttggg aatttagttt 1620
 ttatggtatt gttttttag gtagtagttt tgggtttata tttgaggtt tttttttat 1680
 30 agatattttt gggtagtgt agtgttgtt agtttagaat ttttgaagt atgagaggag 1740
 gatgagttt gttgtttgtt ggtgtgtga tggagggttt tttgtagtt atattgtgt 1800
 35 tttttatata atattttaat agttttggg ttttgggtta gtgtttagt ttaaagagat 1860
 tttgtagttg tattttttt taggttggg taaattttg tttgtttt ttttattatg 1920
 40 gtttgtgagg gtaggaggt tttgggggtt gattttttt atttgaatga gttaggtatt 1980
 ttttttatt ttaagtggg gtgtgattg aggagataaa ggaaggatga tagatttggg 2040
 45 gagtggggtt ttgtagttt ttgggttgg agtgtattt tttttgttt tttttttat 2100
 tggtaggatt ggggtttt aaattggagt tgtttattg ttggtgatg agagataaag 2160
 gattttggtg gtgttgaagt gttgattagt agatttggat ttggattttt ttgtgtttta 2220
 50 gagttagtatt tgaatttta ttttttaggt ggtattgatg gttgagtagt ttttttggg 2280
 ttatatggtg ttttaggta tagatagttg tgtttgtgtg tgtattgtg atgtataaga 2340
 55 ttgttaggat atgttttga tgagttttgt attagttttt ttgttttag ttagaggtat 2400

60

ES 2 533 767 T3

	atgtatatgt gtgtgtatgt gtttaagtgt tagggatagg tgtgtatgtt tatatatatg	2460
	tgatgtgtat gtgtatatat ttgtagaatg ttttagtatt tttttatag gtatgtatag	2520
5	ggttggttgt ttttaataag ttattattat attttgtatt tatatagggt ttttagttt	2580
	tggtatgatt tttagaatta ttttgggaaa ttttttttt ttaaagtta tagatgagga	2640
10	aataagtaaa aagaattaa tgggtatttg tttgtatata ttgtgagaga gattgaggt	2700
	agttaaagat tttgtttatt gttgaatttg ggtgtgttg ttgaatgtgt ttatagatgg	2760
15	gtttgtttag tgttttaggt gattttaggt ttggtgggtt ttgg	2804

20 <210> 44
<400> 44
000

25 <210> 45
<400> 45
000

30 <210> 46
<400> 46
000

35 <210> 47
<400> 47
000

40 <210> 48
<400> 48
000

45 <210> 49
<400> 49
000

50 <210> 50
<400> 50
000

55 <210> 51
<400> 51
000

60 <210> 52

<400> 52
000

5 <210> 53

<400> 53
000

10 <210> 54

<400> 54
000

15 <210> 55

<400> 55
000

20 <210> 56

<400> 56
000

25 <210> 57

<400> 57
000

30 <210> 58

<400> 58
000

35 <210> 59

<400> 59
000

40 <210> 60

<400> 60
000

45 <210> 61

<400> 61
000

50 <210> 62
<211> 20
<212> ADN.
<213> Secuencia artificial

55 <220>
<223> ADN genómico químicamente tratado (Homo sapiens)

<400> 62
ggagtggagg aaattgagat 20

60

<210> 63
<211> 22
<212> ADN.
<213> Secuencia artificial
5
<220>
<223> ADN genómico químicamente tratado (Homo sapiens)
10 <400> 63
ccacacaaca aatactcaaa ac 22
<210> 64
<211> 33
<212> ADN.
15 <213> Secuencia artificial
<220>
<223> ADN genómico químicamente tratado (Homo sapiens)
20 <400> 64
tgggtgttg taattttgt ttgtgtag gtt 33
<210> 65
25 <400> 65
000
<210> 66
30 <400> 66
000
<210> 67
35 <400> 67
000
<210> 68
40 <400> 68
000
<210> 69
45 <400> 69
000
<210> 70
50 <400> 70
000
<210> 71
55 <400> 71
000
<210> 72

<400> 72
000

5 <210> 73
<400> 73
000

10 <210> 74
<400> 74
000

15 <210> 75
<400> 75
000

20 <210> 76
<400> 76
000

<210> 77

25 <400> 77
000

<210> 78

30 <400> 78
000

<210> 79

35 <400> 79
000

<210> 80

40 <400> 80
000

<210> 81

45 <400> 81
000

<210> 82

50 <400> 82
000

<210> 83

55 <400> 83
000

<210> 84

<400> 84
000

5 <210> 85
<400> 85
000

10 <210> 86
<400> 86
000

15 <210> 87
<400> 87
000

20 <210> 88
<400> 88
000

25 <210> 89
<400> 89
000

<210> 90

30 <400> 90
000

<210> 91

35 <400> 91
000

<210> 92

40 <400> 92
000

<210> 93

45 <400> 93
000

<210> 94

50 <400> 94
000

<210> 95

55 <400> 95
000

<210> 96

<400> 96
 000
 5 <210> 97
 <400> 97
 000
 10 <210> 98
 <400> 98
 000
 15 <210> 99
 <400> 99
 000
 20 <210> 100
 <400> 100
 000
 25 <210> 101
 <400> 101
 000
 30 <210> 102
 <211> 20
 <212> ADN.
 <213> Secuencia artificial
 35 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 <400> 102 ccaaaaccta aactacaac 20
 40 <210> 103
 <211> 17
 <212> ADN.
 <213> Secuencia artificial
 45 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 <400> 103
 ggaaattga ggggtaa 17
 50 <210> 104
 <211> 31
 <212> ADN.
 <213> Secuencia artificial
 55 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 <400> 104
 tacaacacca ccaacaaacc caaaaacaca a 31
 60

<210> 105
<211> 19
<212> ADN.
<213> Secuencia artificial
5
<220>
<223> ADN genómico químicamente tratado (Homo sapiens)
<400> 105
10 gttaattgcg ggcgatcga 19
<210> 106
<211> 17
<212> ADN.
15 <213> Secuencia artificial
<220>
<223> ADN genómico químicamente tratado (Homo sapiens)
20 <400> 106
cgtcgtagc gggggg 17
<210> 107
25 <400> 107
000
<210> 108
30 <400> 108
000
<210> 109
35 <400> 109
000
<210> 110
<211> 24
40 <212> ADN.
<213> Secuencia artificial
<220>
<223> ADN genómico químicamente tratado (Homo sapiens)
45 <400> 110
tctaaataac aaaatacctc catt 24
<210> 111
50 <400> 111
000
<210> 112
55 <400> 112
000
<210> 113
60

<400> 113
000

5 <210> 114
<400> 114
000

10 <210> 115
<400> 115
000

15 <210> 116
<400> 116
000

20 <210> 117
<400> 117
000

25 <210> 118
<400> 118
000

30 <210> 119
<400> 119
000

35 <210> 120
<400> 120
000

40 <210> 121
<400> 121
000

<210> 122

45 <400> 122
000

<210> 123

50 <400> 123
000

<210> 124

55 <400> 124
000

<210> 125

<400> 125
000

5 <210> 126
<400> 126
000

10 <210> 127
<400> 127
000

15 <210> 128
<400> 128
000

20 <210> 129
<400> 129
000

25 <210> 130
<400> 130
000

30 <210> 131
<400> 131
000

35 <210> 132
<400> 132
000

40 <210> 133
<400> 133
000

45 <210> 134
<400> 134
000

<210> 135

50 <400> 135
000

<210> 136

55 <400> 136
000

<210> 137

<400> 137
000

5 <210> 138
<400> 138
000

10 <210> 139
<400> 139
000

15 <210> 140
<400> 140
000

20 <210> 141
<400> 141
000

25 <210> 142
<400> 142
000

30 <210> 143
<400> 143
000

35 <210> 144
<400> 144
000

40 <210> 145
<400> 145
000

45 <210> 146
<400> 146
000

50 <210> 147
<400> 147
000

55 <210> 148
<400> 148
000

60 <210> 149

ES 2 533 767 T3

<400> 149
000
<210> 150

5 <400> 150
000

<210> 151

10 <400> 151
000

<210> 152

15 <400> 152
000

<210> 153

20 <400> 153
000

<210> 154

25 <400> 154
000

<210> 155

30 <400> 155
000

<210> 156

35 <400> 156
000

<210> 157

40 <400> 157
000

<210> 158

45 <400> 158
000

<210> 159

50 <400> 159
000

<210> 160

55 <400> 160
000

<210> 161

<400> 161
000

5 <210> 162
<400> 162
000

10 <210> 163
<400> 163
000

15 <210> 164
<400> 164
000

20 <210> 165
<400> 165
000

25 <210> 166
<400> 166
000

30 <210> 167
<400> 167
000

35 <210> 168
<400> 168
000

40 <210> 169
<400> 169
000

45 <210> 170
<400> 170
000

50 <210> 171
<400> 171
000

55 <210> 172
<400> 172
000

60 <210> 173

<400> 173
000

5 <210> 174
<400> 174
000

10 <210> 175
<400> 175
000

15 <210> 176
<400> 176
000

20 <210> 177
<400> 177
000

25 <210> 178
<400> 178
000

30 <210> 179
<400> 179
000

35 <210> 180
<400> 180
000

40 <210> 181
<400> 181
000

45 <210> 182
<400> 182
000

50 <210> 183
<400> 183
000
<210> 184

55 <400> 184
000
<210> 185

60 <400> 185
000

<210> 186
5 <400> 186
000
<210> 187
10 <400> 187
000
<210> 188
15 <400> 188
000
<210> 189
20 <400> 189
000
<210> 190
25 <400> 190
000
<210> 191
30 <400> 191
000
<210> 192
35 <400> 192
000
<210> 193
40 <400> 193
000
<210> 194
45 <400> 194
000
<210> 195
50 <400> 195
000
<210> 196
55 <400> 196
000
<210> 197
60 <400> 197
000

<210> 198
5 <400> 198
000
<210> 199
10 <400> 199
000
<210> 200
15 <400> 200
000
<210> 201
20 <400> 201
000
<210> 202
25 <400> 202
000
<210> 203
30 <400> 203
000
<210> 204
35 <400> 204
000
<210> 205
40 <400> 205
000
<210> 206
45 <400> 206
000
<210> 207
50 <400> 207
000
<210> 208
55 <400> 208
000
<210> 209
60 <400> 209
000

<210> 210
5 <400> 210
000
<210> 211
10 <400> 211
000
<210> 212
15 <400> 212
000
<210> 213
20 <400> 213
000
<210> 214
25 <400> 214
000
<210> 215
30 <400> 215
000
<210> 216
35 <400> 216
000
<210> 217
40 <400> 217
000
<210> 218
45 <400> 218
000
<210> 219
50 <400> 219
000
<210> 220
55 <400> 220
000
<210> 221
60 <400> 221
000

<210> 222
5 <400> 222
000
<210> 223
10 <400> 223
000
<210> 224
15 <400> 224
000
<210> 225
20 <400> 225
000
<210> 226
25 <400> 226
000
<210> 227
30 <400> 227
000
<210> 228
35 <400> 228
000
<210> 229
40 <400> 229
000
<210> 230
45 <400> 230
000
<210> 231
50 <400> 231
000
<210> 232
55 <400> 232
000
<210> 233
60 <400> 233
000

<210> 234
5 <400> 234
000
<210> 235
10 <400> 235
000
<210> 236
15 <400> 236
000
<210> 237
20 <400> 237
000
<210> 238
25 <400> 238
000
<210> 239
30 <400> 239
000
<210> 240
35 <400> 240
000
<210> 241
40 <400> 241
000
<210> 242
45 <400> 242
000
<210> 243
50 <400> 243
000
<210> 244
55 <400> 244
000
<210> 245
60 <400> 245
000

<210> 246
5 <400> 246
000
<210> 247
10 <400> 247
000
<210> 248
15 <400> 248
000
<210> 249
20 <400> 249
000
<210> 250
25 <400> 250
000
<210> 251
30 <400> 251
000
<210> 252
35 <400> 252
000
<210> 253
<400> 253
40 000
<210> 254
<211> 20
<212> ADN.
45 <213> Secuencia artificial
<220>
<223> ADN genómico químicamente tratado (Homo sapiens)
50 <400> 254
ccaaaccta aactacaac 20
<210> 255
<211> 17
55 <212> ADN.
<213> Secuencia artificial
<220>
<223> ADN genómico químicamente tratado (Homo sapiens)
60

<400> 255
 ggaaattga ggggtaa 17

5 <210> 256
 <211> 31
 <212> ADN.
 <213> Secuencia artificial

10 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

<400> 256
 tacaacacca ccaacaaacc caaaaacaca a 31

15 <210> 257
 <211> 17
 <212> ADN.
 <213> Secuencia artificial

20 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

25 <400> 257
 ggcgatcgag gcgtcgt 17

30 <210> 258
 <211> 16
 <212> ADN.
 <213> Secuencia artificial

<223> ADN genómico químicamente tratado (Homo sapiens)

35 <400> 258
 gcgggtgggg agcgag 16

40 <210> 259

<400> 259
 000

45 <210> 260

<400> 260
 000

50 <210> 261

<400> 261
 000

55 <210> 262

<400> 262
 000

60 <210> 263

<400> 263
 000

<210> 264
 <211> 20
 <212> ADN.
 <213> Secuencia artificial
 5
 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 <400> 264
 10 ccaaaaccta aactacaac 20
 <210> 265
 <211> 17
 <212> ADN.
 15 <213> Secuencia artificial
 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 20 <400> 265
 ggaaattga ggggtaa 17
 <210> 266
 <211> 31
 25 <212> ADN.
 <213> Secuencia artificial
 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 30 <400> 266
 tacaacacca ccaacaaacc caaaacaca a 31
 <210> 267
 <211> 17
 <212> ADN.
 <213> Secuencia artificial
 <220>
 40 <223> ADN genómico químicamente tratado (Homo sapiens)
 <400> 267
 ggcgatcgag gcgtcgt 17
 45 <210> 268
 <211> 16
 <212> ADN.
 <213> Secuencia artificial
 50 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)
 <400> 268
 55 gcgggtgggg agcgag 16
 <210> 269
 <400> 269
 60 000

<210> 270
<400> 270
000
5 <210> 271
<400> 271
000
10 <210> 272
<400> 272
000
15 <210> 273
<400> 273
000
20 <210> 274
<400> 274
000
25 <210> 275
<400> 275
000
30 <210> 276
<400> 276
000
35 <210> 277
<400> 277
000
40 <210> 278
<400> 278
000
45 <210> 279
<400> 279
000
50 <210> 280
<400> 280
000
55 <210> 281
<400> 281
000
60

<210> 282
<400> 282
000
5
<210> 283
<400> 283
000
10
<210> 284
<400> 284
000
15
<210> 285
<400> 285
000
20
<210> 286
<400> 286
000
25
<210> 287
<400> 287
000
30
<210> 288
<400> 288
000
35
<210> 289
<400> 289
000
40
<210> 290
<400> 290
000
45
<210> 291
<400> 291
000
50
<210> 292
<400> 292
000
55
<210> 293
<400> 293
000
60

<210> 294
<400> 294
000
5
<210> 295
<400> 295
000
10
<210> 296
<400> 296
000
15
<210> 297
<400> 297
000
20
<210> 298
<400> 298
000
25
<210> 299
<400> 299
000
30
<210> 300
<400> 300
000
35
<210> 301
<400> 301
000
40
<210> 302
<400> 302
000
45
<210> 303
<400> 303
000
50
<210> 304
<400> 304
000
55
<210> 305
<400> 305
000
60

<210> 306
<400> 306
000
5 <210> 307
<400> 307
000
10 <210> 308
<400> 308
000
15 <210> 309
<400> 309
000
20 <210> 310
<400> 310
000
25 <210> 311
<400> 311
000
30 <210> 312
<400> 312
000
35 <210> 313
<400> 313
000
40 <210> 314
<400> 314
000
45 <210> 315
<400> 315
000
50 <210> 316
<400> 316
000
55 <210> 317
<400> 317
000
60 <210> 318

ES 2 533 767 T3

<400> 318
000

5 <210> 319

<400> 319
000

10 <210> 320

<400> 320
000

15 <210> 321

<400> 321
000

20 <210> 322

<400> 322
000

25 <210> 323

<400> 323
000

30 <210> 324

<400> 324
000

35 <210> 325

<400> 325
000

40 <210> 326

<400> 326
000

45 <210> 327

<400> 327
000

50 <210> 328

<400> 328
000

55 <210> 329

<400> 329
000

60 <210> 330

ES 2 533 767 T3

<400> 330
000

5 <210> 331

<400> 331
000

10 <210> 332

<400> 332
000

15 <210> 333

<400> 333
000

20 <210> 334

<400> 334
000

<210> 335

25 <400> 335
000

<210> 336

30 <400> 336
000

<210> 337

35 <400> 337
000

<210> 338

40 <400> 338
000

<210> 339

45 <400> 339
000

<210> 340

50 <400> 340
000

<210> 341

55 <400> 341
000

<210> 342

<400> 342
000

5 <210> 343
<400> 343
000

10 <210> 344
<400> 344
000

15 <210> 345
<400> 345
000

20 <210> 346
<400> 346
000

25 <210> 347
<400> 347
000

30 <210> 348
<400> 348
000

35 <210> 349
<400> 349
000

40 <210> 350
<400> 350
000

45 <210> 351
<400> 351
000

50 <210> 352
<400> 352
000

55 <210> 353
<400> 353
000

60 <210> 354

	<400> 354 000
5	<210> 355 <400> 355 000
10	<210> 356 <400> 356 000
15	<210> 357 <400> 357 000
20	<210> 358 <400> 358 000
25	<210> 359 <400> 359 000
30	<210> 360 <400> 360 000
35	<210> 361 <400> 361 000
40	<210> 362 <400> 362 000
45	<210> 363 <400> 363 000
50	<210> 364 <400> 364 000
55	<210> 365 <211> 20 <212> ADN. <213> Secuencia artificial
60	<220> <223> ADN genómico químicamente tratado (Homo sapiens)

<400> 365
 ccaaaaccta aactacaac 20

5 <210> 366
 <211> 17
 <212> ADN.
 <213> Secuencia artificial

10 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

<400> 366
 ggaaattga ggggtaa 17

15 <210> 367
 <211> 31
 <212> ADN.
 <213> Secuencia artificial

20 <220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

<400> 367
 25 tacaacacca ccaacaaacc caaaaacaca a 31

<210> 368
 <211> 19
 <212> ADN.
 30 <213> Secuencia artificial

<220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

35 <400> 368
 gttaattgcg ggcgatcga 19

<210> 369
 <211> 17
 40 <212> ADN.
 <213> Secuencia artificial

<220>
 <223> ADN genómico químicamente tratado (Homo sapiens)

45 <400> 369
 cgtcgtagc gggtagg 17

REIVINDICACIONES

- 5 1. Un método para detectar y/o clasificar el carcinoma colorrectal en un sujeto que comprende determinar la presencia o ausencia de metilación CpG de FOXL2 en una muestra biológica aislada de dicho sujeto en donde la hipermetilación CpG de FOXL2 es indicativa de la presencia o clase de dicho carcinoma colorrectal.
- 10 2. El método de acuerdo con la reivindicación 1, que comprende poner en contacto el ADN genómico aislado de una muestra biológica que se obtuvo de dicho sujeto con al menos un reactivo, o serie de reactivos que distinguen entre dinucleótidos CpG metilados y no metilados dentro de al menos una región objetivo del ADN genómico, en donde la región objetivo comprende, o hibrida bajo condiciones rigurosas a una secuencia de al menos 16 nucleótidos contiguos de FOXL2 respectivamente, en donde dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótidos CpG, y en el cual se proporciona, al menos en parte, la detección y/o clasificación del carcinoma colorrectal.
- 15 3. El método de acuerdo con la reivindicación 1, que comprende:
- 20 a) extraer o de cualquier otra forma aislar ADN genómico de una muestra biológica que se obtuvo del sujeto
 b) tratar el ADN genómico de a), o un fragmento de este, con uno o más reactivos para convertir las bases de citosina que no son metiladas en la posición 5 de esta a uracilo o a otra base que se detecta diferente a la citosina en términos de propiedades de hibridación;
 c) poner en contacto el ADN genómico tratado, o el fragmento tratado del mismo con una enzima de amplificación y al menos un iniciador que comprende, una secuencia contigua de al menos 9 nucleótidos que es complementaria a, o hibrida en condiciones rigurosas o moderadamente rigurosas a una secuencia que se selecciona del grupo que consiste en la SEQ ID NO: 30, SEQ ID NO: 31, SEQ ID NO: 42, SEQ ID NO: 43, y complementos de estas, en donde el ADN genómico tratado o el fragmento de este se amplifica para producir al menos un amplificado, o no se amplifica; y
 d) determinar, basad en una presencia o ausencia de, o en una propiedad de dicho amplificado, el estado o nivel de metilación de al menos un dinucleótido CpG, de FOXL2, o un promedio, o un valor que refleja un estado o nivel de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG of FOXL2, de manera que al menos se proporciona al menos en parte uno de detección y clasificación del carcinoma colorrectal.
- 25 4. El método de acuerdo con la reivindicación 5, en donde tratar el ADN genómico, o el fragmento de este en b), comprende usar un reactivo que se seleccione del grupo que comprende bisulfito, sulfito de hidrógeno, disulfito, y combinaciones de estos.
- 30 5. El método de acuerdo con la reivindicación 1, que comprende:
- 35 a) extraer o de cualquier otra forma aislar ADN genómico de una muestra biológica que se obtiene del sujeto;
 b) digerir el ADN genómico de a), o un fragmento del mismo, con una o más enzimas de restricción sensibles a la metilación;
 c) poner en contacto la digestión con enzima de restricción de ADN de b), con una enzima de amplificación y al menos dos iniciadores adecuados para la amplificación de una secuencia que comprende al menos un dinucleótido CpG de FOXL2 ; y
 d) determinar, sobre la base de una presencia o ausencia de un amplificado el estado de metilación o el nivel de al menos un dinucleótido CpG de FOXL2, de manera que al menos uno de entre la detección y clasificación del carcinoma colorrectal, al menos en parte, se logre.
- 40 6. El uso de un método de acuerdo con cualquiera de las reivindicaciones 1 a 5 en el diagnóstico y/o clasificación de carcinoma colorrectal.
- 45
- 50

Figura 1

