

19



OFICINA ESPAÑOLA DE
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 801 379**

51 Int. Cl.:

C12Q 1/6886 (2008.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

96 Fecha de presentación y número de la solicitud europea: **01.12.2004** **E 17186907 (6)**

97 Fecha y número de publicación de la concesión europea: **11.03.2020** **EP 3269826**

54 Título: **Métodos y ácidos nucleicos para el análisis de la expresión génica asociada con el desarrollo de trastornos proliferativos de células de próstata**

30 Prioridad:

01.12.2003 EP 03090414

10.02.2004 EP 04090040

10.05.2004 EP 04090187

21.07.2004 EP 04090292

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

11.01.2021

73 Titular/es:

EPIGENOMICS AG (100.0%)

Geneststrasse 5

10829 Berlin, DE

72 Inventor/es:

DAY, KEVIN;

COTTRELL, SUSAN;

DISTLER, JÜRGEN;

MOROTTI, ANDREW M.;

YAMAMURA, SU;

DEKKER, SHARON;

OCAMPO, YREKA y

DEVOS, THEO

74 Agente/Representante:

SÁEZ MAESO, Ana

ES 2 801 379 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 y ácidos nucleicos para el análisis de la expresión génica asociada con el desarrollo de trastornos proliferativos de células de próstata

Campo de la invención

- 5 La presente invención se refiere a secuencias de ADN humano que exhiben patrones de expresión heterogéneos alterados en pacientes con cáncer de próstata. Realizaciones particulares de la invención proporcionan métodos para el diagnóstico de individuos con dicho trastorno.

Antecedentes

- 10 Correlación de la metilación de ADN aberrante con cáncer. La metilación aberrante de ADN dentro de las 'islas' de CpG se caracteriza por la hiper o hipometilación de secuencias de dinucleótidos CpG que conducen a la abrogación o sobreexpresión de un amplio espectro de genes, y se encuentra entre las alteraciones más tempranas y más comunes encontradas y correlacionadas con tumores malignos humanos. Además, se ha demostrado que se produce metilación anormal en elementos reguladores ricos en CpG en partes intrónicas y codificantes de genes para ciertos tumores. En el cáncer de colon, la metilación aberrante del ADN constituye una de las alteraciones más prominentes e inactiva muchos genes supresores de tumores que incluyen, entre otros, p14ARF, p16INK4a, THBS1, MINT2 y MINT31 y genes de reparación de desajuste de ADN como hMLH1.

- 15 Además de la hipermetilación específica de los genes supresores de tumores, se puede observar una hipometilación global del ADN en las células tumorales. Esta disminución en la metilación global se puede detectar temprano, mucho antes del desarrollo de la formación de tumores francos. Se ha determinado una correlación entre la hipometilación y el aumento de la expresión génica para muchos oncogenes.

- 20 Cáncer de próstata. El cáncer de próstata es la neoplasia maligna más común entre los hombres en los Estados Unidos (200.000 casos nuevos por año), y la sexta causa principal de muertes masculinas relacionadas con el cáncer en todo el mundo (204.000 por año). El cáncer de próstata es principalmente una enfermedad de los ancianos, con aproximadamente el 16% de los hombres entre las edades de 60 y 79 con la enfermedad. Según algunas estimaciones en la autopsia, el 80% de todos los hombres mayores de 80 años tienen alguna forma de enfermedad de la próstata (por ejemplo, cáncer, HBP, prostatitis, etc.). La hipertrofia benigna de próstata está presente en aproximadamente el 50% de los hombres de 50 años o más, y en el 95% de los hombres de 75 años o más. Es obvio a partir de estos informes que el cáncer de próstata a menudo no es una enfermedad por la cual, sino con la cual, los hombres mueren. La evidencia reciente sugiere que la incidencia de cáncer de próstata puede estar disminuyendo, probablemente como resultado de un mejor tratamiento, una mejor cirugía y una detección más temprana.

- 25 Diagnóstico de cáncer de próstata; enfoques moleculares. La American Cancer Society ha sugerido las pautas actuales para la detección del cáncer de próstata y son las siguientes: a los 50 años de edad, los profesionales de la salud deben ofrecer un análisis de sangre para el antígeno prostático específico (PSA) y realizar un examen rectal digital (DRE). Se recomienda que las poblaciones de alto riesgo, como los afroamericanos y aquellos con antecedentes familiares de enfermedad de la próstata, comiencen el estudio a los 45 años de edad. Los hombres sin patología prostática anormal generalmente tienen un nivel de PSA en sangre inferior a 4 ng/ml. Los niveles de PSA entre 4 ng/ml y 10 ng/ml (llamada 'zona gris') tienen un 25% de posibilidades de tener cáncer de próstata. El resultado es que el 75% de las veces, los hombres con un DRE anormal y un PSA en esta zona gris tienen una biopsia negativa o aparentemente innecesaria. Por encima de la zona gris, la probabilidad de tener cáncer de próstata es significativa (>67%) y aumenta aún más a medida que aumentan los niveles de PSA. Existen numerosos métodos para medir el PSA (porcentaje de PSA libre, velocidad del PSA, densidad del PSA, etc.), y cada uno tiene una precisión asociada para detectar la presencia de cáncer. Sin embargo, incluso con las mejoras menores en la detección y las caídas reportadas en la mortalidad asociadas con la detección, la frecuencia de falsos positivos sigue siendo alta. La especificidad reducida se debe en parte al aumento de PSA en sangre asociado con HBP y prostatitis. También se ha estimado que hasta el 45% de las biopsias de próstata según las pautas actuales son falsamente negativas, lo que resulta en una disminución de la sensibilidad incluso con la biopsia.

- 30 La biopsia guiada por TRUS se considera el "estándar de oro" para diagnosticar el cáncer de próstata. Las recomendaciones para la biopsia se basan en niveles anormales de PSA y/o un DRE anormal. Para el PSA, hay una zona gris donde quizás no sea necesario un alto porcentaje de biopsias. Sin embargo, la capacidad de detectar cáncer en esta zona gris (niveles de PSA de 4.0 a 10 ng/ml) es difícil sin una biopsia. Debido a esta falta de especificidad, el 75% de los hombres que se someten a una biopsia no tienen cáncer. Sin embargo, sin biopsia, las personas con cáncer serían pasadas por alto, lo que daría lugar a una mayor morbilidad y mortalidad. Desafortunadamente, los riesgos asociados con una biopsia innecesaria también son altos.

- 35 Los marcadores moleculares ofrecerían la ventaja de que pueden usarse para analizar eficientemente incluso muestras de tejido muy pequeñas, y muestras cuya arquitectura de tejido no se ha mantenido. En la última década, se han estudiado numerosos genes con respecto a la expresión diferencial entre la hiperplasia benigna de próstata y los diferentes grados de cáncer de próstata. Sin embargo, todavía no se ha demostrado que un marcador único sea suficiente para la detección de tumores de próstata en un entorno clínico.

Como alternativa, las metodologías basadas en ARNm de alta dimensión pueden, en casos particulares, proporcionar un medio para distinguir entre diferentes tipos de tumores y lesiones benignas y malignas. Sin embargo, la aplicación de tales metodologías como una herramienta de diagnóstico de rutina en un entorno clínico se ve obstaculizada y sustancialmente limitada por la extrema inestabilidad del ARNm, los cambios de expresión que ocurren rápidamente después de ciertos desencadenantes (por ejemplo, recolección de muestras) y, lo más importante, por la gran cantidad de ARNm necesaria para el análisis que a menudo no se puede obtener de una biopsia de rutina (véase, por ejemplo, Lipshutz, R.J. et al., *Nature Genetics* 21: 20-24, 1999; Bowtell, D.D.L. *Nature Genetics Suppl.* 21: 25-32, 1999).

La metilación genética aberrante en el cáncer de próstata se ha observado en varios genes que incluyen Gstp1, AR, p16 (CDKN2a/INK4a), CD44, CDH1. La hipometilación del genoma completo, por ejemplo, del elemento repetitivo LINE-1 también se ha asociado con la progresión tumoral (Santourlidis, S., et al., *Prostate* 39: 166-74, 1999).

Marcadores de metilación del cáncer de próstata. Se ha demostrado que la región promotora central del gen Glutona S-Transferasa P (GSTP1; número de acceso NM_000852) está hipermetilada en el tejido tumoral de próstata. La enzima glutatión S-transferasa pi está involucrada en la desintoxicación de carcinógenos electrofílicos, y los niveles alterados o disminuidos de actividad enzimática (deterioro de GSTP1) se han asociado con el desarrollo de neoplasias, particularmente en la próstata. Los mecanismos del deterioro de GSTP1 incluyen la mutación (el alelo GSTP*B se ha asociado con un mayor riesgo de cáncer) y la metilación. Sin embargo, con respecto al uso de marcadores Gstp1, la técnica anterior está limitada con respecto al número de secuencias CpG del promotor Gstp1 que se han caracterizado por el estado de metilación diferencial. Además, no hay divulgaciones, sugerencias o enseñanzas en la técnica anterior de cómo tales marcadores podrían usarse para distinguir entre hiperplasia benigna de próstata y diferentes grados de cáncer de próstata.

La metilación genética aberrante también se ha observado en varios otros genes, incluidos AR, p16 (CDKN2a/INK4a), CD44, CDH1. La hipometilación del genoma completo, por ejemplo, del elemento repetitivo LINE-1 también se ha asociado con la progresión tumoral (Santourlidis, S., et al., *Prostate* 39: 166-74, 1999).

Sin embargo, no se ha habilitado el uso de estos genes como marcadores alternativos o suplementarios diagnósticos o clínicamente útiles en un entorno comercial. La aplicación de genes diferencialmente metilados a plataformas clínicamente utilizables requiere mucha más investigación sobre la sensibilidad y especificidad de los genes. Por ejemplo, en el caso del gen CD44, un supresor de metástasis conocido, la regulación a la baja se asoció con hipermetilación. Sin embargo, el uso de este gen como marcador comercialmente disponible no estaba habilitado, porque también estaba metilado en tejidos normales (Véase Vis, et al., *Mol Urol.* 5: 199-203, 2001).

Desarrollo de pruebas médicas. Dos medidas de evaluación clave de cualquier examen médico o prueba de diagnóstico son su sensibilidad y especificidad, que miden cuán bien funciona la prueba para detectar con precisión a todas las personas afectadas sin excepción y sin incluir falsamente a las personas que no tienen la enfermedad objetivo (valor predictivo). Históricamente, muchas pruebas de diagnóstico han sido criticadas debido a la poca sensibilidad y especificidad. Un verdadero resultado positivo (TP) es donde la prueba es positiva y la condición está presente. Un resultado falso positivo (FP) es donde la prueba es positiva pero la condición no está presente. Un verdadero resultado negativo (TN) es donde la prueba es negativa y la condición no está presente. Un resultado falso negativo (FN) es donde la prueba es negativa pero la condición 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 una prueba para detectar correctamente la enfermedad objetivo en un individuo sometido a prueba. Una prueba que tiene poca sensibilidad produce una alta tasa de falsos negativos, es decir, individuos que tienen la enfermedad pero son identificados falsamente como libres de esa enfermedad en particular. El peligro potencial de un falso negativo es que el individuo enfermo permanecerá sin diagnóstico y sin tratamiento durante un período de tiempo, durante el cual la enfermedad puede progresar a una etapa posterior en la que los tratamientos, si los hay, pueden ser menos efectivos. Un ejemplo de una prueba que tiene baja sensibilidad es una prueba de sangre basada en proteínas para el VIH. Este tipo de prueba exhibe poca sensibilidad porque no puede detectar la presencia del virus hasta que la enfermedad esté bien establecida y el virus haya invadido el torrente sanguíneo en cantidades sustanciales. En contraste, un ejemplo de una prueba que tiene alta sensibilidad es la detección de carga viral usando la reacción en cadena de la polimerasa (PCR). Se logra una alta sensibilidad porque este tipo de prueba puede detectar cantidades muy pequeñas del virus. La alta sensibilidad es particularmente importante cuando las consecuencias de perder un diagnóstico son altas.

La especificidad, por otro lado, es una medida de la capacidad de una prueba para identificar con precisión pacientes que están libres del estado de la enfermedad. Una prueba que tiene poca especificidad produce una alta tasa de falsos positivos, es decir, individuos que se identifican falsamente como portadores de la enfermedad. Una desventaja de los falsos positivos es que obligan a los pacientes a someterse a tratamientos de procedimientos médicos innecesarios con sus riesgos asociados, tensiones emocionales y financieras, y que podrían tener efectos adversos en la salud del paciente. Una característica de las enfermedades que dificulta el desarrollo de pruebas de diagnóstico con alta especificidad es que los mecanismos de la enfermedad, particularmente en el cáncer, a menudo involucran una pluralidad de genes y proteínas. Además, ciertas proteínas pueden estar elevadas por razones no relacionadas con un estado de enfermedad. Un ejemplo de una prueba que tiene una alta especificidad es una prueba basada en genes

que puede detectar una mutación p53. La especificidad es importante cuando el coste o el riesgo asociado con otros procedimientos de diagnóstico o una intervención médica adicional son muy altos.

La prueba de sangre de PSA tiene una sensibilidad del 73%, una especificidad del 60% y un valor predictivo del 31.5%. La sensibilidad y la especificidad del PSA pueden mejorarse, pero implican compensaciones. La sensibilidad al PSA se puede mejorar ajustando el nivel de PSA "normal" a un valor más bajo para los hombres más jóvenes o siguiendo los valores de PSA en suero en un paciente individual a lo largo del tiempo (velocidad del PSA). Ambos métodos aumentarán la cantidad de cánceres detectados, pero también aumentarán la cantidad de hombres sometidos a biopsia. Por el contrario, la especificidad se puede mejorar mediante el uso de niveles de PSA "normales" más altos para los hombres mayores, mediante el uso de la relación de PSA libre a total, o ajustando el valor normal de acuerdo con el tamaño de la próstata. Estos tres métodos disminuyen la cantidad de biopsias innecesarias, pero aumentan el riesgo de que se pierdan algunos tipos de cáncer.

Por lo tanto, se puede ver que existe la necesidad de un medio de diagnóstico de cáncer de próstata con sensibilidad, especificidad y/o valor predictivo mejorados.

La sensibilidad y especificidad del ensayo cuantitativo de reacción en cadena de la polimerasa específica de metilación (QMSP) solo (sin análisis histológico) en el análisis de cáncer de próstata de biopsias con aguja ha variado desde 30% de sensibilidad y 100% de especificidad a 89% de sensibilidad y 64% de especificidad (Harden et al., J. Natl Cancer Inst 2003; 95: 1634-1637). Sin embargo, no se analizó el valor predictivo de dicha técnica como herramienta de detección clínica. Además, las pruebas genéticas de suero y fluidos corporales como la orina y la saliva reducirían la necesidad de biopsias para detectar el cáncer y, por lo tanto, serían la herramienta de detección o monitorización más efectiva. Sin embargo, el desarrollo de tales pruebas requiere un grado extremadamente alto de sensibilidad y especificidad. El análisis de la hipermetilación del gen Gstp1 (Cairns P, et al. Clin Cancer Res 7: 2727-30 2001) en sedimento de orina de pacientes con cáncer de próstata mostró que solo 6 de 22 individuos con niveles elevados de metilación en tumores biopsiados mostraron hipermetilación correspondiente en muestras de orina.

Enfoque multifactorial. El diagnóstico del cáncer se ha basado tradicionalmente en la detección de marcadores moleculares únicos (por ejemplo, mutaciones genéticas, niveles elevados de PSA). Desafortunadamente, el cáncer es un estado de enfermedad en donde los marcadores únicos no han podido detectar o diferenciar muchas formas de la enfermedad. Por lo tanto, los ensayos que reconocen un solo marcador han demostrado tener un valor predictivo limitado. Un aspecto fundamental de esta divulgación es que el diagnóstico de cáncer basado en la metilación y la detección, diagnóstico y monitorización terapéutica de tales enfermedades proporcionarán mejoras significativas sobre el estado de la técnica que utiliza análisis de marcadores únicos mediante el uso de una selección de múltiples marcadores. La metodología analítica multiplexada es particularmente adecuada 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 citológica como clínicamente.

Necesidad pronunciada en la técnica. Por lo tanto, en vista de la incidencia de hiperplasia de próstata (50% de los hombres de 50 años o más, y 95% de los hombres de 75 años o más) y cáncer de próstata (180 por 100.000), existe una necesidad sustancial en la técnica para el desarrollo de marcadores moleculares que podrían usarse para proporcionar métodos sensibles, precisos y no invasivos (en oposición a, por ejemplo, biopsia y ecografía transrectal) para el cribado de poblaciones para proporcionar un diagnóstico precoz de trastornos proliferativos de células de próstata.

Resumen de la invención

La presente divulgación proporciona nuevos métodos y ácidos nucleicos para la detección y/o diferenciación entre trastornos proliferativos de células de próstata. De este modo, la divulgación aborda la larga necesidad de nuevos medios para el diagnóstico precoz de trastornos proliferativos de células de próstata, en particular para la detección de cáncer de próstata, carcinoma de próstata y neoplasias de próstata.

Se proporcionan en el presente documento secuencias genómicas y/o regiones reguladoras de las mismas de acuerdo con la SEQ ID NO: 31, siendo sus niveles de expresión indicativos de la presencia de trastornos proliferativos de células de próstata o características de los mismos. En particular, el estado de metilación de las posiciones CpG de los mismos es indicativo de la presencia de trastornos proliferativos de células de próstata o características de los mismos.

El método es para el diagnóstico de trastornos proliferativos de células de próstata, más preferiblemente cáncer de próstata, carcinoma de próstata y/o neoplasias.

El método es para el diagnóstico de cáncer de próstata, carcinoma de próstata y/o neoplasias y/o su diferenciación de la hiperplasia benigna de próstata.

En donde el método es para la diferenciación de una de próstata normal y/o HBP del cáncer de próstata, se prefiere particularmente que la secuencia o secuencias diana comprendan o hibriden en condiciones restrictivas con una o más secuencias que comprenden al menos 16 nucleótidos contiguos de las secuencias de acuerdo con la Tabla 4, comprendiendo dichos nucleótidos contiguos al menos una secuencia de dinucleótidos CpG, TpG o CpA.

En donde el método es para la diferenciación de uno de próstata normal, tejido normal de otros tejidos, cáncer de otros tejidos y/o BPH del cáncer de próstata, se prefiere particularmente que la(s) secuencia(s) diana comprendan o hibridan bajo condiciones restrictivas condiciones para, una o más secuencias que comprenden al menos 16 nucleótidos contiguos de las secuencias de acuerdo con la Tabla 5, comprendiendo dichos nucleótidos contiguos al menos una secuencia de dinucleótidos CpG, TpG o CpA.

En donde el método es para la diferenciación del cáncer de próstata de los cánceres de otros tejidos, se prefiere particularmente que la(s) secuencia(s) diana comprendan o hibriden en condiciones restrictivas con una o más secuencias que comprenden al menos 16 nucleótidos contiguos de las secuencias de acuerdo con la Tabla 6, comprendiendo dichos nucleótidos contiguos al menos una secuencia de dinucleótidos CpG, TpG o CpA.

Se prefiere particularmente que dicho trastorno proliferativo de células de próstata sea un cáncer de próstata, carcinoma de próstata o neoplasia de próstata. Además se proporcionan métodos y ácidos nucleicos para la diferenciación entre los tipos no cancerosos de tejido prostático (incluyendo HBP y normal) del carcinoma de próstata. Además se describen métodos y ácidos nucleicos para la diferenciación del cáncer de próstata del tejido prostático normal, tejidos que se originan de otros tejidos y BPH. También se describen métodos y ácidos nucleicos para la diferenciación del cáncer de próstata de cánceres que se originan en otros tejidos.

Para permitir este análisis, en este documento se proporciona un método para el análisis de muestras biológicas para la metilación genómica asociada con el desarrollo de trastornos proliferativos de células de próstata. Dicho método se caracteriza porque un ácido nucleico, o un fragmento del mismo, de acuerdo con la SEQ ID NO: 31 se pone en contacto con un reactivo o serie de reactivos capaces de distinguir entre dinucleótidos CpG metilados y no metilados dentro de la secuencia genómica, o secuencias de interés.

Se prefiere particularmente que el método y los ácidos nucleicos se utilicen para al menos uno de: detección de; estudio de poblaciones para; diferenciación entre; monitorización de; y detección y monitorización de trastornos proliferativos de células de próstata.

La presente divulgación proporciona un método para determinar los parámetros genéticos y/o epigenéticos del ADN genómico. El método tiene utilidad para la detección mejorada y/o diferenciación entre trastornos proliferativos de células de próstata. Aunque se conocen ensayos de metilación para la detección de cáncer de próstata, actualmente no existe un sistema de clasificación molecular para la detección de trastornos proliferativos de células de próstata, ni uno que diferencie con precisión las afecciones benignas de los carcinomas y neoplasmas de próstata.

La fuente puede ser cualquier fuente adecuada. Preferiblemente, la fuente de la muestra de ADN se selecciona del grupo que consiste en células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculación, orina, sangre y combinaciones de los mismos. Preferiblemente, la fuente son biopsias, fluidos corporales, eyaculación, orina o sangre.

Específicamente, en este documento se proporciona un método para la detección de trastornos proliferativos de células de próstata, que comprende: obtener una muestra biológica que comprende ácido(s) nucleico(s) genómico(s); poner en contacto el ácido o los ácidos nucleicos, o un fragmento del mismo, con un reactivo o una pluralidad de reactivos suficientes para distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas dentro de una secuencia diana del ácido nucleico sujeto, en donde la secuencia diana comprende o hibrida bajo condiciones restrictivas para, una secuencia que comprende al menos 16 nucleótidos contiguos de SEQ ID NO: 31, comprendiendo dichos nucleótidos contiguos al menos una secuencia de dinucleótidos CpG; y determinar, con base en al menos en parte en dicha distinción, el estado de metilación de al menos una secuencia de dinucleótidos CpG diana, o un promedio, o un valor que refleje un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG diana.

Preferiblemente, distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas dentro de la secuencia diana comprende la conversión dependiente del estado de metilación o la no conversión de al menos una de dichas secuencias de dinucleótidos CpG en la secuencia de dinucleótidos convertida o no convertida correspondiente dentro de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121, 238, 239 y regiones contiguas de las mismas correspondientes a la secuencia diana.

Aspectos adicionales proporcionan un método para la detección y/o diferenciación entre trastornos proliferativos de células de próstata, comprendiendo: obtener una muestra biológica que tiene ADN genómico sujeto; extraer el ADN genómico; tratar el ADN genómico, o un fragmento del mismo, con uno o más reactivos para convertir bases de citosina no metiladas en 5 posiciones en uracilo u otra base que sea detectablemente diferente a la citosina en términos de propiedades de hibridación; poner en contacto el ADN genómico tratado, o el fragmento tratado del mismo, con una enzima de amplificación y al menos dos cebadores que comprenden, en cada caso, una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a o hibrida en condiciones restrictivas o moderadamente restrictivas a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121, 238, 239 y sus complementos, en donde el ADN tratado o el fragmento del mismo se amplifica para producir un amplificado, o no se amplifica; y determinar, con base en la presencia o ausencia de, o en una propiedad de dicho amplificado, el estado de metilación de al menos una secuencia de dinucleótidos CpG de SEQ ID NO: 31, o un promedio, o un valor que

refleje un estado de metilación promedio de un pluralidad de secuencias de dinucleótidos CpG de los mismos. Preferiblemente, al menos una de tales moléculas de ácido nucleico hibridante o molécula de ácido nucleico peptídico está unida a una fase sólida. Preferiblemente, la determinación comprende el uso de al menos un método seleccionado del grupo que consiste en: hibridar al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria o hibrida en condiciones restrictivas o moderadamente restrictivas a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121, 238, 239 y sus complementos; 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 o hibrida en condiciones restrictivas o moderadamente restrictivas a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121, 238, 239 y sus complementos; hibridar al menos una molécula de ácido nucleico que comprende una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria o hibrida en condiciones restrictivas o moderadamente restrictivas con una secuencia seleccionada del grupo que consiste en las SEQ ID NO: 120, 121, 238, 239, y sus complementos, y extendiendo al menos una de tales moléculas de ácido nucleico hibridado en al menos una base de nucleótidos; y secuenciación del amplificado.

Otros aspectos proporcionan un método para la detección y/o diferenciación entre trastornos proliferativos de células de próstata, que comprende: obtener una muestra biológica que tiene ADN genómico sujeto; extraer el ADN genómico; en contacto con el ADN genómico, o un fragmento del mismo, que comprende una secuencia de acuerdo con la SEQ ID NO: 31 o una secuencia que se hibrida en condiciones restrictivas con el mismo, con una o más enzimas de restricción sensibles a la metilación, en donde el ADN genómico se digiere de ese modo para producir fragmentos de digestión, o no se digiere de ese modo; y determinar, con base en la presencia o ausencia de, o en la propiedad de al menos uno de dichos fragmentos, el estado de metilación de al menos una secuencia de dinucleótidos CpG de una SEQ ID NO: 31, o un valor promedio, o un valor que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG de los mismos. Preferiblemente, el ADN genómico digerido o no digerido se amplifica antes de dicha determinación.

Los aspectos adicionales proporcionan secuencias de ácido nucleico genómicas y químicamente modificadas novedosas, así como oligonucleótidos y/u oligómeros de PNA para el análisis de patrones de metilación de citosina dentro de la SEQ ID NO: 31.

La presente invención está definida por las reivindicaciones 1-14.

Breve descripción de los dibujos

Las Figuras 1 a 3 muestran matrices clasificadas producidas a partir de datos de secuenciación con bisulfito. La matriz global representa los datos de secuencia para un fragmento. Cada fila de la matriz representa un único sitio CpG dentro del fragmento y cada columna representa una muestra individual. La barra de la izquierda representa una escala del porcentaje de metilación, con el grado de metilación representado por la tonalidad de cada posición dentro de la columna, desde el negro que representa el 100% de metilación hasta el gris claro que representa el 0% de metilación. No hubo datos disponibles en las posiciones blancas.

La figura 1 muestra los datos de secuenciación de un fragmento del gen Receptor de prostaglandina E2, subtipo EP4 en donde las muestras secuenciadas son de carcinoma de próstata.

La Figura 2 muestra los datos de secuenciación de un fragmento del gen Receptor Nuclear Huérfano (Factor de Transcripción de Fetoproteína a-1 en donde las muestras secuenciadas son de carcinoma de próstata.

La Figura 3 muestra los datos de secuenciación de un fragmento del gen 1-Acil-SN-Glicerol-3-Fosfato Aciltransferasa Gamma en donde las muestras secuenciadas son de carcinoma de próstata.

La Figura 4 muestra las clasificaciones de marcadores de próstata normal y BPH frente a cáncer de próstata de acuerdo con el ejemplo 3. Cada región genómica de interés individual se representa como un punto. El gráfico de la izquierda proporciona valores p no corregidos del modelo de regresión logística gen a gen. Las líneas punteadas inferiores y superiores muestran límites de Bonferroni y FDR del 5% respectivamente. El eje X muestra los valores de p para las posiciones individuales de CpG. Los valores p son las probabilidades de que la distribución observada ocurra por casualidad en el conjunto de datos. La gráfica derecha proporciona precisión, sensibilidad y especificidad de un SVM lineal guiado en mediciones de metilación de todos los oligonucleótidos. La precisión de cada región genómica se representa como cuadrados negros, la especificidad como diamantes sin relleno, la sensibilidad como cuadrados sin relleno. La precisión medida en el eje X muestra la fracción de muestras clasificadas correctamente.

La Figura 5 muestra los mejores 12 marcadores para la diferenciación de próstata normal y HBP versus cáncer de próstata de acuerdo con el ejemplo 3. Las muestras normales de próstata y HBP se muestran a la izquierda. El cáncer de próstata está a la derecha. Cada columna representa una muestra; cada fila un oligonucleótido. Los oligonucleótidos se agrupan por marcador candidato. Los marcadores indicados se ordenan de arriba a abajo con mayor precisión. En el lado derecho de cada marcador, se enumeran los valores p corregidos de Bonferroni. Los datos de metilación se centran y normalizan a una desviación estándar para oligonucleótidos individuales. El color representa la distancia relativa del estado de metilación del oligonucleótido desde el valor medio. El color verde representa las

CpG hipometiladas dentro de un oligonucleótido, mientras que el rojo indica las CpG hipermetiladas dentro de un oligonucleótido.

La Figura 6 muestra las clasificaciones de los marcadores de próstata normal, BPH y otros tejidos versus cáncer de próstata de acuerdo con el ejemplo 3. Cada región genómica de interés individual se representa como un punto. La gráfica izquierda proporciona valores p no corregidos del modelo de regresión logística gen a gen. Las líneas punteadas inferiores y superiores muestran límites de Bonferroni y FDR del 5% respectivamente. El eje X muestra los valores de p para las posiciones individuales de CpG. Los valores p son las probabilidades de que la distribución observada ocurra por casualidad en el conjunto de datos. El gráfico de la derecha muestra la precisión, sensibilidad y especificidad de un SVM lineal trazado en mediciones de metilación de todos los oligonucleótidos. La precisión de cada región genómica se representa como cuadrados negros, la especificidad como diamantes sin relleno, la sensibilidad como cuadrados sin relleno. La precisión medida en el eje X muestra la fracción de muestras clasificadas correctamente.

La figura 7 muestra los mejores 12 marcadores para la diferenciación de próstata normal, HBP y otros tejidos frente al cáncer de próstata de acuerdo con el ejemplo 3. A la izquierda se muestran muestras de próstata normal, HBP y otros tejidos. Los "Otros tejidos" incluían tejido normal de otros órganos y cáncer de otros orígenes además de la próstata, según la tabla 7. El cáncer de próstata está a la derecha. Cada columna representa una muestra; cada fila un oligonucleótido. Los oligonucleótidos se agrupan por marcador candidato. Los marcadores indicados se ordenan de arriba a abajo con mayor precisión. En el lado derecho de cada marcador, se enumeran los valores p corregidos de Bonferroni. Los datos de metilación se centran y normalizan a una desviación estándar para oligonucleótidos individuales. El color representa la distancia relativa del estado de metilación del oligonucleótido desde el valor medio. El color verde representa las CpG hipometiladas dentro de un oligonucleótido, mientras que el rojo indica las CpG hipermetiladas dentro de un oligonucleótido.

La Figura 8 muestra las clasificaciones de marcadores de próstata normal, BPH y otros tejidos frente a cáncer de próstata de acuerdo con el ejemplo 3. Cada región genómica de interés individual se representa como un punto. El gráfico de la izquierda proporciona valores p no corregidos del modelo de regresión logística gen a gen. Las líneas punteadas inferiores y superiores muestran límites de Bonferroni y FDR del 5% respectivamente. El eje X muestra los valores de p para las posiciones individuales de CpG. Los valores p son las probabilidades de que la distribución observada ocurra por casualidad en el conjunto de datos. Los siguientes cánceres se muestran de izquierda a derecha: vejiga, melanoma, testículos, riñón, cáncer de endometrio, pulmón, mama, páncreas, hígado, ovario, glándula salival y próstata.

Descripción detallada de la invención

Definiciones:

Como se usa en el presente documento, el término expresión se tomará para significar la transcripción y traducción de un gen. El nivel de expresión de un gen puede determinarse mediante el análisis de cualquier factor asociado o indicativo del nivel de transcripción y traducción de un gen, incluidos, entre otros, el análisis de metilación, la pérdida de heterocigosidad (en lo sucesivo, también denominado LOH), niveles de expresión de ARN y niveles de expresión de proteínas.

Además, la actividad del gen transcrito puede verse afectada por variaciones genéticas tales como, pero sin limitación, mutaciones genéticas (incluidas, entre otras, SNP, mutaciones puntuales, eliminaciones, inserciones, longitud de repetición, reordenamientos y otros polimorfismos).

El término "Relación observada/esperada" ("Relación O/E") se refiere a la frecuencia de los dinucleótidos CpG dentro de una secuencia de ADN particular, y corresponde al [número de sitios CpG/(número de bases C x número de Bases G)].

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 correspondiente a una "Relación observada/esperada" >0.6, y (2) que tiene un "Contenido de GC" >0.5. Las islas CpG tienen típicamente, pero no siempre, entre aproximadamente 0.2 y aproximadamente 1 kb, o aproximadamente 2 kb de longitud.

El término "estado de metilación" o "estatus 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 particulares de metilación de CpG (cada uno con dos secuencias de dinucleótidos CpG CpG) dentro de una secuencia de ADN incluyen "no metilado", "totalmente metilado" y "hemimetilado".

El término "hemi-metilación" o "hemimetilación" se refiere al estado de metilación de un sitio de metilación de CpG palindrómico, donde solo una única citosina en una de las dos secuencias de dinucleótidos CpG del sitio de metilación de CpG palindrómico está metilada (por ejemplo, 5'-CC^MGG-3' (filamento superior): 3'-GGCC-5' (filamento inferior)).

El término 'AUC' como se usa en el presente documento es una abreviatura para el área bajo una curva. En particular, se refiere al área bajo una curva de características operativas del receptor (ROC). La curva ROC es un gráfico de la

- tasa positiva verdadera contra la tasa de falsos positivos para los diferentes puntos de corte posibles de una prueba de diagnóstico. Muestra la compensación entre sensibilidad y especificidad según el punto de corte seleccionado (cualquier aumento en la sensibilidad irá acompañado de una disminución en la especificidad). El área bajo una curva ROC (AUC) es una medida para la precisión de una prueba de diagnóstico (cuanto mayor sea el área, mejor, óptimo es 1, una prueba aleatoria tendría una curva ROC en la diagonal con un área de 0.5; para referencia: J.P. Egan. Signal Detection Theory and ROC Analysis, Academic Press, New York, 1975).
- El término "hipermetilación" se refiere al estado de metilación promedio correspondiente a una presencia aumentada de 5-mCyt en uno o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN de una muestra de ADN de prueba, en relación con la cantidad de 5-mCyt encontrada en los dinucleótidos CpG correspondientes dentro de una muestra de ADN de control normal.
- El término "hipometilación" se refiere al estado de metilación promedio correspondiente a una presencia disminuida de 5-mCyt en uno o una pluralidad de dinucleótidos CpG dentro de una secuencia de ADN de una muestra de ADN de prueba, en relación con la cantidad de 5-mCyt encontrada en los dinucleótidos CpG correspondientes dentro de una muestra de ADN de control normal.
- El término "microarreglos" se refiere en general a "microarreglos de ADN" y "chip(s) de ADN", como se reconoce en la técnica, abarca todos los soportes sólidos reconocidos en la técnica y abarca todos los métodos para fijar moléculas de ácido nucleico a los mismos o síntesis de ácidos nucleicos en estos.
- Los "parámetros genéticos" son mutaciones y polimorfismos de genes y secuencias adicionales necesarios para su regulación. Se designarán como mutaciones, en particular, inserciones, eliminaciones, mutaciones puntuales, inversiones y polimorfismos y, particularmente preferidos, SNP (polimorfismos de un solo nucleótido).
- Los "parámetros epigenéticos" son, en particular, metilaciones de citosina. Otros parámetros epigenéticos incluyen, por ejemplo, la acetilación de histonas que, sin embargo, no pueden analizarse directamente usando el método descrito pero que, a su vez, se correlacionan con la metilación del ADN.
- El término "reactivo de bisulfito" se refiere a un reactivo que comprende bisulfito, disulfito, sulfito de hidrógeno o combinaciones de los mismos, útil como se describe aquí para distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas.
- El término "ensayo de metilación" se refiere a cualquier ensayo para determinar el estado de metilación de una o más secuencias de dinucleótidos CpG dentro de una secuencia de ADN.
- El término "MS.AP-PCR" (Reacción en cadena de polimerasa preparada arbitrariamente sensible a la metilación) se refiere a la tecnología reconocida en la técnica que permite una exploración global del genoma usando cebadores ricos en CG para centrarse en las regiones más probable que contenga dinucleótidos CpG, y descrito por Gonzalgo et al., Cancer Research 57: 594-599, 1997.
- El término "MethyLight™" se refiere a la técnica de PCR en tiempo real basada en fluorescencia reconocida en la técnica descrita por Eads et al., Cancer Res. 59: 2302-2306, 1999.
- El término ensayo "HeavyMethyl™", en la realización del mismo implementado en el presente documento, se refiere a un ensayo, en donde las sondas de bloqueo específicas de metilación (también denominadas en este documento bloqueadores) que cubren posiciones CpG entre, o cubiertas por los cebadores de amplificación permiten la metilación. Amplificación selectiva específica de una muestra de ácido nucleico.
- El término ensayo "HeavyMethyl™ MethyLight™", en la realización del mismo implementado en el presente documento, se refiere a un ensayo HeavyMethyl™ MethyLight™, que es una variación del ensayo MethyLight™, en donde el ensayo MethyLight™ se combina con bloqueo específico de metilación sondas que cubren las posiciones CpG entre los cebadores de amplificación.
- El término "Ms-SNuPE" (Extensión de cebador de nucleótido único sensible a la metilación) se refiere al ensayo reconocido en la técnica descrito por Gonzalgo & Jones, Nucleic Acids Res. 25: 2529-2531, 1997.
- El término "MSP" (PCR específica de metilación) se refiere al ensayo de metilación reconocido en la técnica descrito por Herman et al. Proc. Natl. Acad. Sci. USA 93: 9821-9826, 1996, y por la Patente de los Estados Unidos No. 5,786,146.
- El término "COBRA" (Análisis combinado de restricción con bisulfito) se refiere al ensayo de metilación reconocido en la técnica descrito por Xiong & Laird, Nucleic Acids Res. 25: 2532-2534, 1997.
- El término "MCA" (amplificación de isla CpG metilada) se refiere al ensayo de metilación descrito por Toyota et al., Cancer Res. 59: 2307-12, 1999, y en el documento WO 00/26401A1.
- El término "hibridación" debe entenderse como un enlace de un oligonucleótido a una secuencia complementaria a lo largo de las líneas de los pares de bases de Watson-Crick en el ADN de la muestra, formando una estructura dúplex.

Las "condiciones de hibridación restrictivas", tal como se definen en el presente documento, implican hibridación a 68°C en 5x SSC/5x solución de Denhardt/SDS al 1.0%, y lavado en 0.2x SSC/SDS al 0.1% a temperatura ambiente, o implican la técnica equivalente reconocido del mismo (por ejemplo, condiciones en las que se lleva a cabo una hibridación a 60°C en 2.5 x regulador SSC, seguido de varios pasos de lavado a 37°C en una baja concentración de regulador, y permanece estable). Las condiciones moderadamente restrictivas, como se definen en el presente documento, implican el lavado en 3x SSC a 42°C, o el equivalente reconocido en la técnica del mismo. Los parámetros de concentración de sal y temperatura se pueden variar para lograr el nivel óptimo de identidad entre la sonda y el ácido nucleico diana. La orientación sobre tales condiciones está disponible en la técnica, por ejemplo, por Sambrook et al., 1989, *Molecular Cloning, A Laboratory Manual*, Cold Spring Harbor Press, N.Y.; and Ausubel et al. (eds.), 1995, *Current Protocols in Molecular Biology*, (John Wiley & Sons, N.Y.) at Unit 2.10.

Los términos "SEQ ID NO de arreglo", "SEQ ID NO de arreglo compuesta" o "secuencia de arreglo compuesta" se refieren a una secuencia, hipotética o de otro tipo, que consiste en una cabeza a cola (5' a 3') compuesto lineal de todas las secuencias contiguas individuales de un arreglo sujeto (por ejemplo, un compuesto de cabeza a cola de SEQ ID NO: x-y, en ese orden).

Los términos "nodo de SEQ ID NO de arreglo", "nodo de SEQ ID NO de arreglo compuesta" o "nodo de secuencia de arreglo compuesta" se refieren a una unión entre cualesquiera de dos secuencias contiguas individuales de la "SEQ ID NO de arreglo", la "SEQ ID NO de arreglo compuesta" o la "secuencia de arreglo compuesta".

En referencia a las secuencias de arreglo compuestas, la expresión "nucleótidos contiguos" se refiere a una región de secuencia contigua de cualquier secuencia contigua individual de los arreglos compuesta, pero no incluye una región de la secuencia de arreglo compuesta que incluye un "nodo". como se define aquí anteriormente.

Visión general:

La presente divulgación proporciona marcadores genéticos moleculares que tienen una utilidad novedosa para la detección y/o diferenciación entre trastornos proliferativos de células de próstata. Se prefiere particularmente que el método y los ácidos nucleicos de acuerdo con la invención se utilicen para al menos uno de: detección de; detección de poblaciones para; diferenciación entre; monitorización de; y detección y monitorización de trastornos proliferativos de células de próstata.

La modificación del ADN con bisulfito es una herramienta reconocida en la técnica utilizada para evaluar el estado de metilación de CpG. La 5-metilcitosina es la modificación de la base covalente más frecuente en el ADN de las células eucariotas. Desempeña un papel, por ejemplo, en la regulación de la transcripción, en la impresión genética y en la tumorigénesis. Por lo tanto, la identificación de 5-metilcitosina como componente de información genética es de considerable interés. Sin embargo, las posiciones de 5-metilcitosina no se pueden identificar por secuenciación, porque la 5-metilcitosina tiene el mismo comportamiento de emparejamiento de bases que la citosina. Además, la información epigenética transportada por la 5-metilcitosina se pierde por completo durante, por ejemplo, la amplificación por PCR.

El método utilizado con mayor frecuencia para analizar la presencia de 5-metilcitosina en el ADN se basa en la reacción específica de bisulfito con citosina, por lo que, tras la hidrólisis alcalina posterior, la citosina se convierte en uracilo que corresponde a la timina en su comportamiento de apareamiento de bases. Sin embargo, significativamente, la 5-metilcitosina permanece sin modificar en estas condiciones. En consecuencia, 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 como la única citosina restante utilizando técnicas estándar de biología molecular reconocidas en la técnica, por ejemplo, mediante amplificación e hibridación, o por secuenciación. Todas estas técnicas se basan en propiedades de emparejamiento de bases diferenciales, que ahora se pueden explotar por completo.

La técnica anterior, en términos de sensibilidad, se define mediante un método que comprende encerrar el ADN a analizar en una matriz de agarosa, evitando así la difusión y renaturalización del ADN (el bisulfito solo reacciona con el ADN monocatenario), y reemplazando todos los pasos de precipitación y purificación con diálisis rápida (Olek A, et al., *Un método modificado y mejorado para el análisis de metilación de citosina basado en bisulfito*, *Nucleic Acids Res.* 24: 5064-6, 1996). Por lo tanto, es posible analizar el estado de metilación de las células individuales, ilustrando la utilidad y la sensibilidad del método. Rein, T., et al., *Nucleic Acids Res.*, 26: 2255, 1998 proporcionan una visión general de los métodos reconocidos en la técnica para detectar 5-metilcitosina.

La técnica de bisulfito, salvo algunas excepciones (por ejemplo, Zeschnigk M, et al., *Eur J Hum Genet.* 5: 94-98, 1997), actualmente solo se usa en investigación. En todos los casos, los fragmentos cortos y específicos de un gen conocido se amplifican después de un tratamiento con bisulfito y se secuencian completamente (Olek and Walter, *Nat Genet.* 1997 17: 275-6, 1997), sometidos a una o más reacciones de extensión del cebador. (Gonzalzo & Jones, *Nucleic Acids Res.*, 25: 2529-31, 1997; WO 95/00669; Patente de los Estados Unidos No. 6,251,594) para analizar posiciones individuales de citosina, o tratadas por digestión enzimática (Xiong & Laird, *Nucleic Acids Res.*, 25: 2532-4, 1997). La detección por hibridación también se ha descrito en la técnica (Olek et al., WO 99/28498). Además, se ha descrito el uso de la técnica de bisulfito para la detección de metilación con respecto a genes individuales (Grigg and Clark,

Bioessays, 16: 431-6, 1994; Zeschnigk M, et al., Hum Mol Genet., 6: 387-95, 1997; Feil R, et al., Nucleic Acids Res., 22: 695-, 1994; Martin V, et al., Gene, 157: 261-4, 1995; WO 9746705 y WO 9515373).

En un aspecto, la presente divulgación 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 secuencias de dinucleótidos CpG dentro de la secuencia de SEQ ID NO: 31.

Esta determinación tiene utilidad para la detección y/o diferenciación entre trastornos proliferativos de células de próstata.

Procedimientos de ensayo de metilación. Se conocen diversos procedimientos de ensayo de metilación en la técnica, y se pueden usar junto con la presente divulgación. 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. Dichos ensayos implican, entre otras técnicas, la secuenciación de ADN de ADN tratado con bisulfito, PCR (para amplificación específica de secuencia), análisis de transferencia Southern y uso de enzimas de restricción sensibles a la metilación.

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

COBRA. El análisis COBRA es un ensayo cuantitativo de metilación útil para determinar los niveles de metilación del ADN en loci genéticos específicos en pequeñas cantidades de ADN genómico (Xiong & Laird, Nucleic Acids Res. 25: 2532-2534, 1997). En resumen, la digestión con enzimas de restricción se usa para revelar diferencias de secuencia dependientes de la metilación en productos de PCR de ADN tratado con bisulfito de sodio. Las diferencias de secuencia dependientes de la metilación se introducen primero en el ADN genómico mediante tratamiento con bisulfito estándar de acuerdo con el procedimiento descrito por Frommer et al. (Proc. Natl. Acad. Sci. USA 89: 1827-1831, 1992). Luego, la amplificación por PCR del ADN convertido con bisulfito se realiza usando cebadores específicos para las islas de interés CpG, seguido de digestión con endonucleasa de restricción, electroforesis en gel y detección usando sondas de hibridación marcadas específicas. Los niveles de metilación en la muestra de ADN original están representados por las cantidades relativas de producto de PCR digerido y no digerido de forma linealmente cuantitativa en un amplio espectro de niveles de metilación de ADN. Además, esta técnica se puede aplicar de manera confiable al ADN obtenido de muestras de tejido embebido en parafina microdisecionadas. Los reactivos típicos (por ejemplo, como se pueden encontrar en un kit típico basado en COBRA) para el análisis de COBRA pueden incluir, pero no se limitan a: cebadores de PCR para un gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); enzima de restricción y regulador apropiado; oligo de hibridación génica; control de hibridación oligo; kit de etiquetado de quinasa para sonda oligo; y nucleótidos marcados. Además, los reactivos de conversión con bisulfito pueden incluir: regulador de desnaturalización de ADN; regulador de sulfonación; reactivos o kits de recuperación de ADN (por ejemplo, precipitación, ultrafiltración, columna de afinidad); regulador de desulfonación; y componentes de recuperación de ADN.

Preferiblemente, ensayos como "MethyLight™" (una técnica de PCR en tiempo real basada en fluorescencia) (Eads et al., Cancer Res. 59: 2302-2306, 1999), Ms-SNuPE (nucleótido único sensible a la metilación Primer Extension) reacciones (Gonzalzo & Jones, Nucleic Acids Res. 25: 2529-2531, 1997), PCR específica de metilación ("MSP"; Herman et al., Proc. Natl. Acad. Sci. USA 93: 9821-9826, 1996; Patente de los Estados Unidos No. 5,786,146), y la amplificación de isla metílica CpG ("MCA"; Toyota et al., Cancer Res. 59: 2307-12, 1999) se usan solos o en combinación con otros de estos métodos.

MethyLight™. El ensayo MethyLight™ es un ensayo de metilación cuantitativa de alto rendimiento que utiliza tecnología de PCR en tiempo real basada en fluorescencia (TaqMan™) que no requiere más manipulaciones después del paso de PCR (Eads et al., Cancer Res. 59: 2302-2306, 1999). En resumen, el proceso MethyLight™ comienza con una muestra mixta de ADN genómico que se convierte, en una reacción de bisulfito de sodio, en un conjunto mixto de diferencias de secuencia dependientes de la metilación de acuerdo con los procedimientos estándar (el proceso de bisulfito convierte los residuos de citosina no metilados en uracilo). La PCR basada en fluorescencia se realiza en una reacción de PCR "imparcial" (con cebadores que no se superponen a los sitios de metilación de CpG conocidos), o en una reacción "sesgada" (con cebadores de PCR que se superponen a los dinucleótidos de CpG conocidos). La discriminación de secuencia puede ocurrir a nivel del proceso de amplificación o al nivel del proceso de detección de fluorescencia, o ambos.

El ensayo MethyLight™ puede usarse como una prueba cuantitativa para los patrones de metilación en la muestra de ADN genómico, en donde la discriminación de secuencia ocurre al nivel de hibridación de la sonda. En esta versión cuantitativa, la reacción de PCR proporciona una amplificación imparcial en presencia de una sonda fluorescente que se superpone a un sitio de metilación putativo particular. Un control imparcial de la cantidad de ADN de entrada es proporcionado por una reacción en la que ni los cebadores ni la sonda se superponen a ningún dinucleótido CpG. Alternativamente, se logra una prueba cualitativa para la metilación genómica mediante el sondeo del conjunto de

PCR sesgado con oligonucleótidos de control que no "cubren" los sitios de metilación conocidos (una versión basada en fluorescencia de la técnica "MSP"), o con oligonucleótidos que cubren los sitios de metilación potencial.

El proceso MethyLight™ puede usarse con una sonda "TaqMan®" en el proceso de amplificación. Por ejemplo, el ADN genómico bicatenario se trata con bisulfito de sodio y se somete a uno de dos conjuntos de reacciones de PCR usando sondas TaqMan®; por ejemplo, con cebadores sesgados y sonda TaqMan®, o cebadores no sesgados y sonda TaqMan®. La sonda TaqMan® tiene doble etiqueta con moléculas fluorescentes de "indicador" y "desactivador", y está diseñada para ser específica para una región de contenido de GC relativamente alta de modo que se derrita a una temperatura de aproximadamente 10°C más alta en el ciclo de PCR que el cebadores hacia adelante o hacia atrás. Esto permite que la sonda TaqMan® permanezca completamente hibridada durante el paso de fusión/extensión de PCR. A medida que la polimerasa Taq sintetiza enzimáticamente una nueva cadena durante la PCR, eventualmente alcanzará la sonda TaqMan® recocida. La actividad de endonucleasa de 5' a 3' de la polimerasa Taq desplazará la sonda TaqMan® al digerirla para liberar la molécula informadora fluorescente para la detección cuantitativa de su señal ahora no apagada utilizando un sistema de detección fluorescente en tiempo real.

Los reactivos típicos (por ejemplo, como se pueden encontrar en un kit típico basado en MethyLight™) para el análisis MethyLight™ pueden incluir, pero no se limitan a: cebadores de PCR para un gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); Sondas TaqMan®; reguladores de PCR optimizados y desoxinucleótidos; y polimerasa Taq.

La técnica Ms-SNuPE es un método cuantitativo para evaluar las diferencias de metilación en sitios CpG específicos basados en el tratamiento con bisulfito de ADN, seguido de la extensión del cebador de un solo nucleótido (Gonzalzo & Jones, Nucleic Acids Res. 25: 2529-2531, 1997). En resumen, el ADN genómico se hace reaccionar con bisulfito de sodio para convertir la citosina no metilada en uracilo, dejando la 5-metilcitosina sin cambios. La amplificación de la secuencia diana deseada se realiza luego usando cebadores de PCR específicos para ADN convertido con bisulfito, y el producto resultante se aísla y se usa como plantilla para el análisis de metilación en el sitio o sitios de CpG de interés. Se pueden analizar pequeñas cantidades de ADN (por ejemplo, secciones de patología microdisseccionadas), y evita la utilización de enzimas de restricción para determinar el estado de metilación en los sitios CpG.

Los reactivos típicos (por ejemplo, como se pueden encontrar en un kit típico basado en Ms-SNuPE) para el análisis de Ms-SNuPE pueden incluir, pero no se limitan a: cebadores de PCR para un gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); reguladores de PCR optimizados y desoxinucleótidos; kit de extracción en gel; cebadores de control positivo; cebadores Ms-SNuPE para un gen específico; regulador de reacción (para la reacción Ms-SNuPE); y nucleótidos marcados. Además, los reactivos de conversión con bisulfito pueden incluir: regulador de desnaturalización de ADN; regulador de sulfonación; reactivos o kit de recuperación de ADN (por ejemplo, precipitación, ultrafiltración, columna de afinidad); regulador de desulfonación; y componentes de recuperación de ADN.

MSP. La MSP (PCR específica de metilación) permite evaluar el estado de metilación de prácticamente cualquier grupo de sitios CpG dentro de una isla CpG, independientemente del uso de enzimas de restricción sensibles a la metilación (Herman et al. Proc. Natl. Acad. Sci. USA 93: 9821-9826, 1996; Patente de los Estados Unidos No. 5,786,146). En resumen, el ADN se modifica mediante bisulfito de sodio que convierte todas las citosinas no metiladas, pero no metiladas, en uracilo, y posteriormente se amplifica con cebadores específicos para ADN metilado versus no metilado. La MSP requiere solo pequeñas cantidades de ADN, es sensible a los alelos metilados al 0.1% de un locus de isla CpG dado y puede realizarse en ADN extraído de muestras incluidas en parafina. Los reactivos típicos (por ejemplo, como se pueden encontrar en un kit típico basado en MSP) para el análisis de MSP pueden incluir, pero no se limitan a: cebadores de PCR metilados y no metilados para un gen específico (o secuencia de ADN tratada con bisulfito o isla CpG), PCR optimizada reguladores y desoxinucleótidos, y sondas específicas.

MCA. La técnica MCA es un método que se puede usar para detectar patrones de metilación alterados en el ADN genómico y para aislar secuencias específicas asociadas con estos cambios (Toyota et al., Cancer Res. 59: 2307-12, 1999). En resumen, las enzimas de restricción con diferentes sensibilidades a la metilación de citosina en sus sitios de reconocimiento se usan para digerir ADN genómico de tumores primarios, líneas celulares y tejidos normales antes de la amplificación por PCR cebada arbitrariamente. Los fragmentos que muestran metilación diferencial se clonan y secuencian después de resolver los productos de PCR en geles de poli(acrilamida de alta resolución). Los fragmentos clonados se usan luego como sondas para el análisis Southern para confirmar la metilación diferencial de estas regiones. Los reactivos típicos (por ejemplo, como se pueden encontrar en un kit típico basado en MCA) para el análisis de MCA pueden incluir, entre otros: cebadores de PCR para cebado arbitrario de ADN genómico; reguladores y nucleótidos de PCR, enzimas de restricción y reguladores apropiados; oligos o sondas de hibridación génica; controlar la hibridación de oligos o sondas.

Se determinó que la secuencia genómica de acuerdo con la SEQ ID NO: 31, y las variantes tratadas de forma no natural de la misma de acuerdo con la SEQ ID NOs: 120, 121, 238, 239, tenían utilidad para la detección, clasificación y/o monitorización de trastornos proliferativos de células de próstata.

En una realización, la divulgación proporciona un método para la detección y/o diferenciación entre trastornos proliferativos de células de próstata en un sujeto. Dicho método comprende los siguientes pasos

i) determinar los niveles de expresión de las secuencias de genes de acuerdo con la Tabla 26 y/o regiones reguladoras de la misma

ii) determinar la presencia o ausencia de trastornos proliferativos de células de próstata de acuerdo con dicho nivel de expresión.

5 En una realización particularmente preferida, el nivel de expresión de la secuencia genómica de acuerdo con la Tabla 26 se determina por análisis del nivel de metilación de secuencias genómicas y/o regiones reguladoras de las mismas.

Dicho análisis comprende los siguientes pasos:

10 i) poner en contacto el ADN genómico obtenido del sujeto con al menos un reactivo, o una serie de reactivos que distingue entre dinucleótidos CpG metilados y no metilados dentro de al menos una región diana del ADN genómico, en donde dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótidos CpG,

ii) determinar la presencia o ausencia de trastornos proliferativos de células de próstata y/o una clasificación de los mismos de acuerdo con el estado de metilación de dichas regiones objetivo analizadas en i).

15 En donde el método es para la diferenciación de una de próstata normal y/o BPH del cáncer de próstata, se prefiere particularmente que la(s) secuencia(s) diana comprendan, o hibriden en condiciones restrictivas, una o más secuencias que comprendan al menos 16 nucleótidos contiguos de las secuencias de acuerdo con la Tabla 4, dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótidos CpG, TpG o CpA.

20 En donde el método es para la diferenciación de uno de próstata normal, tejido normal de otros tejidos, cáncer de otros tejidos y/o BPH del cáncer de próstata, se prefiere particularmente que la(s) secuencia(s) diana comprendan o hibriden bajo condiciones restrictivas condiciones para, una o más secuencias que comprenden al menos 16 nucleótidos contiguos de las secuencias de acuerdo con la Tabla 5, comprendiendo dichos nucleótidos contiguos al menos una secuencia de dinucleótidos CpG, TpG o CpA.

25 En donde el método es para la diferenciación del cáncer de próstata de los cánceres de otros tejidos, se prefiere particularmente que la(s) secuencia(s) diana comprendan o hibriden en condiciones restrictivas con una o más secuencias que comprenden al menos 16 nucleótidos contiguos de las secuencias de acuerdo con la Tabla 6, dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótidos CpG, TpG o CpA.

30 En el primer paso de dicho análisis de metilación, se aísla el ADN genómico por analizar. El ADN genómico puede aislarse por cualquier medio estándar en la técnica, incluido el uso de kits disponibles comercialmente. En resumen, en donde el ADN de interés está encapsulado en una membrana celular, la muestra biológica debe ser interrumpida y sometida a lisis por medios enzimáticos, químicos o mecánicos. La solución de ADN se puede limpiar de proteínas y otros contaminantes, por ejemplo, por digestión con proteinasa K. El ADN genómico se recupera de la solución. Esto puede llevarse a cabo mediante una variedad de métodos que incluyen la salazón, la extracción orgánica o la unión del ADN a un soporte de fase sólida. La elección del método se verá afectada por varios factores, incluidos el tiempo, los gastos y la cantidad requerida de ADN.

35 Preferiblemente, la fuente de la muestra de ADN se selecciona del grupo que consiste en células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculación, orina, sangre y combinaciones de los mismos. Preferiblemente, la fuente son biopsias, fluidos corporales, eyaculación, orina o sangre.

La muestra de ADN genómico se trata entonces de tal manera que las bases de citosina que no están metiladas en la posición 5' se conviertan en uracilo, timina u otra base que sea diferente a la citosina en términos de comportamiento de hibridación. Esto se entenderá como "tratamiento" en este documento.

40 El tratamiento del ADN genómico descrito anteriormente se lleva a cabo preferiblemente con bisulfito (sulfito de hidrógeno, disulfito) y posterior hidrólisis alcalina que da como resultado una conversión de nucleobases de citosina no metiladas en uracilo o en otra base que es diferente a la citosina en términos del comportamiento de emparejamiento de bases.

45 El ADN tratado se analiza luego para determinar el estado de metilación de una o más secuencias de genes diana (antes del tratamiento) asociadas con la presencia de trastornos proliferativos de células de próstata y/o indicativos de la presencia de subtipos de los mismos. Se prefiere particularmente que la región diana comprenda o hibride en condiciones restrictivas con al menos 16 nucleótidos contiguos de una secuencia genómica enumerada en la Tabla 26. Se prefiere además que se analice la secuencia de dichos genes en la Tabla 26 como se describe en el listado de secuencias adjunto.

50 El método de análisis puede seleccionarse de los conocidos en la técnica, incluidos los enumerados aquí. Particularmente preferidos son MethyLight, MSP y el uso de oligonucleótidos de bloqueo como se describirá aquí. Se prefiere además que cualquier oligonucleótido utilizado en dicho análisis (incluidos cebadores, oligonucleótidos de bloqueo y sondas de detección) sea complementario inverso, idéntico o hibridado en condiciones restrictivas o altamente restrictivas a un segmento largo de la base de al menos 16 pares de bases secuencias de una o más de

las SEQ ID NO: 120, 121, 238, 239 y secuencias complementarias a las mismas. Se prefiere además que cualquier oligonucleótido utilizado en dicho análisis (incluidos cebadores, oligonucleótidos de bloqueo y sondas de detección) sea complementario inverso, idéntico o hibridado en condiciones restrictivas o altamente restrictivas a un segmento largo de la base de al menos 16 pares de bases secuencias de una o más de SEQ ID NO: 120, 121, y secuencias complementarias a las mismas.

La metilación aberrante, más preferiblemente la hipermetilación de una secuencia genómica tomada de la Tabla 26, está asociada con la presencia de trastornos proliferativos de células de próstata. El análisis de la secuencia permite la detección y/o diferenciación entre trastornos proliferativos de células de próstata.

En un aspecto, el método describe el uso de una secuencia genómica de la Tabla 26 como marcadores para la detección y/o diferenciación entre trastornos proliferativos de células de próstata.

También se describe en el presente documento un kit para su uso en la detección y/o diferenciación entre un estado de trastorno proliferativo de células de próstata mediante el análisis de una muestra biológica.

En este documento se proporciona una nueva aplicación del análisis de los niveles y/o patrones de metilación dentro de dichas secuencias que permite la detección temprana y, por lo tanto, el tratamiento mejorado de los trastornos proliferativos de células de próstata, además, dicho método permite la diferenciación de afecciones benignas de afecciones tales como carcinoma de próstata y neoplasias que pueden ser potencialmente malignos y tener consecuencias de salud mucho más graves para el paciente. El tratamiento de los trastornos proliferativos de las células de la próstata se mejora mediante la detección temprana, y el método divulgado permite al médico y al paciente tomar decisiones de tratamiento mejores y más informadas.

Mejoras adicionales

La presente divulgación proporciona nuevos usos para las secuencias genómicas de SEQ ID NO: 31. Realizaciones adicionales proporcionan variantes modificadas de SEQ ID NO: 31, así como oligonucleótidos y/u oligómeros de PNA para el análisis de patrones de metilación de citosina dentro de SEQ ID NO: 31.

Un objetivo comprende el análisis del estado de metilación de uno o más dinucleótidos CpG dentro de la secuencia genómica de SEQ ID NO: 31 y secuencias complementarias a la misma.

En el presente documento se proporcionan ácidos nucleicos tratados, derivados de la SEQ ID NO: 31 genómica, en donde el tratamiento es adecuado para convertir al menos una base de citosina no metilada de la secuencia de ADN genómico en uracilo u otra base que sea detectablemente diferente a la citosina en términos de hibridación. Las secuencias genómicas en cuestión pueden comprender una, o más, posiciones CpG metiladas consecutivas o aleatorias. Dicho tratamiento comprende preferiblemente el uso de un reactivo seleccionado del grupo que consiste en bisulfito, sulfito de hidrógeno, disulfito y combinaciones de los mismos. En una realización preferida de la invención, el objetivo comprende el análisis de un ácido nucleico modificado de origen no natural que comprende una secuencia de al menos 16 bases de nucleótidos contiguas en longitud de una secuencia seleccionada del grupo que consiste en SEQ ID NOS: 120, 121, 238, 239, en donde dicha secuencia comprende al menos un dinucleótido CpG, TpA o CpA y secuencias complementarias a la misma. En una realización preferida adicional de la invención, el objetivo comprende el análisis de un ácido nucleico modificado que no se produce de forma natural que comprende una secuencia de al menos 16 bases de nucleótidos contiguas en longitud de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121, en donde dicha secuencia comprende al menos un dinucleótido CpG, TpA o CpA y secuencias complementarias a la misma. Las secuencias de SEQ ID NO: 120, 121, 238, 239 proporcionan versiones modificadas de forma no natural del ácido nucleico de acuerdo con SEQ ID NO: 31, en donde la modificación de cada secuencia genómica da como resultado 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 en sentido, por ejemplo, SEQ ID NO: 31, se describen cuatro versiones convertidas. Una primera versión en la que "C" se convierte en "T", pero "CpG" sigue siendo "CpG" (es decir, corresponde al caso en donde, para la secuencia genómica, todos los residuos "C" de las secuencias de dinucleótidos CpG están metilados y, por lo tanto, no están convertidos); una segunda versión revela el complemento de la secuencia de ADN genómico descrita (es decir, la cadena antisentido), en la que "C" se convierte en "T", pero "CpG" sigue siendo "CpG" (es decir, corresponde al caso donde, para toda "C" los residuos de las secuencias de dinucleótidos CpG se metilan y, por lo tanto, no se convierten).

Alternativamente, dicho análisis comprende el uso de un oligonucleótido u oligómero para detectar el estado de metilación de citosina dentro de ADN genómico o tratado (modificado de forma no natural), de acuerdo con la SEQ ID NO: 31, SEQ ID NO: 120, 121, 238, 239. En otra alternativa, dicho análisis comprende el uso de un oligonucleótido u oligómero para detectar el estado de metilación de la citosina dentro del ADN genómico o tratado (modificado químicamente), de acuerdo con la SEQ ID NOS: 120, 121.

Dicho oligonucleótido u oligómero comprende cada uno una secuencia de ácidos nucleicos que tiene una longitud de al menos nueve (9) nucleótidos que se hibrida, en condiciones restrictivas o moderadamente restrictivas (como se define anteriormente en el presente documento), a una secuencia de ácidos nucleicos tratada de acuerdo con la SEQ ID NO: 120, 121, 238, 239 y/o secuencias complementarias a la misma, o a una secuencia genómica de acuerdo con la SEQ ID NO: 31 y/o secuencias complementarias a la misma.

Se prefiere particularmente un oligonucleótido u oligómero que comprende una secuencia de ácidos nucleicos que tiene una longitud de al menos nueve (9) nucleótidos que hibrida, en condiciones restrictivas o moderadamente restrictivas (como se define aquí anteriormente), a una secuencia de ácidos nucleicos tratada de acuerdo con SEQ ID NOS: 120, 121, y/o secuencias complementarias a la misma.

- 5 Por lo tanto, la presente divulgación incluye moléculas de ácido nucleico (por ejemplo, oligonucleótidos y moléculas de ácido nucleico peptídico (PNA) (oligómeros de PNA)) que hibridan en condiciones de hibridación moderadamente restrictivas y/o restrictivas con todas o una parte de las secuencias SEQ ID NO: 31, SEQ ID NOS: 120, 121, 238, 239, o sus complementos. La porción de hibridación de los ácidos nucleicos de hibridación tiene típicamente al menos 9, 15, 20, 25, 30 o 35 nucleótidos de longitud. Sin embargo, las moléculas más largas tienen utilidad y, por lo tanto, están dentro del alcance de la presente divulgación.

Preferiblemente, la porción de hibridación de los ácidos nucleicos de hibridación de la invención es al menos 95%, o al menos 98%, o 100% idéntica a la secuencia, o a una porción de la misma de SEQ ID NOS: 120, 121, 238, 239 o sus complementos.

- 15 Los ácidos nucleicos de hibridación del tipo descrito en el presente documento pueden usarse, por ejemplo, como un cebador (por ejemplo, un cebador de PCR), o una sonda o cebador de diagnóstico y/o clasificación. Preferiblemente, la hibridación de la sonda de oligonucleótidos a una muestra de ácido nucleico se realiza en condiciones restrictivas y la sonda es 100% idéntica a la secuencia diana. La estabilidad del dúplex o híbrido del ácido nucleico se expresa como la temperatura de fusión o T_m , que es la temperatura a la cual una sonda se disocia de un ADN diana. Esta temperatura de fusión se usa para definir las condiciones de rigurosidad requeridas.

- 20 Para secuencias diana que están relacionadas y son sustancialmente idénticas a la secuencia correspondiente de SEQ ID NO: 31 (tales como variantes alélicas y SNP), en lugar de idénticas, es útil establecer primero la temperatura más baja a la que solo se produce la hibridación homóloga con una concentración particular de sal (por ejemplo, SSC o SSPE). Luego, suponiendo que el 1% de falta de coincidencia da como resultado una disminución de 1°C en la T_m , la temperatura del lavado final en la reacción de hibridación se reduce en consecuencia (por ejemplo, si se buscan secuencias que tienen >95% de identidad con la sonda, la temperatura final la temperatura de lavado disminuye en 5°C). En la práctica, el cambio en T_m puede estar entre 0.5°C y 1.5°C por 1% de desajuste.

- 25 Ejemplos de oligonucleótidos de longitud X (en nucleótidos), como se indica por las posiciones de polinucleótidos con referencia a, por ejemplo, SEQ ID NO: x, incluyen los correspondientes a conjuntos (conjuntos en sentido y antisentido) de oligonucleótidos superpuestos consecutivamente de longitud X, donde los oligonucleótidos dentro de cada conjunto de superposición consecutiva (correspondiente a un valor X dado) se definen como el conjunto finito de oligonucleótidos Z de las posiciones de nucleótidos:

$n \text{ a } (n + (X-1));$

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

donde Y es igual a la longitud (nucleótidos o pares de bases) de SEQ ID NO: x (2299);

- 35 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 meros consecutivamente superpuestos); y

donde el número (Z) de oligómeros superpuestos consecutivamente de longitud X para una SEQ ID NO dada de longitud Y es igual a $Y - (X-1)$. Por ejemplo, $Z = 2299 - 19 = 2280$ para conjuntos en sentido o antisentido de SEQ ID NO: x, donde X=20.

- 40 Preferiblemente, el conjunto está limitado a aquellos oligómeros que comprenden al menos un dinucleótido CpG, TpG o CpA.

La presente divulgación abarca, para la SEQ ID NO: 31 (en sentido y antisentido), conjuntos de oligonucleótidos superpuestos consecutivamente múltiples u oligonucleótidos modificados de longitud X, donde, por ejemplo, X= 9, 10, 17, 20, 22, 23, 25, 27, 30 o 35 nucleótidos.

- 45 Los oligonucleótidos u oligómeros descritos en el presente documento constituyen herramientas eficaces útiles para determinar los parámetros genéticos y epigenéticos de la secuencia genómica correspondiente a la SEQ ID NO: 31. Preferiblemente, dichos oligómeros comprenden al menos un dinucleótido CpG, TpG o CpA.

- 50 Los oligonucleótidos u oligómeros particularmente preferidos son aquellos en los que la citosina del dinucleótido CpG (o de las correspondientes secuencias convertidas de TpG o CpA dinucleótido) está dentro del tercio medio del oligonucleótido; es decir, donde el oligonucleótido tiene, por ejemplo, 13 bases de longitud, el dinucleótido CpG, TpG o CpA se coloca dentro del quinto al noveno nucleótido desde el extremo 5'.

Los oligonucleótidos también pueden modificarse uniando químicamente el oligonucleótido a una o más unidades estructurales o conjugados para potenciar la actividad, la estabilidad o la detección del oligonucleótido. Dichas unidades estructurales o conjugados incluyen cromóforos, fluoróforos, lípidos tales como colesterol, ácido cólico,

- 5 tioéter, cadenas alifáticas, fosfolípidos, poliaminas, polietilenglicol (PEG), unidades estructurales de palmitoilo y otros como se describe en, por ejemplo, las Patentes de los Estados Unidos Nos. 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 también pueden existir en forma de un PNA (ácido nucleico peptídico) que tiene propiedades de emparejamiento particularmente preferidas. Por lo tanto, el oligonucleótido puede incluir otros grupos adjuntos tales como péptidos, y puede incluir agentes de escisión activados por hibridación (Krol et al., BioTechniques 6: 958-976, 1988) o agentes intercalantes (Zon, Pharm. Res. 5: 539- 549, 1988). Para este fin, el oligonucleótido puede conjugarse con otra molécula, por ejemplo, un cromóforo, fluoróforo, péptido, agente de reticulación activado por hibridación, agente de transporte, agente de escisión activado por hibridación, etc.
- 10 El oligonucleótido también puede comprender al menos una unidad estructural de azúcar y/o base modificada reconocida en la técnica, o puede comprender una cadena principal modificada o un enlace internucleósido no natural.
- 15 Los oligonucleótidos u oligómeros 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: 31 y secuencias complementarias a los mismos, o a los correspondientes CpG, TpG o dinucleótido CpA dentro de una secuencia de los ácidos nucleicos tratados de acuerdo con SEQ ID NO: 120, 121, 238, 239 y secuencias complementarias a los mismos. Sin embargo, se anticipa que para 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 en consecuencia.
- 20 Por lo tanto, se proporciona un conjunto de al menos dos (2) oligonucleótidos y/u oligómeros de PNA) útiles para detectar el estado de metilación de citosina en el ADN genómico tratado (SEQ ID NO: 120, 121, 238, 239), o en ADN genómico (SEQ ID NO: 31 y secuencias complementarias a la misma). Estas sondas permiten el diagnóstico y/o clasificación de parámetros genéticos y epigenéticos de los trastornos proliferativos de células de próstata. El conjunto de oligómeros también se puede usar para detectar polimorfismos de un solo nucleótido (SNP) en el ADN genómico tratado (SEQ ID NO: 120, 121, 238, 239), o en el ADN genómico (SEQ ID NO: 31 y secuencias complementarias al mismo).
- 25 Se prefiere particularmente un conjunto de al menos dos (2) oligonucleótidos y/u oligómeros de PNA) útiles para detectar el estado de metilación de citosina en el ADN genómico tratado de acuerdo con la SEQ ID NO: 120, 121.
- Al menos uno, y más preferiblemente todos los miembros de un conjunto de oligonucleótidos están unidos a una fase sólida.
- 30 También se proporciona un conjunto de al menos dos (2) oligonucleótidos que se usan como oligonucleótidos 'cebadores' para amplificar secuencias de ADN de una de las SEQ ID NO: 31, SEQ ID NO 120, 121, 238, 239 y secuencias complementarias a las mismas, o segmentos de las mismas.
- 35 Se prefiere particularmente un conjunto de al menos dos (2) oligonucleótidos que se usan como oligonucleótidos 'cebadores' para amplificar secuencias de ADN de una de las SEQ ID NO: 120, 121. Se anticipa que los oligonucleótidos pueden constituir todo o parte de un "arreglo" o "chip de ADN" (es decir, una disposición de diferentes oligonucleótidos y/u oligómeros de PNA unidos a una fase sólida). Tal conjunto de diferentes secuencias de oligonucleótidos y/u oligómeros de PNA puede caracterizarse, por ejemplo, en que está dispuesto en la fase sólida en forma de una red rectangular o hexagonal. La superficie en fase sólida puede estar compuesta de silicio, vidrio, poliestireno, aluminio, acero, hierro, cobre, níquel, plata u oro. También se pueden usar nitrocelulosa, así como plásticos como el nylon, que pueden existir en forma de gránulos o también como matrices de resina. Puede obtenerse una visión general de la técnica anterior en la fabricación de conjuntos de oligómeros de una edición especial de Nature Genetics (Nature Genetics Supplement, Volume 21, January 1999, y de la literatura allí citada). Las sondas marcadas con fluorescencia a menudo se usan para escanear matrices de ADN inmovilizadas. La unión simple de los tintes Cy3 y Cy5 al 5'-OH de la sonda específica es particularmente adecuada para marcadores de fluorescencia. La detección de la fluorescencia de las sondas hibridadas puede llevarse a cabo, por ejemplo, a través de un microscopio confocal. Los tintes Cy3 y Cy5, además de muchos otros, están disponibles comercialmente.
- 40 También se anticipa que los oligonucleótidos, o secuencias particulares de los mismos, pueden constituir todo o parte de un "arreglo virtual" en la que los oligonucleótidos, o secuencias particulares de los mismos, se usan, 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, por ejemplo, se describe en US 2003/0013091 (número de serie de los Estados Unidos 09/898,743, publicado el 16 de enero de 2003). En tales métodos, se generan suficientes etiquetas para que cada ácido nucleico en la mezcla compleja (es decir, cada analito) pueda unirse de manera única por una etiqueta única y así detectarse (cada etiqueta se cuenta directamente, lo que resulta en una lectura digital de cada especie moleculares en la mezcla).
- 50 Se prefiere particularmente que los oligómeros se utilicen para al menos uno de: detección de; detección de poblaciones para; diferenciación entre; monitorización de; y detección y monitorización de trastornos proliferativos de células de próstata. Esto se habilita mediante el uso de dichos conjuntos para la detección y/o diferenciación entre trastornos proliferativos de células de próstata en una muestra biológica aislada de un paciente. Se prefieren
- 55

particularmente aquellos conjuntos de oligómeros que comprenden al menos dos oligonucleótidos seleccionados de uno de los siguientes conjuntos de oligonucleótidos.

En un aspecto del método, esto se logra mediante el análisis del estado de metilación de al menos una secuencia diana que comprende o hibrida en condiciones restrictivas con al menos 16 nucleótidos contiguos de un gen o secuencia de acuerdo con la Tabla 26 y sus complementos.

Se proporciona además un método para determinar parámetros genéticos y/o epigenéticos de la secuencia genómica de acuerdo con la SEQ ID NO: 31 dentro de un sujeto mediante el análisis de metilación de citosina y polimorfismos de un solo nucleótido. Dicho método comprende poner en contacto un ácido nucleico que comprende la SEQ ID NO: 31 en una muestra biológica obtenida de dicho sujeto con al menos un reactivo o una serie de reactivos, en donde dicho reactivo o serie de reactivos distingue entre dinucleótidos CpG metilados y no metilados dentro del ácido nucleico diana.

Preferiblemente, dicho método comprende los siguientes pasos: En el primer paso, se obtiene una muestra del tejido por analizar. Preferiblemente, la fuente de la muestra de ADN se selecciona del grupo que consiste en células o líneas celulares, portaobjetos histológicos, biopsias, tejido embebido en parafina, fluidos corporales, eyaculación, orina, sangre y combinaciones de los mismos. Preferiblemente, la fuente son biopsias, fluidos corporales, eyaculación, orina o sangre.

El ADN genómico se aísla luego de la muestra. El ADN genómico puede aislarse por cualquier medio estándar en la técnica, incluido el uso de kits disponibles comercialmente. En resumen, en donde el ADN de interés está encapsulado en una membrana celular, la muestra biológica debe ser interrumpida y sometida a lisis por medios enzimáticos, químicos o mecánicos. La solución de ADN se puede eliminar de proteínas y otros contaminantes, por ejemplo, por digestión con proteinasa K. El ADN genómico se recupera de la solución. Esto puede llevarse a cabo mediante una variedad de métodos que incluyen la extracción con solución salina, la extracción orgánica o la unión del ADN a un soporte de fase sólida. La elección del método se verá afectada por varios factores, incluidos el tiempo, los gastos y la cantidad requerida de ADN.

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

En el segundo paso del método, la muestra de ADN genómico se trata de tal manera que las bases de citosina que no están metiladas en la posición 5' se convierten en uracilo, timina u otra base que sea diferente a la citosina en términos del comportamiento de hibridación. Esto se entenderá como "pretratamiento" o "tratamiento" en este documento.

El tratamiento de ADN genómico descrito anteriormente se lleva a cabo preferiblemente con bisulfito (sulfito, disulfito de hidrógeno) y posterior hidrólisis alcalina que da como resultado una conversión de nucleobases de citosina no metiladas en uracilo o en otra base que sea diferente a la citosina en términos de comportamiento de emparejamiento de bases.

En el tercer paso del método, los fragmentos del ADN tratado se amplifican, usando conjuntos de oligonucleótidos cebadores como se describe en el presente documento, y una enzima de amplificación. La amplificación de varios segmentos de ADN puede llevarse a cabo simultáneamente en un mismo recipiente de reacción. Típicamente, la amplificación se lleva a cabo usando una reacción en cadena de la polimerasa (PCR). El conjunto de oligonucleótidos cebadores incluye al menos dos oligonucleótidos cuyas secuencias son cada una inversamente complementaria, idéntica o hibridada en condiciones restrictivas o altamente restrictivas con un segmento largo de al menos 16 pares de bases de las secuencias de bases de una de las SEQ ID NO: 120, 121, 238, 239 y secuencias complementarias a los mismos.

Particularmente preferido es un conjunto de oligonucleótidos cebadores que incluye al menos dos oligonucleótidos cuyas secuencias son cada una inversamente complementaria, idéntica o hibridada en condiciones restrictivas o altamente restrictivas con un segmento largo de al menos 16 pares de bases de las secuencias de bases de uno de SEQ ID NO: 120, 121 y secuencias complementarias a la misma.

En un aspecto alternativo del método, el estado de metilación de las posiciones CpG preseleccionadas dentro de las secuencias de ácido nucleico que comprenden SEQ ID NO: 31 puede detectarse mediante el uso de oligonucleótidos cebadores específicos de metilación. Esta técnica (MSP) se ha descrito en la Patente de los Estados Unidos No. 6,265,171 de Herman. El uso de cebadores 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 cebadores de MSP contienen al menos un cebador que se hibrida con un dinucleótido CpG tratado con bisulfito. Por lo tanto, la secuencia de dichos cebadores comprende al menos un dinucleótido CpG. Los cebadores de MSP específicos para ADN no metilado contienen una "T" en la posición de la posición C en el CpG. Preferiblemente, por lo tanto, se requiere que la secuencia de bases de dichos cebadores comprenda una secuencia que tenga una longitud de al menos 9 nucleótidos que hibrida a una secuencia de ácidos nucleicos tratada de acuerdo con una de las SEQ ID NO: 120, 121, 238, 239 y secuencias complementarias a la misma, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG.

Preferiblemente, se requiere que la secuencia de bases de dichos cebadores comprenda una secuencia que tenga una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácidos nucleicos tratada de acuerdo con una de las SEQ ID NO: 120, 121 y secuencias complementarias a la misma, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG.

- 5 Un aspecto preferido adicional del método comprende el uso de oligonucleótidos bloqueadores. El uso de tales oligonucleótidos bloqueadores ha sido descrito por Yu et al., *BioTechniques* 23: 714-720, 1997. Los oligonucleótidos de la sonda de bloqueo se hibridan con el ácido nucleico tratado con bisulfito simultáneamente con los cebadores de PCR. La amplificación por PCR del ácido nucleico se termina en la posición 5' de la sonda de bloqueo, de modo que se suprime la amplificación de un ácido nucleico donde está presente la secuencia complementaria a la sonda de
10 bloqueo. Las sondas pueden diseñarse para hibridar con el ácido nucleico tratado con bisulfito de una manera específica del 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 están metilados en la posición en cuestión se llevaría a cabo mediante el uso de sondas de bloqueo que comprenden un 'CpA' o 'TpA' en la posición en cuestión, en oposición a un 'CpG' si se desea la supresión de la amplificación de ácidos
15 nucleicos metilados.

Para los métodos de PCR que usan oligonucleótidos bloqueadores, la interrupción eficiente de la amplificación mediada por polimerasa requiere que los oligonucleótidos bloqueadores no sean alargados por la polimerasa. Preferiblemente, esto se logra mediante el uso de bloqueadores que son 3'-desoxioligonucleótidos u oligonucleótidos derivados en la posición 3' con un grupo hidroxilo "libre". Por ejemplo, los oligonucleótidos de 3'-O-acetilo son representativos de una clase preferida de molécula bloqueadora.
20

Además, la descomposición mediada por polimerasa de los oligonucleótidos bloqueadores debería descartarse. Preferiblemente, dicha preclusión comprende el uso de una polimerasa que carece de actividad exonucleasa 5'-3', o el uso de oligonucleótidos bloqueadores modificados que tienen, por ejemplo, puentes de tioato en los extremos 5' del mismo que hacen que la molécula bloqueadora sea resistente a la nucleasa. Las aplicaciones particulares pueden no
25 requerir tales modificaciones 5' del bloqueador. Por ejemplo, si los sitios de unión del bloqueador y el cebador se superponen, evitando así la unión del cebador (por ejemplo, con un bloqueador en exceso), se impedirá sustancialmente la degradación del oligonucleótido bloqueador. Esto se debe a que la polimerasa no extenderá el cebador hacia y a través (en la dirección 5'-3') del bloqueador, un proceso que normalmente da como resultado la degradación del oligonucleótido bloqueador hibridado.

- 30 Un bloqueador/PCR particularmente preferido comprende el uso de oligómeros de ácido nucleico peptídico (PNA) como oligonucleótidos de bloqueo. Dichos oligómeros bloqueadores de PNA son ideales, porque la polimerasa no los descompone ni los extiende.

Preferiblemente, por lo tanto, se requiere que la secuencia de bases de dichos oligonucleótidos de bloqueo comprenda una secuencia que tenga una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácidos nucleicos tratada de acuerdo con una de las SEQ ID NO: 120, 121, 238, 239 y secuencias complementarias a las mismas, en donde la secuencia de bases de dichos oligonucleótidos comprende al menos un dinucleótido CpG, TpG o CpA.
35

Se prefiere además que la secuencia de bases de dichos oligonucleótidos de bloqueo se requiera que comprenda una secuencia que tenga una longitud de al menos 9 nucleótidos que hibrida con una secuencia de ácidos nucleicos tratada de acuerdo con una de las SEQ ID NO: 120, 121, y secuencias complementarias a las mismas, en donde la secuencia de bases de dichos oligonucleótidos comprende al menos un dinucleótido CpG, TpG o CpA.
40

Los fragmentos obtenidos por medio de la amplificación pueden llevar un marcador detectable directa o indirectamente. Se prefieren los marcadores en forma de marcadores de fluorescencia, radionucleidos o fragmentos de moléculas desmontables que tienen una masa típica que se puede detectar en un espectrómetro de masas. Cuando dichas etiquetas son etiquetas de masa, se prefiere que los amplificados marcados tengan una sola carga neta positiva o negativa, lo que permite una mejor detectabilidad en el espectrómetro de masas. La detección puede llevarse a cabo y visualizarse mediante, por ejemplo, espectrometría de masas por ionización/desorción asistida por matriz (MALDI) o utilizando espectrometría de masas por pulverización electrónica (ESI).
45

La espectrometría de masas por ionización/desorción asistida por matriz (MALDI-TOF) es un desarrollo muy eficiente para el análisis de biomoléculas (Karas and Hillenkamp, *Anal Chem.*, 60: 2299-301, 1988). Un analito está incrustado en una matriz absorbente de luz. La matriz se evapora mediante un pulso láser corto que transporta la molécula de analito a la fase de vapor de manera no fragmentada. El analito se ioniza por colisiones con moléculas de matriz. Un voltaje aplicado acelera los iones en un tubo de vuelo libre de campo. Debido a sus diferentes masas, los iones se aceleran a diferentes tasas. Los iones más pequeños llegan al detector antes que los más grandes. La espectrometría MALDI-TOF es muy adecuada para el análisis de péptidos y proteínas. El análisis de los ácidos nucleicos es algo más difícil (Gut & Beck, *Current Innovations and Future Trends*, 1: 147-57, 1995). La sensibilidad con respecto al análisis de ácido nucleico es aproximadamente 100 veces menor que para los péptidos, y disminuye de manera desproporcionada al aumentar el tamaño del fragmento. Además, para los ácidos nucleicos que tienen una cadena principal con carga negativa múltiple, el proceso de ionización a través de la matriz es considerablemente menos eficiente. En la espectrometría MALDI-TOF, la selección de la matriz juega un papel eminentemente importante. Para
50
55

la desorción de péptidos, se han encontrado varias matrices muy eficientes que producen una cristalización muy fina. Ahora hay varias matrices sensibles para el ADN, sin embargo, la diferencia en sensibilidad entre péptidos y ácidos nucleicos no se ha reducido. Sin embargo, esta diferencia en la sensibilidad puede reducirse modificando químicamente el ADN de tal manera que se vuelva más similar a un péptido. Por ejemplo, los ácidos nucleicos de fosforotioato, en los que los fosfatos habituales de la columna vertebral se sustituyen con tiofosfatos, se pueden convertir en un ADN de carga neutra utilizando química de alquilación simple (Gut & Beck, *Nucleic Acids Res.* 23: 1367-73, 1995). El acoplamiento de una etiqueta de carga a este ADN modificado da como resultado un aumento en la sensibilidad MALDI-TOF al mismo nivel que el encontrado para los péptidos. Una ventaja adicional del etiquetado de carga es la mayor estabilidad del análisis contra impurezas, lo que hace que la detección de sustratos no modificados sea considerablemente más difícil.

En el cuarto paso del método, los amplificados obtenidos durante el tercer paso del método se analizan para determinar el estado de metilación de los dinucleótidos CpG antes del tratamiento.

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

Los amplificadores obtenidos por medio de PCR estándar y de metilación específica pueden analizarse adicionalmente por medio de métodos basados en hibridación tales como, pero sin limitación, tecnología de arreglos y tecnologías basadas en sondas, así como mediante técnicas tales como secuenciación y extensión dirigida por plantilla.

En un aspecto del método, los amplificados sintetizados en el paso tres se hibridan posteriormente a un arreglo 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 utilizadas durante la hibridación se compone preferiblemente de al menos 2 oligonucleótidos u oligómeros de PNA; en el proceso, los amplificados sirven como sondas que hibridan con oligonucleótidos previamente unidos a una fase sólida; los fragmentos no hibridados se eliminan posteriormente; dichos oligonucleótidos contienen al menos una secuencia de bases que tiene una longitud de al menos 9 nucleótidos que es inversamente complementaria o idéntica a un segmento de las secuencias de bases especificadas en el presente Listado de Secuencias; y el segmento comprende al menos un dinucleótido CpG, TpG o CpA.

Preferiblemente, dicho dinucleótido está presente en el tercio central del oligómero. Por ejemplo, en donde el oligómero comprende un dinucleótido CpG, dicho dinucleótido es preferiblemente del quinto al noveno nucleótido desde el extremo 5' de un 13-mero. Existe un oligonucleótido para el análisis de cada dinucleótido CpG dentro de la secuencia de acuerdo con la SEQ ID NO: 31, y las posiciones equivalentes dentro de la SEQ ID NO: 120, 121, 238, 239.

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

En otro aspecto adicional del método, el estado de metilación genómica de las posiciones CpG se puede determinar por medio de sondas de oligonucleótidos que se hibridan con el ADN tratado con bisulfito simultáneamente con los cebadores de amplificación por PCR (en donde dichos cebadores pueden ser metilación específica o estándar).

Un aspecto particularmente preferido de este método es el uso de PCR cuantitativa en tiempo real basada en fluorescencia (Heid et al., *Genome Res.* 6: 986-994, 1996; véase también la Patente de los Estados Unidos No. 6,331,393) que emplea una doble sonda oligonucleotídica fluorescente marcada (PCR TaqMan™, utilizando 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 extensible, llamado sonda TaqMan™, que, en realizaciones preferidas, está diseñado para hibridarse con una secuencia rica en GpC ubicada entre los cebadores de amplificación directa e inversa. La sonda TaqMan™ comprende además una "unidad estructural indicador" fluorescente y una "unidad estructural desactivadora" unido covalentemente a unidades estructurales conectores (por ejemplo, fosforamidas) unidos a los nucleótidos del oligonucleótido TaqMan™. Para el análisis de la metilación dentro de los ácidos nucleicos posteriores al tratamiento con bisulfito, se requiere que la sonda sea específica de metilación, como se describe en la Patente de los Estados Unidos No. 6,331,393, también conocida como el ensayo MethylLight™. Las variaciones en la metodología de detección TaqMan™ que también son adecuadas para usar con la invención descrita incluyen el uso de tecnología de doble sonda (Lightcycler™) o cebadores de amplificación fluorescentes (tecnología Sunrise™). Ambas técnicas pueden adaptarse de una manera adecuada para su uso con ADN tratado con bisulfito y, además, para el análisis de metilación dentro de los dinucleótidos CpG.

Un método adicional adecuado para el uso de oligonucleótidos de sonda para la evaluación de la metilación mediante análisis de ácidos nucleicos tratados con bisulfito. En otro aspecto preferido del método, el cuarto paso del método comprende el uso de la extensión de oligonucleótidos dirigida por plantilla, tales como MS-SNuPE según lo descrito por Gonzalgo & Jones, *Nucleic Acids Res.* 25: 2529-2531, 1997.

En otro aspecto adicional del método, el cuarto paso del método comprende secuenciación y posterior análisis de secuencia del amplificado generado en el tercer paso del método (Sanger F., et al., Proc Natl Acad Sci USA 74: 5463-5467, 1977).

Mejor modo

5 En el aspecto más preferido del método, los ácidos nucleicos genómicos se aíslan y se tratan de acuerdo con los primeros tres pasos del método descrito anteriormente, a saber:

a) obtener, de un sujeto, una muestra biológica que tiene ADN genómico sujeto;

b) extraer o aislar de otro modo el ADN genómico;

10 c) tratar el ADN genómico de b), o un fragmento del mismo, con uno o más reactivos para convertir las bases de citosina que no están metiladas en la posición 5 del mismo a uracilo u otra base que sea detectablemente diferente a la citosina en términos de propiedades de hibridación; y en donde

d) la amplificación posterior al tratamiento en c) se lleva a cabo de una manera específica de metilación, concretamente mediante el uso de cebadores específicos de metilación u oligonucleótidos de bloqueo, y, además

15 e) la detección de los amplificados se lleva a cabo mediante una sonda de detección en tiempo real, como se describió anteriormente.

Preferiblemente, cuando la amplificación posterior de d) se lleva a cabo por medio de cebadores específicos de metilación, como se describió anteriormente, dichos cebadores específicos de metilación comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que se hibrida con una secuencia de ácidos nucleicos tratada de acuerdo con una de las SEQ ID NO: 120, 121, 238, 239 y secuencias complementarias a la misma, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG.

20 Se prefiere además que, cuando la amplificación posterior de d) se realiza por medio de cebadores específicos de metilación, como se describió anteriormente, dichos cebadores específicos de metilación comprendan una secuencia que tiene una longitud de al menos 9 nucleótidos que se hibrida con una secuencia de ácidos nucleicos de acuerdo con una de las SEQ ID NO: 120, 121 y secuencias complementarias a la misma, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG.

En un aspecto alternativo más preferido del método, la amplificación posterior de d) se lleva a cabo en presencia de oligonucleótidos de bloqueo, como se describió anteriormente. Dichos oligonucleótidos de bloqueo que comprenden una secuencia que tiene una longitud de al menos 9 nucleótidos que se hibrida con una secuencia de ácidos nucleicos tratada de acuerdo con una de las SEQ ID NO: 120, 121, 238, 239 y secuencias complementarias a la misma, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG, TpG o CpA.

30 Se prefiere particularmente que dichos oligonucleótidos bloqueadores comprendan una secuencia que tenga una longitud de al menos 9 nucleótidos que se hibrida con una secuencia de ácidos nucleicos tratada de acuerdo con una de las SEQ ID NO: 120, 121 y secuencias complementarias a la misma, en donde la base de la secuencia de dichos oligómeros comprende al menos un dinucleótido CpG, TpG o CpA.

35 El paso e) del método, a saber, la detección de los amplificados específicos indicativos del estado de metilación de una o más posiciones CpG de acuerdo con la SEQ ID NO: 31 se lleva a cabo mediante métodos de detección en tiempo real como se describió anteriormente.

Aspectos adicionales de la divulgación proporcionan un método para el análisis del estado de metilación del ADN genómico de acuerdo con la divulgación (SEQ ID NO: 31, y sus complementos) sin la necesidad de un tratamiento previo.

40 En el primer paso de tales realizaciones adicionales, la muestra de ADN genómico se aísla de fuentes de tejido o celulares. Preferiblemente, tales fuentes incluyen líneas celulares, portaobjetos histológicos, fluidos corporales o tejido embebido en parafina. En el segundo paso, se extrae el ADN genómico. La extracción puede realizarse por medios que son estándar para un experto en la materia, que incluyen, pero no se limitan al uso de lisados de detergente, sonificación y agitación vortical con perlas de vidrio. Una vez que se han extraído los ácidos nucleicos, se utiliza el ADN de doble cadena genómico en el análisis.

En una realización preferida, el ADN puede escindirse antes del tratamiento, y esto puede ser por cualquier medio estándar en el estado de la técnica, en particular con endonucleasas de restricción sensibles a la metilación.

50 En el tercer paso, el ADN se digiere con una o más enzimas de restricción sensibles a la metilación. La digestión se lleva a cabo de tal manera que la hidrólisis del ADN en el sitio de restricción informa el estado de metilación de un dinucleótido CpG específico.

En el cuarto paso, que es opcional pero una realización preferida, los fragmentos de restricción se amplifican. Esto se lleva a cabo preferiblemente usando una reacción en cadena de la polimerasa, y dichos amplificados pueden llevar marcadores detectables adecuados como se discutió anteriormente, a saber, marcadores de fluoróforo, radionucleidos y marcadores de masa.

- 5 En el quinto paso se detectan los amplificados. La detección puede ser por cualquier medio estándar en la técnica, por ejemplo, pero no limitado a, análisis de electroforesis en gel, análisis de hibridación, incorporación de etiquetas detectables dentro de los productos de PCR, análisis de ADN por arreglos, análisis MALDI o ESI.

- Después de la determinación del estado de metilación de los ácidos nucleicos genómicos, la presencia, ausencia o subclase de trastorno proliferativo de células de próstata se deduce con base en el estado de metilación de al menos una secuencia de dinucleótidos CpG de SEQ ID NO: 31, 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: 31.

Kits

- Además, un aspecto adicional de la presente divulgación es un kit que comprende, por ejemplo: un reactivo que contiene bisulfito; un conjunto de oligonucleótidos cebadores que contienen al menos dos oligonucleótidos cuyas secuencias en cada caso corresponden, son complementarias o se hibridan en condiciones restrictivas o muy restrictivas con un segmento largo de 16 bases de las secuencias SEQ ID NO: 31, SEQ ID NOs: 120, 121, 238, 239; oligonucleótidos y/u oligómeros de PNA; así como instrucciones para llevar a cabo y evaluar el método descrito. En una realización preferida adicional, dicho kit puede comprender además reactivos estándar para realizar un análisis de metilación de posición específica de CpG, en donde dicho análisis comprende una o más de las siguientes técnicas: MS-SNuPE, MSP, MethyLight™, HeavyMethyl™, COBRA y secuenciación de ácidos nucleicos.

- Se prefiere además que dicho compuesto comprenda un reactivo que contenga bisulfito; un conjunto de oligonucleótidos cebadores que contienen al menos dos oligonucleótidos cuyas secuencias en cada caso corresponden, son complementarias o se hibridan en condiciones restrictivas o altamente restrictivas con un segmento largo de 16 bases de las secuencias SEQ ID NO: 120, 121; oligonucleótidos y/u oligómeros de PNA; así como instrucciones para llevar a cabo y evaluar el método descrito. En una realización preferida adicional, dicho kit puede comprender además reactivos estándar para realizar un análisis de metilación de posición específica de CpG, en donde dicho análisis comprende una o más de las siguientes técnicas: MS-SNuPE, MSP, MethyLight™, HeavyMethyl™, COBRA y secuenciación de ácidos nucleicos.

Sin embargo, un kit también puede contener solo una parte de los componentes mencionados anteriormente.

Ejemplos

En los siguientes Ejemplos, se considera que 'uL' significa 'microlitro', es decir, 10^{-6} litros, en consecuencia, 'uM' significa 'micromolar'.

- Se aisló y analizó ADN genómico agrupado usando los métodos de descubrimiento, AP-PCR y MCA (Ejemplo 1). Estas tecnologías distinguen entre sitios CpG metilados y no metilados mediante el uso de enzimas sensibles a la metilación. En general, todo el ADN genómico se digiere primero para aumentar la manejabilidad y luego se digiere con una enzima de restricción sensible a la metilación. Los fragmentos metilados se amplifican preferiblemente porque la escisión en los sitios no metilados impide la amplificación de estos productos. Los fragmentos metilados diferencialmente identificados usando estas técnicas se secuencian (Ejemplo 2) y se comparan con el genoma humano usando la utilidad BLAST en la base de datos Ensembl. El conjunto de muestras se seleccionó en función del objetivo inicial del problema de diagnóstico a resolver, a saber, la detección y discriminación mejoradas de carcinomas de próstata en condiciones normales o benignas. Las siguientes comparaciones se realizaron utilizando tres grupos de muestras de cáncer de próstata "All Cancer" (10, 10 y 20 muestras cada uno), dos grupos de muestras de hiperplasia prostática benigna (BPH) (10 muestras cada uno), tres grupos de muestras de cáncer de próstata de bajo grado (10 muestras cada una), tres grupos de muestras de cáncer de próstata de alto grado (10 muestras cada uno) y un grupo de linfocitos de sangre periférica (PBL) (9 muestras):

1. BPH versus todos los cánceres (puntaje de Gleason alto y bajo; zonas de transición (TZ) y periféricas (PZ), 2 comparaciones)
2. BPH vs. Puntaje de Gleason bajo (Gleason <6, TZ y PZ representados, 2 comparaciones)
3. BPH vs. Puntaje de Gleason alto (Gleason >7, TZ y PZ representados, 2 comparaciones)
4. Puntaje de Gleason bajo versus puntaje de Gleason alto (para MCA, cada grupo se usó como probador y conductor)
5. BPH vs. PBL
6. Todo cáncer versus PBL

La comparación de BPH frente a PBL no se realizó para APPCR.

Para todas las comparaciones de MCA que incluían muestras de cáncer, el cáncer fue el probador. La comparación de puntaje de Gleason de bajo a alto se realizó dos veces, una con bajo como probador y otra con alto como probador, lo que eleva el número total de comparaciones para ACM a diez. En los experimentos con PBL, la muestra de PBL fue el conductor. Véase la tabla 1.

5 Tabla 1: Grupos de muestras utilizados en estudios de comparación (AP-PCR y MCA)

Comparación	Apodo	Tipo grupo	Grupo #	Muestras por grupo	Desglose de muestras			
					Gleason <6/zona Transición	Gleason <6/zona periférica	Gleason <7/zona Transición	Gleason <7/zona periférica
BPH vs. Todos los cánceres	BA1	HBP	1	10	10 BPH			
		Todos	1	10	2	3	2	3
BPH vs. Todos los cánceres	BA2	HBP	2	10	5 HBP, 5 Normal			
		Todos	2	10	3	2	3	2
BPH vs Bajo	BL1	HBP	1	10	10 HBP			
		Bajo	1	10	5	5	0	0
BPH vs Bajo	BL2	HBP	2	10	5 HBP, 5 Normal			
		Bajo	2	10	5	5	0	0
BPH vs Alto	BH1	BPH	1	10	10 BPH			
		Alto	1	10	0	0	5	5
BPH vs Alto	BH2	BPH	2	10	5 BPH, 5 Normal			
		Alto	2	10	0	0	5	5
Bajo vs Alto	HL	Bajo	3	10	5	5	0	0
		Alto	3	10	0	0	5	5
BPH vs PBL	BP	BPH	1	10	10 BPH			
		PBL	1	9	9 PBL			
Cáncer vs PBL	CP	AH	3	20	5	5	5	5
		PBL	1	9	9 PBL			

Ejemplo 1: MCA y AP-PCR

10 El objetivo de la siguiente investigación fue identificar una o más secuencias de dinucleótidos CpG metiladas diferencialmente primarias usando un ensayo controlado adecuado para identificar al menos una secuencia de dinucleótidos CpG metilada diferencialmente dentro del genoma completo, o una fracción representativa de la misma.

Todos los procesos se realizaron en muestras agrupadas y/o individuales, y el análisis se llevó a cabo utilizando dos métodos de descubrimiento diferentes; a saber, amplificación de CpG metilada (MCA) y PCR cebada arbitrariamente (AP-PCR).

AP-PCR. El análisis AP-PCR se realizó en clases de muestras de ADN genómico de la siguiente manera:

- 15 1. Aislamiento de ADN; El ADN genómico se aisló de las clases de muestra utilizando el kit Wizard™ disponible comercialmente;
2. Restricción de la digestión enzimática; cada grupo de muestras de ADN se digirió con 3 conjuntos diferentes de enzimas de restricción durante 16 horas a 37°C: RsaI (sitio de reconocimiento: GTAC); RsaI (sitio de reconocimiento: GTAC) más HpaII (sitio de reconocimiento: CCGG; sensible a la metilación); y RsaI (sitio de reconocimiento: GTAC) más MspI (sitio de reconocimiento: CCGG; insensible a la metilación);

20

3. Análisis de AP-PCR; cada una de las muestras de ADN digeridas por restricción se amplificó con los cebadores enumerados en la Tabla 2 a una temperatura de fusión de 40°C, y con ³²P dATP en los conjuntos de cebadores descritos en la Tabla 3.

5 4. Electroforesis en gel de poliacrilamida; 1.6 µl de cada muestra de AP-PCR se cargó en un gel de tamaño de secuenciación de poliacrilamida al 5% y se sometió a electroforesis durante 4 horas a 130 vatios. Los geles se transfirieron a papel de cromatografía, se cubrieron con una envoltura de sarán y se secaron en un secador de gel durante un período de aproximadamente 1 hora.

10 5. Exposición de película autorradiográfica; La película se expuso a geles secos durante 20 horas a menos 80°C, y luego se desarrolló. Se agregaron marcadores de Autorad Glogos II (Stratagene) al gel seco y la exposición se repitió con una nueva película. El primer autorradio se retuvo para registros, mientras que el segundo se usó para eliminar bandas; y

6. Las bandas correspondientes a la metilación diferencial se identificaron visualmente en el gel.

Dichas bandas se seccionaron y el ADN se aisló y se clonó usando el Kit de clonación Invitrogen TA.

Tabla 2: Cebadores utilizados de acuerdo con el Protocolo AP-PCR Ejemplo 1

Nombre	SEQ ID NO:	Secuencia
GC1	928	GGGCCGCGGC
GC2	929	CCCCGCGGGG
GC3	930	CGCGGGGGCG
GC4	931	GCGCGCCGCG
GC5	932	GCGGGGCGGC
G1	933	GCGCCGACGT
G2	934	CGGGACGCGA
G3	935	CCGCGATCGC
G4	936	TGGCCGCCGA
G5	937	TGCGACGCCG
G6	938	ATCCCGCCCG
G7	939	GCGCATGCGG
G8	940	GCGACGTGCG
G9	941	GCCGCGNGNG
G10	942	GCCCGCGNNG
APBS1	943	AGCGGCCGCG
APBS5	944	CTCCCACGCG

Nombre	SEQ ID NO:	Secuencia
APBS7	945	GAGGTGCGCG
APBS10	946	AGGGGACGCG
APBS11	947	GAGAGGCGCG
APBS12	948	GCCCCGCGA
APBS13	949	CGGGGCGCGA
APBS17	950	GGGGACGCGA
APBS18	951	ACCCACCCG

Tabla 3

Combinación	Cebador 1	Cebador 2	Cebador 3
101	GC1	G2	APBS1
103	GC3	G4	APBS1
105	GC5	G6	APBS1
107	GC2	G8	APBS5
109	GC4	G10	APBS5
111	GC1	G8	APBS7
113	GC3	G6	APBS7
115	GC5	G4	APBS7
117	GC2	G2	APBS10
119	GC4	G2	APBS10
121	GC1	G4	APBS11
123	GC3	G5	APBS11
125	GC5	G7	APBS11
127	GC2	G9	APBS12
129	GC4	G9	APBS12

ES 2 801 379 T3

Combinación	Cebador 1	Cebador 2	Cebador 3
131	GC1	G7	APBS13
133	GC3	G5	APBS13
135	GC5	G3	APBS13
137	GC2	G1	APBS17
139	GC4	G3	APBS17
141	GC1	G5	APBS18
143	GC3	G7	APBS18
145	GC5	G9	APBS18
147	G2	G3	APBS17
149	G4	G5	APBS17
151	G6	G7	APBS17
153	G8	G9	APBS13
155	G8	G10	APBS13
157	G6	G8	APBS12
159	G4	G6	APBS12
161	G2	G4	APBS12
163	G2	G10	APBS11
165	G2	G5	APBS11
167	G4	G7	APBS10
169	G6	G9	APBS10
171	G1	G8	APBS10
173	G6	G10	APBS7
175	G4	G8	APBS7
177	G2	G6	APBS5
179	G4	G10	APBS5

Combinación	Cebador 1	Cebador 2	Cebador 3
181	G2	G8	APBS5
183	APBS1	APBS10	APBS11
185	APBS5	APBS7	APBS17
187	APBS1	APBS12	APBS18
189	APBS10	APBS13	APBS17
191	APBS5	APBS11	APBS12
193	APBS7	APBS10	APBS13
195	APBS1	APBS5	APBS11
197	APBS7	APBS17	APBS18
199	APBS1	APBS12	APBS13

- 5 MCA. Se usó MCA para identificar secuencias hipermetiladas en una población de ADN genómico en comparación con una segunda población mediante la eliminación selectiva de secuencias que no contienen las regiones hipermetiladas. Esto se logró, como se describe en detalle aquí anteriormente, mediante la digestión del ADN genómico con una enzima sensible a la metilación que escinde los sitios de restricción no metilados para dejar extremos romos, seguido de la escisión con un isoesquizómero que es insensible a la metilación y deja extremos adhesivos. Esto es seguido por la ligación de adaptadores, la generación de amplicones y la hibridación sustractiva de la población de probadores con la población de conductores.

Las soluciones de reacción de digestión de restricción inicial contenían lo siguiente:

- 10 Conductores:

ADN	510 uL
regulador 4	60 uL
100x BSA	6 uL
SmaI (20U/uL)	24 uL

Probadores:

ADN	68 uL
regulador 4	10 uL
10x BSA	10 uL
SmaI (20U/uL)	2 uL

ES 2 801 379 T3

Las mezclas de reacción se incubaron durante la noche a temperatura ambiente.

Los agrupados se digirieron luego adicionalmente con Xmal (2 µl = 100 U), 6 horas a 37°C. Se agregaron 2 µl (20U) de Xmal a cada digestor del probador y 8 µl (80U) a cada digerido del controlador.

- 5 El material digerido y limpio se ligó al cebador adaptador RXMA24 + RXMA12 (Secuencia: RXMA24: AGCACTCTCCAGCCTCTCACCGAC (SEQ ID NO: 952); RXMA12: CCGGGTCGGTGA (SEQ ID NO: 953). Estos se hibridaron para crear el adaptador calentando juntos a 70°C y enfriándose lentamente a temperatura ambiente (RT) en una reacción de 30 uL:

Cada ADN	33 uL
Regulador T4	6 uL
Adaptador-cebador RXMA (100 uM)	20 uL
Ligasa	1 uL

- 10 La solución de reacción se incubó durante la noche a temperatura ambiente.

Se usaron 3 µl de la mezcla de ligación para las poblaciones de probadores y conductores en cada PCR inicial para generar los amplicones de partida. Las soluciones de reacción fueron las siguientes:

Probadores:

100uM RXMA24	1 uL
regulador PCR	10 uL
25 mM dNTPs	1.2 uL
ddH2O	68.8 uL
Taq de titanio	1 uL
100% DMSO	2 uL
5M Betaína	10 uL

- 15 Se agregaron 3 uL de ADN de prueba ligado a cada cóctel probador de 97 uL.

Conductores:

Los conductores se amplifican con dUTP en lugar de dTTP:

100uM RXMA24	1 uL
regulador PCR	10 uL
25 mM dNTPs	1.2 uL (25 mM cada dATP, dCTP, dGTP, y dUTP)
ddH2O	68.8 uL

ES 2 801 379 T3

Taq de titanio	1 uL
100% DMSO	2 uL
5M Betaína	10 uL

Se añadieron 3 uL de ADN conductor ligado a cada cóctel conductor de 97 uL.

Condiciones de PCR:

72 grados 5 min

5 30 ciclos:

95 grados 1 min

72 grados 3 min

: Extensión final:

72 grados 10 min.

10 Los amplicones probadores se digirieron luego con XmaI, produciendo extremos sobresalientes, y los amplicones impulsores se digirieron con SmaI, produciendo fragmentos extremos romos.

Conductores (SmaI):

ADN	500 uL
Regulador 4	100 uL
100x BSA	10 uL
H2O	340 uL
SmaI(20U/uL)	50 uL

Volumen total: 1 ml. Se incubó durante la noche a temperatura ambiente.

15 Probadores (XmaI):

ADN	20 uL
regulador 4	10 uL
10x BSA	10 uL
H2O	59 uL
XmaI (50U/uL)	1 uL
Total vol: 100 uL.	Incubado durante la noche a 37 grados.

Un nuevo conjunto de cebadores adaptadores (hibridados como se describe para los cebadores RXMA anteriores) JXMA24 + JXMA12 (Secuencia: JXMA24: ACCGACGTCGACTATCCATGAACC (SEQ ID NO: 954); JXMA12:

ES 2 801 379 T3

CCGGGGTTCATG (SEQ ID NO: 955) se ligó al Probador en un termociclador a 16°C durante 2 horas en la siguiente solución de reacción:

ADN	16 uL
regulador T4	3 uL
JXMA-P adaptador (100uM)	10 uL
T4 Ligasa (400U/uL)	1 uL

5 El probador digerido y los amplicones del conductor se hibridaron juntos. Se realizó una reacción de PCR selectiva usando el cebador JXMA24 (SEQ ID NO: 954). La solución de reacción contenía:

JXMA24	0.5 uL
regulador taq	5 uL
dNTPs	0.6 uL
ddH2O	27.4 uL
Betaina	5 uL
DMSO	1 uL
taq de titanio	0.5 uL
ADN	10 uL

Condiciones de PCR:

72 grados 8 min (rellenar extremos)

5 ciclos:

95 grados 1 min

10 72 grados 3 min

Extensión final:

72 grados 10 min

15 Subsecuentemente, se añadieron 20 µl de regulador de nucleasa de judías Mung más 10 µl de nucleasa de judías Mung (10U) y se incubaron a 37°C durante 30 minutos. Esta reacción se limpió y se usó como plantilla para 25 ciclos más de PCR usando el cebador JXMA24 en la siguiente solución de reacción:

JXMA24	1 uL
regulador taq	10 uL
dNTPs	1.2 uL
ddH2O	27 uL

Betaína	10 uL
DMSO	2 uL
taq de titanio	1 uL
ADN	48 uL

bajo las siguientes condiciones.

95 grados 2 min

30 ciclos:

95 grados 1 min

5 72 grados 3 min

Extensión final:

72 grados 10 min

Mantener a 4 grados

El producto de PCR resultante (probador) se digirió nuevamente usando XmaI:

10 45 uL ADN

15 uL Regulador 4

15 uL 10x BSA

71 uL H₂O

4 uL XmaI

15 Incubado durante la noche a 37 grados

Se ligó un tercer adaptador, NXMA24 (AGGCAACTGTGCTATCCGAGTGAC; SEQ ID NO: 956) + NXMA12 (CCGGGTCACTCG; SEQ ID NO: 957). El probador (500 ng) se hibridó por segunda vez con el controlador digerido original (40 ug) en 4 uL EE (EPPS 30 mM, EDTA 3 mM) y 1 uL NaCl 5M a 67°C durante 20 horas. La PCR selectiva se realizó usando el cebador NXMA24 de la siguiente manera:

NXMA24	0.5 uL
regulador taq	5 uL
dNTPs	0.6 uL
ddH ₂ O	27.4 uL
Betaína	5 uL
DMSO	1 uL
taq de titanio	0.5 uL
ADN	10 uL

20

Programa PCR:

72 grados 8 min (rellenar extremos)

8 ciclos:

95 grados 1 min

72 grados 3 min

5 Extension final:

72 grados 10 min

La solución de reacción se mantuvo a 4 grados

10 Posteriormente, se añadieron 20 µl de regulador de nucleasa de judías Mung más 10 µl de nucleasa de judías Mung (10U) y se incubaron a 30°C durante 30 minutos. Esta reacción se limpió y se usó como plantilla para 25 ciclos más de PCR usando el cebador NXMA24 de la siguiente manera:

Solución de reacción

NXMA24	1 uL
regulador taq	10 uL
dNTPs	1.2 uL
ddH2O	27 uL
Betaína	10 uL
DMSO	2 uL
taq de titanio	1 uL
ADN	48 uL

Programa PCR:

95 grados 2 min

30 ciclos:

15 95 grados 1 min

72 grados 3 min

Extension final:

72 grados 10 min

Mantener a 4 grados

20 El producto de PCR resultante se digirió con XmaI:

Solución de reacción:

ADN	38 uL
regulador 4	5 uL
10x BSA	5 uL

Xma I	2 uL
Incubado durante la noche a 37 grados.	

La digestión de ADN se ligó luego en el vector pBC Sk-predigerido con XmaI y se trató con fosfatasa (675 ng). Se usaron 5 µl de la mezcla de ligación para transformar células TOP10™ químicamente competentes de acuerdo con las instrucciones del fabricante. Las transformaciones se sembraron en placas LB/XGal/IPTG/CAM. Las colonias de inserto seleccionadas se secuenciaron de acuerdo con el Ejemplo 2.

El Ejemplo 1 dio como resultado un gran número de secuencias únicas que eran candidatos potenciales para marcadores de ensayo. Se eliminó un subconjunto de estas secuencias debido a su alto contenido de repetición (>50%). Se identificaron un total de 480 secuencias únicas en las comparaciones realizadas para este estudio. Un subconjunto de estas secuencias se seleccionó adicionalmente utilizando el siguiente procedimiento de puntuación:

Apariencia usando múltiples métodos

Apariencia en múltiples agrupados

Ubicado dentro de la isla CpG

Ubicado dentro de la región promotora de un gen

Cerca o dentro del gen predicho o conocido

Se sabe que está asociado con la enfermedad.

Clase de gen (factor de transcripción, factor de crecimiento, etc.)

Elemento repetitivo (puntaje negativo)

Bajo este esquema de puntuación, una secuencia de MeST recibe un punto para cada uno de los criterios anteriores, y recibe una puntuación de (-)8 por tener un contenido de secuencia repetitiva superior al 50%. El puntaje más alto posible es 7, el más bajo es (-)8. Las puntuaciones se generan automáticamente utilizando una base de datos registrada. Del conjunto inicial de 480 secuencias de MeST, 277 obtuvieron 0 o más. Usando los criterios de puntuación anteriores, junto con la revisión manual de las secuencias, el número de candidatos MeST se redujo aún más a 126 secuencias únicas.

El diseño del cebador para las 126 secuencias se inició luego con el fin de la Secuenciación con bisulfito. Treinta y cinco de las secuencias se descartaron por diversos motivos, incluida la incapacidad de diseñar cebadores adecuados, el fallo de amplificación del ADN de control o si un examen adicional de la secuencia o las actualizaciones de la base de datos Ensembl revelaron secuencias de mala calidad o repetidas que no se habían observado previamente.

Ejemplo 2: Secuenciación con bisulfito

Para la Secuenciación con bisulfito, se diseñaron cebadores de amplificación para cubrir cada secuencia de MeST identificada cuando fuera posible o parte de las regiones flanqueantes de 1000 pb corriente arriba o 1000 pb corriente abajo que rodean la posición. Las muestras utilizadas en el Ejemplo 1 se utilizaron para la producción de amplicones en esta fase del estudio. Cada muestra fue tratada con bisulfito de sodio y secuenciada. Los datos de secuencia se obtuvieron utilizando la tecnología de secuenciación ABI 3700. Las trazas de secuencia obtenidas se normalizaron y se calculó el porcentaje de metilación utilizando el programa de análisis de trazas de secuencia de bisulfito ESME registrado por Epigenomic.

Resultados de la secuenciación con bisulfito

Se observaron las siguientes propiedades:

1. La secuenciación con bisulfito indica la metilación diferencial de un sitio CpG entre clases seleccionadas de muestras (puntaje de Fisher)

2. Se observa cometilación

3. Si solo un sitio tiene un puntaje de Fisher >1, ¿hay sitios adicionales alrededor con un puntaje de Fisher >0.5?

Las regiones genómicas que se consideraban que demostraban una cometilación significativa según lo evaluado por estos criterios luego avanzaron a una investigación adicional.

Las Figuras 1 a 3 son matrices clasificadas producidas a partir de datos de secuenciación con bisulfito analizados por el software registrado de Epigenomics (Véase WO 2004/000463 para más información). La matriz general representa los datos de secuencia para una región de interés. Cada fila de la matriz representa un único sitio CpG dentro del fragmento y cada columna representa una muestra de ADN individual. La barra de la izquierda representa una escala del porcentaje de metilación, con el grado de metilación representado por el tono de cada posición dentro de la columna, desde el negro que representa el 100% de metilación hasta el gris claro que representa el 0% de metilación. Las posiciones blancas representaban una medida sin datos disponibles.

La Figura 1 muestra los datos de secuenciación de un fragmento del gen Receptor de Prostaglandina E2, Subtipo EP4. Aquí, la secuenciación con bisulfito mostró patrones de metilación diferenciales, pero no concluyentes entre las muestras. El gen se investigó más a fondo en un conjunto de muestras más grande utilizando el proceso de arreglos (Ejemplo 3) ya que la precisión de este gen como marcador podría mejorarse cuando se analiza en combinación con otros genes.

La Figura 2 muestra los datos de secuenciación de un fragmento del gen Receptor Nuclear Huérfano (Factor de Transcripción de Fetoproteína alfa-1). En este caso, la Secuenciación con bisulfito indicó metilación diferencial o cometilación entre tipos de muestra.

La Figura 3 muestra los datos de secuenciación de un fragmento del gen 1-acil-SN-glicerol-3-fosfato aciltransferasa gamma. Este era representativo de un subconjunto de ROI para el que solo se obtuvieron lecturas de secuencia de baja calidad y el gen solo pudo analizarse significativamente usando el proceso de arreglos (Ejemplo 3).

Ejemplo 3; Análisis por arreglos

Una selección de las regiones genómicas metiladas diferencialmente se analizó adicionalmente por medio de análisis por arreglos de alto rendimiento. El ensayo final más útil adecuado para una prueba de detección de diagnóstico/clasificación permitiría el análisis de fluidos corporales como suero, plasma u sedimento de orina (obviando la necesidad de procedimientos invasivos). Por lo tanto, el conjunto de muestras incluía muestras de ADN de otros cánceres que pueden estar presentes en la sangre para proporcionar conjuntos de marcadores más específicos para ensayos sensibles. //

Descripción del conjunto de muestras para el estudio en chips.

El conjunto de muestras para el análisis de microarreglos se diseñó para proporcionar información relativa tanto a la sensibilidad como a la especificidad de los candidatos marcadores. Se analizó una gran cantidad de muestras (Tabla 7) de cáncer de próstata, BPH y próstata normal. Las muestras de cáncer de próstata se agruparon por puntuación de Gleason (alta (≥ 8), moderada (7) y baja (≤ 6)) y por zona (periférica o de transición). La distribución de las muestras de BPH fue aleatoria, pero debido a que la mayoría de las BPH se derivan de la zona de transición, se puede suponer que la mayoría de las muestras eran de ese origen. Además de las muestras de próstata, se incluyeron otros tipos de cáncer para evaluar la especificidad de la próstata. Se incluyeron muestras de PBL para analizar el uso potencial de estos marcadores en un análisis de sangre. También se incluyeron hígado normal y cáncer de hígado debido a la metilación observada de GSTP1 en estas muestras.

Tabla 7. Descripción general de las muestras para el estudio por arreglos.

Tipo de muestra	Tipo de muestra
Cánceres de próstata	Cánceres endocrinos relacionados
• Grado alto (Gleason ≥ 8)	Mama
• Zona de transición	Masculino
• Zona periférica	Femenino
• Grado bajo (Gleason ≤ 6)	• Ovario
• Zona de transición	• Uterino
• Zona periférica	Otros cánceres
• Grado moderado (Gleason = 7)	• Hígado

Tipo de muestra	Tipo de muestra
• Zona de transición	• Pulmón
• Zona periférica	• Esofágico
• Cánceres de próstata adicionales	• Glándula salival
• Terapia posthormonal	• Estómago
Enfermedad benigna de la próstata	• Pancreático
• BPH	• Melanoma
• Fibroma benigno	• Colon
• Prostatitis	Otros tejidos normales
Cánceres del tracto genitourinario	• Próstata
• Vejiga	• Transicional
• Testicular	• Periférico
• Riñón	• Vejiga
	• Riñón
	• Hígado
	• Testículos
	• Esperma
	• Uréter
	• PBLs

Extracción de ADN

Se recibieron muestras de tejido congelado o de ADN genómico extraído. Todas las muestras de ADN se extrajeron usando columnas Qiagen Genomic Tip-500 o el dispositivo MagnaPure.

5 Tratamiento con bisulfito y PCR multiplex

El ADN genómico total de todas las muestras se trató con bisulfito para convertir las citosinas no metiladas en uracilo. Las citosinas metiladas permanecieron conservadas como citosinas. El tratamiento con bisulfito se realizó utilizando el proceso de tratamiento con bisulfito registrado de Epigenomics. Se realizaron dos reacciones de bisulfito independientes por muestra de paciente. Después de la bisulfitación, se usaron 10 ng de cada muestra de ADN en posteriores reacciones de PCR multiplex (mPCR) que contenían 7-8 pares de cebadores.

10

Hibridación

Cada reacción contenía lo siguiente:

0.4 mM de cada dNTPS

1 Unidad Taq Polimerasa

2.5 ul de regulador de PCR

MgCl₂ 3.5 mM

5 Conjunto de cebadores 80 nM (cebadores 12-16)

11.25 ng de ADN (tratado con bisulfito)

Se muestran detalles adicionales de los cebadores en la Tabla 8.

10 Se llevaron a cabo cuarenta ciclos de la siguiente manera: desnaturalización a 95°C durante 15 minutos, seguida de fusión a 55°C durante 45 segundos, alargamiento del cebador a 65°C durante 2 minutos. Se realizó un alargamiento final a 65°C durante 10 min.

Hibridación

15 Todos los productos de PCR de cada muestra individual se hibridaron luego con portaobjetos de vidrio que portaban un par de oligonucleótidos inmovilizados para cada posición de CpG bajo análisis. Cada uno de estos oligonucleótidos de detección se diseñó para hibridarse con la secuencia convertida de bisulfito alrededor de un sitio CpG que originalmente estaba no metilado (TG) o metilado (CG). Véase la Tabla 2 para obtener más detalles de todos los oligonucleótidos de hibridación utilizados (tanto informativos como no informativos). Las condiciones de hibridación se seleccionaron para permitir la detección de las diferencias de un solo nucleótido entre las variantes TG y CG.

20 Se diluyeron 5 ul de volumen de cada producto de PCR multiplex en 10 x regulador SSARC. La mezcla de reacción se hibridó luego a los oligonucleótidos de detección como sigue. Desnaturalización a 95°C, enfriamiento a 10°C, hibridación a 42°C durante la noche seguida de lavado con 10 x SSARC y dH₂O a 42°C.

Los detalles adicionales de los oligonucleótidos de hibridación se muestran en la Tabla 9.

Las señales fluorescentes de cada oligonucleótido hibridado se detectaron usando un escáner y un software Genepix. Las relaciones para las dos señales (del oligonucleótido CG y el oligonucleótido TG usado para analizar cada posición CpG) se calcularon con base en la comparación de la intensidad de las señales fluorescentes.

25 Para cada paciente, 2 alícuotas de ADN fueron tratadas con bisulfito y para cada muestra de ADN tratada con bisulfito se realizaron dos hibridaciones, dando como resultado un total de 4 chips procesados por paciente. Para la hibridación, las muestras se agruparon en 2 rondas de procesamiento para evitar un posible sesgo de proceso. Como se indicó, cada una de las 2 rondas incluyó una redundancia de 2 veces para cada muestra de ADN para la redundancia de 4 veces por paciente. Las muestras se hibridaron en lotes de 112 muestras aleatorizadas por sexo, diagnóstico, tejido y lote de bisulfito.

Métodos de análisis de datos:

Análisis de los datos del chip.

35 Para el análisis de los datos del chip se usó el software registrado de Epigenomics "EpiScape®". Abarca una variedad de herramientas estadísticas y nuevos métodos de aprendizaje automático para analizar y visualizar los datos de arreglos de metilación. En las siguientes secciones se resumen las técnicas de análisis de datos más importantes que se han aplicado para analizar los datos.

Desde intensidades de hibridación sin procesar hasta relaciones de metilación

40 La relación de metilación logarítmica (log (CG/TG)) en cada posición CpG se determina de acuerdo con un flujo de preprocesamiento estandarizado. Esta relación logarítmica tiene la propiedad de que el ruido de hibridación tiene una variación aproximadamente constante en todo el rango de posibles tasas de metilación.

Prueba de hipótesis

45 La tarea principal fue identificar marcadores que puedan hacer una contribución significativa a la predicción de clase de las muestras. Para las "realizaciones particularmente preferidas" de la divulgación, se detectó la contribución significativa cuando la hipótesis nula de que un modelo de predicción que incluya el marcador no mejora el rendimiento de clasificación con respecto a un modelo sin el marcador podría rechazarse con $p < 0.05$. Debido a que se aplicó esta prueba a un conjunto completo de marcadores potenciales, se corrigieron los valores p para múltiples pruebas. Se hizo esto aplicando la corrección conservadora de Bonferroni, que simplemente multiplica los valores p de un solo marcador con el número de marcadores potenciales probados. También se proporcionan resultados con el método menos conservador de Rata de Falso Descubrimiento (FDR).

En todo este ejemplo, un marcador (a veces también denominado simplemente gen o amplicón) también se denomina región genómica de interés (ROI). Se compone de varias posiciones CpG en la región genómica respectiva. Para probar la hipótesis nula de que un marcador no tiene poder predictivo, se usó la prueba de relación de probabilidad para los modelos de regresión logística. El modelo de regresión logística para un solo marcador es una combinación lineal de mediciones de metilación de todas las posiciones CpG en el ROI respectivo. El modelo de regresión logística ajustado se compara con un modelo de probabilidad constante que es independiente de la metilación y representa la hipótesis nula. El valor p del marcador se calculó mediante la prueba de relación de probabilidad.

Un valor p significativo para un marcador significa que la metilación de este ROI tiene alguna correlación sistemática con la pregunta de interés dada por las dos clases. En general, un valor p significativo no implica necesariamente un buen rendimiento de clasificación. Sin embargo, debido a que con la regresión logística se utiliza un predictor lineal como base de la presente estadística de prueba, los valores p pequeños son indicativos de un buen rendimiento clínico.

Predicción de clase por aprendizaje supervisado

Para dar una estimación confiable de cuán bien el conjunto CpG de un marcador seleccionado puede diferenciar entre diferentes clases de tejidos, se puede determinar su Precisión de predicción por clasificación. Para ese propósito, se calculó una función de predicción basada en el perfil de metilación utilizando un cierto conjunto de muestras de tejido con una etiqueta de clase específica. Este paso se llama trazado y explota el conocimiento previo representado por las etiquetas de datos. La precisión de predicción de esa función se probó en un conjunto de muestras independientes. Como método de elección, se utilizó el algoritmo de máquina de vectores de soporte (SVM) para aprender la función de predicción. En este análisis, la sensibilidad y la especificidad se ponderaron por igual. Esto se logró estableciendo que el riesgo asociado con las clasificaciones de falsos positivos y falsos negativos sea inversamente proporcional a los tamaños de clase respectivos. Por lo tanto, se podría esperar que la sensibilidad y especificidad del clasificador resultante sea aproximadamente igual. Téngase en cuenta que esta ponderación se puede adaptar de acuerdo con los requisitos clínicos.

Estimación del rendimiento de la predicción de la clase de tejido: validación cruzada

Con un tamaño de muestra limitado, el método de validación cruzada proporcionó una estimación efectiva y confiable para la precisión de predicción de una función discriminadora y, por lo tanto, además de la importancia de los marcadores, se proporcionaron estimaciones de precisión, sensibilidad y especificidad de validación cruzada. Para cada tarea de clasificación, las muestras se dividieron en 5 grupos de aproximadamente el mismo tamaño. Luego, el algoritmo de aprendizaje fue trazado en 4 de estos 5 grupos de muestra. El predictor obtenido por este método se probó en el grupo restante de muestras de prueba independientes. El número de clasificaciones positivas y negativas correctas se contó en 10 ejecuciones para el algoritmo de aprendizaje para todas las opciones posibles del grupo de prueba independiente sin utilizar ningún conocimiento obtenido de las ejecuciones anteriores. Este procedimiento se repitió en 10 permutaciones aleatorias del conjunto de muestras, lo que proporciona una mejor estimación del rendimiento de la predicción que si se realiza simplemente dividiendo las muestras en un conjunto de muestras de entrenamiento y un conjunto de prueba independiente.

Resultados de análisis de datos

Nuestro primer paso en el análisis de los datos de los arreglos fue observar marcadores discriminatorios en una comparación de todos los tejidos de origen prostático. Primero se comparó el tejido prostático normal y BPH con muestras de cáncer de próstata, y se encontró que muchos de los marcadores utilizados en este estudio tienen valores p que cumplen con los criterios deseados (Figura 4). Luego, se compararon los tejidos de cáncer de próstata con todas las otras clases de tejidos utilizados en este estudio (Tabla 7). Casi todos los marcadores cumplieron los criterios estadísticos especificados con este conjunto de muestras. Se sabe que el gen GSTP1 está hipermetilado en el cáncer de próstata, pero también muestra hipermetilación en otros tipos de cáncer. Por lo tanto, la presente comparación final fue un examen más detallado de los niveles de metilación en el cáncer de próstata versus otros tipos de cáncer.

Próstata normal y BPH vs Cáncer de próstata

En esta comparación, la clase negativa consistió en 91 muestras de próstata normal y BPH. La clase positiva consistió en 99 muestras de cáncer de próstata. La mayoría de los marcadores cumplieron los criterios de valor $p < 0.05$ (Figura 4). Los valores p, la precisión, la sensibilidad y la especificidad del análisis se muestran en la Tabla 4. Los mejores 12 marcadores se muestran más en la Figura 5.

Próstata normal y otros tejidos vs cáncer de próstata

Luego se realizaron comparaciones en el conjunto de muestra completo. El grupo negativo se expandió para incluir tejido normal de otros órganos y cáncer de otros orígenes además de la próstata, de acuerdo con la Tabla 7. La clase negativa consta de 254 muestras de próstata normal, BPH y otros tejidos normales y cancerosos. La clase positiva consta de 99 muestras de cáncer de próstata. Nuevamente, los valores p para la mayoría de los marcadores cumple con el nivel de significancia de $p = < 0.05$ (Tabla 5). La precisión del marcador de mayor rendimiento es ~86% (consulte

la Figura 6 y/o la Tabla 5). Los valores p, la precisión, la sensibilidad y la especificidad del análisis se muestran en la Tabla 5. Los mejores 12 marcadores se muestran más en la Figura 7.

Otros cánceres vs cáncer de próstata

- 5 Dado que la hipermetilación de GSTP1 (marcador de cáncer de próstata de metilación de última generación) no es específica de la próstata, se examinó el estado de metilación del cáncer de próstata y otros cánceres con mayor detalle. La Figura 16 muestra que GSTP1 estaba fuertemente hipermetilada en el cáncer de hígado y en menor grado en el cáncer de mama. Sin embargo, varios de los mejores marcadores candidatos distinguen bien entre el cáncer de próstata y el hígado. Los valores de p, la precisión, la sensibilidad y la especificidad del análisis se muestran en la Tabla 6. Los mejores 12 marcadores se muestran en la Figura 8.
- 10 Las Tablas 4-6 a continuación resumen las características de rendimiento de todos los marcadores en las siguientes comparaciones de tejidos: Próstata normal y BPH vs cáncer de próstata (Tabla 4); Próstata normal, BPH y otros tejidos vs cáncer de próstata (Tabla 5); y otros tejidos cancerosos vs cáncer de próstata (Tabla 6).
- 15 Los análisis en las Tablas 4 y 5 contenían BPH y muestras de próstata normales en el grupo de análisis. El análisis de próstata normal y BPH vs cáncer de próstata se diseñó para determinar el rendimiento de los marcadores en un entorno específico de próstata. El análisis que incluyó otros tejidos, tanto cancerosos como normales (Próstata normal, BPH y otros tejidos vs cáncer de próstata) tuvo en cuenta el rendimiento de los marcadores con antecedentes que pueden contribuir o alterar el rendimiento general de los marcadores en muestras remotas.
- 20 Los tipos de cáncer (Tabla 6) también se compararon debido a la propensión a que GSTP1 se metile en múltiples tipos de cáncer. Este tipo de falta de especificidad podría tener un impacto negativo en el rendimiento de un marcador en los ensayos basados en fluidos corporales. GSTP1 está altamente metilado en el cáncer de próstata, pero también en el cáncer de hígado como se anticipó. IGF2 se metila de manera similar en el cáncer de hígado. La mayoría de los marcadores que se muestran en la Figura 8 no están metilados en la mayoría de los tipos de cáncer, con la excepción del cáncer de próstata. De las Figuras 4-8, se puede observar que hay múltiples candidatos que tienen el potencial de ser marcadores informativos y precisos. Se recomienda combinar múltiples marcadores para garantizar una alta
- 25 sensibilidad y especificidad.

Tabla 4: Próstata normal, BPH y otros tejidos vs cáncer de próstata

Genómica SEQ ID NO :	Cadena en sentido metilada tratada SEQ ID NO:	Cadena antisentido metilada tratada SEQ ID NO:	no metilada tratada SEQ ID NO:	Cadena antisentido no metilada tratada SEQ ID NO:	Valor p	Precisión	Sensibilidad	Especificidad
31	120	121	238	239	6.90E- 018	0.7	0.79	0.66

Tabla 5: Próstata normal y BPH vs cáncer de próstata

Genómica SEQ ID NO :	Cadena en sentido metilada tratada SEQ ID NO:	Cadena antisentido metilada tratada SEQ ID NO:	Cadena sentido no metilada tratada SEQ ID NO:	Cadena antisentido no metilada tratada SEQ ID NO:	valor p	Precisión	Sensibilidad	Especificidad
31	120	121	238	239	5.20E- 015	0.76	0.71	0.81

Tabla 6: Otros tipos de cáncer vs cáncer de próstata

Genómica SEQ ID NO :	Cadena en sentido metilada tratada SEQ ID NO:	Cadena antisentido metilada tratada SEQ ID NO:	Cadena sentido no metilada tratada SEQ ID NO:	Cadena antisentido no metilada tratada SEQ ID NO:	valor p	Precisión	Sensibilidad	Especificidad
31	120	121	238	239	1.5e- 13	0.74	0.92	0.58

La Tabla 7 se encuentra al comienzo del Ejemplo 3.

Tabla 8

Nombre y/o SEQ ID No:	Cebador:	Longitud del amplificado:
DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) (SEQ ID NO: 31)	AAACCCTACTTCCTACAAACAA (SEQ ID NO: 365)	420
	AGGGAGGTTTGGTGTATTTT (SEQ ID NO: 364)	

Tabla 9

No:	Nombre del gen y/o SEQ ID NO	Oligo:
287	DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) (SEQ ID NO: 31)	TATCGGATTATCGCGG (SEQ ID NO: 702)
288	DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) (SEQ ID NO: 31)	ATTGGATTATTGTGGGG (SEQ ID NO: 703)
289	DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) (SEQ ID NO: 31)	GTCGGTAGTTTATCGGAT (SEQ ID NO: 704)
290	DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) (SEQ ID NO: 31)	GTTGGTAGTTTATTGGAT (SEQ ID NO: 705)
291	DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) (SEQ ID NO: 31)	TAGGAGACGTTACGTT (SEQ ID NO: 706)
292	DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) (SEQ ID NO: 31)	AGATGTTATGTTAGGGT (SEQ ID NO: 707)

5 Ejemplo 4

En el siguiente análisis, el estado de metilación de los genes de acuerdo con la Tabla 10 se analizó mediante reacción en cadena de polimerasa específica de metilación usando los cebadores de acuerdo con la Tabla 10 (más adelante).

10 El estudio se realizó en 50 muestras de tejido de cáncer de próstata y 50 de hiperplasia prostática benigna (BPH). El ADN genómico se analizó utilizando la técnica MSP después de la conversión con bisulfito. El proceso con bisulfito convierte las citosinas no metiladas en uracilo mientras que las citosinas metiladas permanecen conservadas. El tratamiento con bisulfito se realizó con modificaciones menores de acuerdo con el protocolo descrito en Olek et al. (1996) Las secuencias de interés se amplificaron por medio de cebadores específicos de metilación, y el amplificado se detectó por medio de sondas Taqman (Véase Tabla 10).

Tabla 10

Genómica SEQ ID NO :	Cebador	Cebador	Sonda Taqman
31	Cgttttttcgttttat tttcgc (SEQ ID NO: 976)	Gacaaaaaacgccacgt c (SEQ ID NO: 977)	56- FAM/ccgacaattcaccga atcaccg/3BHQ_1 (SEQ ID NO: 978)

El término BHQ denota colorantes Black Hole Quencher (BHQ™), FAM es una molécula de colorante fluoróforo utilizada comúnmente.

Reactivos:

- 5 Se utiliza un conjunto estándar de reactivos y condiciones de ciclización para el establecimiento de MSP y la amplificación de plantilla. Las condiciones estándar se resumen en las Tablas 11 y 12. Antes de analizar muestras biológicas, se establecieron amplicones utilizando 100 picogramos de ADN completamente metilado como control positivo y 100 nanogramos de ADN no metilado como control negativo. También se verificó la sensibilidad relativa de las condiciones de reacción utilizando 50 picogramos de ADN metilado en un fondo de 50 nanogramos de ADN no metilado. Las concentraciones de reactivos se resumen en la Tabla 11 y las condiciones de ciclización para el ABI 7700 se definen en la Tabla 12.

Tabla 11: Concentraciones de reactivos.

Reactivo	Conc madre (uM)	Conc. Rx final (nM)	Conc.MM (uM)	Volumen MM (uL)
Avance	10.0	500.0	3.33	35.0
Retroceso	10.0	500.0	3.33	35.0
Sonda	100.0	400.0	2.67	2.8
Agua	-	-	-	32.2
Taqmix	-	-	-	245.0
Total	-	-	-	350.0

Tabla 12: Condiciones de ciclización

Temperatura	Tiempo (seg)	# de Ciclos
(C)		
Desnaturalización		1
95	600	
Fusión		50
95	10	
60 o 63	45	

- 15 **Análisis de los datos:**

Predicción de clase por aprendizaje supervisado

- 20 Con el fin de dar una estimación confiable de cuán bien el conjunto CpG de un marcador seleccionado puede diferenciar entre diferentes clases de tejidos, se puede determinar su precisión de predicción por clasificación. Para ese propósito, se calculó una función de predicción basada en el perfil de metilación utilizando un cierto conjunto de muestras de tejido con una etiqueta de clase específica. Este paso se llama capacitación y explota el conocimiento previo representado por las etiquetas de datos. La precisión de predicción de esa función se prueba en un conjunto

- de muestras independientes. Como método de elección, se utilizó el algoritmo de máquina de vectores de soporte (SVM) (véase, por ejemplo, Cristiannini, N. and Shawe-Taylor, J. An introduction to support vector machines. Cambridge, UK: Cambridge University Press, 2000.; Duda, R. O., Hart, P. E., and Stork, D. G. Pattern Classification. New York: John Wiley & Sons, 2001.) para aprender la función de predicción. Para este informe, la sensibilidad y la especificidad se ponderan por igual. Esto se logra estableciendo que el riesgo asociado con las clasificaciones de falsos positivos y falsos negativos sea inversamente proporcional a los tamaños de clase respectivos. Por lo tanto, se puede esperar que la sensibilidad y especificidad del clasificador resultante sea aproximadamente igual. Téngase en cuenta que esta ponderación se puede adaptar de acuerdo con los requisitos clínicos.

Resultados

- Para determinar la sensibilidad y especificidad de dichos marcadores, se seleccionaron 50 muestras de cáncer de próstata y 50 BPH utilizando los parámetros definidos. Las muestras se analizaron previamente siguiendo un criterio técnico de ADN metilado frente a ADN no metilado. Después de asegurarse de que eran específicos para el ADN metilado sin amplificar el ADN no metilado común, los ensayos se ejecutaron usando PCR en tiempo real MethyLight en una plataforma TaqMan (ABI7900). El rendimiento final del ensayo se describe en la Tabla 13. El AUC y los valores de sensibilidad y especificidad correspondientes se calcularon utilizando los algoritmos SVM.

Tabla 13

Nombre del gen	Genómica SEQ ID NO :	AUC	Sensibilidad	Especificidad
RECEPTOR DE PROSTAGLANDINA E2, SUBTIPO EP4		0.921	0.829	0.871
HISTONA H4		0.918	0.88	0.719
PROTEÍNA 16 DOMINIO PR DEDO DE ZINC		0.871	0.768	0.822
RECEPTOR NUCLEAR HUÉRFANO NR5A2		0.859	0.694	0.878
DOMINIO QUINASA 1 LIM	31	0.868	0.791	0.755
Región genómica		0.842	0.815	0.704
PROTEÍNA LIM/HOMEOBOX LHX9		0.745	0.695	0.653

Ejemplo 5

- El objetivo del siguiente estudio fue analizar el estado de metilación de los marcadores de cáncer de próstata en diferentes muestras de fluidos corporales con el fin de identificar la opción preferida de fluido corporal (orina o suero) para la prueba y el marcador, marcadores o combinaciones preferidos de marcadores. El estudio se realizó en muestras de suero y orina de 80 pacientes con una edad promedio de 65 años y representativos de varios tipos raciales (caucásicos, afroamericanos, etc.). En cada caso, el ADN genómico se analizó utilizando la técnica HeavyMethyl o MSP después de la conversión con bisulfito.
- Se preparó sedimento de orina para análisis y se trató con bisulfito de acuerdo con lo siguiente:
- Se purificaron 200 ul de muestras de sedimento utilizando el kit de aislamiento de ADN Magnapure 1 con un volumen de elución de 100 ul.
 - Se aplicaron 5 ul de PCR HD6 en Magnapure Eluate, para determinar la concentración de ADN
 - Se trataron 100 ul de la solución de ADN utilizando la técnica de tratamiento con bisulfito.
 - 10 ul PCR cuantitativa específica para bisulfito C3
 - 5 ul 'Prueba de sulfito de Merck'

Se preparó suero para análisis y se trató con bisulfito de acuerdo con lo siguiente:

- Las muestras de suero de 1 ml se purificaron usando el ácido nucleico total Magnapure DNA Large Volume con un volumen de elución de 100 ul.

- 5 ul PCR HD6 en Magnapure Eluate: para determinar la concentración de ADN

- Se trataron 100 ul de la solución de ADN utilizando una técnica registrada de tratamiento con bisulfito.

5 • 10 ul PCR cuantitativa específica de bisulfito C3

- 5 ul 'Prueba de sulfito de Merck'

HD6 es un ensayo genómico que amplifica el ADN convertido sin bisulfito. Se utiliza para cuantificar el ADN después de la extracción inicial. Es específico para el gen beta-globulina.

10 C3 es un ensayo específico de bisulfito ubicado cerca del exón 4 del gen GSTP1. Se usó de manera similar al presente ensayo HB14 que también se usa para cuantificar el ADN convertido con bisulfito.

Ambos requieren el uso de una curva estándar y luego la localización de las curvas de amplificación y la concentración correspondiente se usa para dar un valor cuantitativo para las muestras de ADN.

Se realizaron ejecuciones de PCR individuales en 10 ul de ADN tratado con bisulfito por muestra para cada uno de los marcadores como se describe a continuación.

15 Ensayo de metilo pesado del gen GSTP1

En el análisis se analizó el estado de metilación del gen GSTP1.

Análisis MSP del gen HISTONA H4.

En el análisis se analizó el estado de metilación del gen HISTONA H4

Análisis MSP del gen RECEPTOR DE PROSTAGLANDINA E2.

20 En el análisis se analizó el estado de metilación del gen RECEPTOR DE PROSTAGLANDINA E2.

Análisis MSP del gen RECEPTOR NUCLEAR HUÉRFANO NR5A2.

En el análisis se analizó el estado de metilación del gen RECEPTOR NUCLEAR HUÉRFANO NR5A2.

Análisis de marcadores

25 Los resultados se analizaron cualitativamente puntuando la amplificación como +/- y cuantitativamente determinando el porcentaje de ADN metilado como una fracción del ADN total calculado usando la PCR específica de bisulfito C3. Para medir el ADN metilado total, se incluyó una curva estándar 100% metilada (ADN tratado con Chemicon SSS1) en cada ensayo.

Resultados

30 Para cada marcador se determinó una curva característica de funcionamiento del receptor (curva ROC) del ensayo. Un ROC es un gráfico de la tasa positiva verdadera contra la tasa de falsos positivos para los diferentes puntos de corte posibles de una prueba de diagnóstico. Muestra la compensación entre sensibilidad y especificidad según el punto de corte seleccionado (cualquier aumento en la sensibilidad irá acompañado de una disminución en la especificidad). El área bajo una curva ROC (AUC) es una medida para la precisión de una prueba de diagnóstico (cuanto mayor sea el área, mejor, óptimo es 1, una prueba aleatoria tendría una curva ROC en la diagonal con un área de 0.5; para referencia: J.P. Egan. Signal Detection Theory and ROC Analysis, Academic Press, New York, 1975).

35 Resultados de AUC:

Suero:

Marcador: HeavyMethyl GSTP1

AUC: 0.51

40 Marcador: MSP HISTONA H4

AUC: 0.59

Marcador: RECEPTOR DE PROSTAGLANDINA E2 MSP

AUC: 0.52

Marcador: MSP RECEPTOR NUCLEAR HUÉRFANO NR5A2

AUC: 0.50

Orina:

Marcador: HeavyMethyl GSTP1

5 AUC: 0.58

Marcador: MSP HISTONA H4

AUC: 0.5

Marcador: RECEPTOR DE PROSTAGLANDINA E2 MSP

AUC: 0.49

10 Marcador: MSP RECEPTOR NUCLEAR HUÉRFANO NR5A2

AUC: 0.56

15 Para proporcionar una detección precisa del cáncer de próstata, se prefiere que se realice un análisis combinado de múltiples marcadores (es decir, un panel de genes). Para el análisis de muestras a base de orina, una combinación preferida de marcadores es GSTP1, RECEPTOR DE PROSTAGLANDINA E2 Y RECEPTOR NUCLEAR HUÉRFANO NR5A2 con una sensibilidad de 0.37 y una especificidad de 0.72.

Para el análisis de muestras a base de suero, la combinación más preferida de marcadores es GSTP1, HISTONA H4 y RECEPTOR NUCLEAR HUÉRFANO NR5A2 con una sensibilidad de 0.35 y una especificidad de 0.75.

Análisis MSP de los genes según la Tabla 18.

20 En el siguiente análisis, el estado de metilación del gen de acuerdo con la SEQ ID NO: 31 se analizó mediante ensayos TaqMan de MSP usando los cebadores y las sondas de acuerdo con la Tabla 6 (más adelante).

25 El estudio se realizó en 10 muestras de tejido de carcinoma de próstata, 10 de tejido de hiperplasia prostática benigna y 5 muestras de tejido prostático normal. El ADN genómico se analizó utilizando la técnica MSP después de la conversión con bisulfito. El ADN genómico total de todas las muestras se trató con bisulfito convirtiendo las citosinas no metiladas en uracilo. Las citosinas metiladas permanecieron conservadas. El tratamiento con bisulfito se realizó con modificaciones menores de acuerdo con el protocolo descrito en Olek et al. (1996).

La secuencia de interés se amplificó luego mediante cebadores específicos para ADN tratado con bisulfito y los amplificados se detectaron mediante sondas TaqMan usando plataformas TaqMan y/o Lightcycler. Los resultados se muestran en la Tabla 20 a continuación.

Los reactivos y las condiciones de ciclado fueron los siguientes:

30 SEQ ID NO: 31

Reactivo	Conc. Final
Cebador de avance	0.5 uM
Cebador de retroceso	0.5 uM
Sonda	0.4 uM
MgCl	2.5 mM
FastStart	1x
Agua	-
Plantilla	-

Reactivo		Conc. Final		
Total		-		
Temperatura (C)	Tiempo (seg)	Velocidad (c/seg)	¿Colectado?	# de Ciclos
Desnaturalización				1
95	600	20		
Fusión				50
95	10	20		
60	45	20	Individual	
Frio				1
40	30	20		

Tabla 19

Secuencia genómica SEQ ID NO:	Cebador de avance SEQ ID NO:	Cebador de retroceso SEQ ID NO:	Sonda SEQ ID NO:
31	1011	1012	1013

Tabla 20

Secuencia genómica SEQ ID NO:	Sensibilidad (%)	Especificidad
31	80	80

5

Tabla 25: Rendimiento basado en el ensayo como se deduce de los Ejemplos 1 a 5 anteriores

Marcador	ROC AUC	AUC St Err	SVM Exac.	SVM Sens	SVM Especif.
GSTP1	0.932	0.001	0.928	0.892	0.965
RECEPTOR DE PROSTAGLANDINA E2	0.921	0.001	0.849	0.829	0.871
HISTONA H4	0.918	0.001	0.801	0.880	0.719
RASSF1A	0.903	0.001	0.828	0.800	0.857

Marcador	ROC AUC	AUC St Err	SVM Exac.	SVM Sens	SVM Especif.
PROTEÍNA 16 DOMINIO PR DEDO DE ZINC	0.871	0.001	0.794	0.768	0.822
DOMINIO QUINASA 1 LIM	0.868	0.001	0.773	0.791	0.755
RECEPTOR NUCLEAR HUÉRFANO NR5A2	0.859	0.001	0.784	0.694	0.878
PROTEÍNA LIM/HOMEOBOX LHX9	0.745	0.003	0.675	0.695	0.653
PROTEÍNA MULTITRANSMEMBRANA ASOCIADA LISOSÓMICA	0.533	0.004	0.491	0.803	0.167

Tabla 26: Genes y secuencias de acuerdo con la presente divulgación

SEQ ID Ref gen:	Genómica SEQ ID NO:	En sentido convertida metilada SEQ ID NO:	Antisentido convertida metilada SEQ ID NO:	En sentido convertida no metilada SEQ ID NO:	Antisentido convertida no metilada SEQ ID NO:
NM_002314	31	120	121	238	239

NM_002314: DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1)

5 Listado de secuencias

<110> Epigenomics AG

<120> MÉTODOS Y ÁCIDOS NUCLEICOS PARA EL ANÁLISIS DE LA EXPRESIÓN GÉNICA ASOCIADA CON EL DESARROLLO DE TRASTORNOS PROLIFERATIVOS DE CÉLULAS DE PRÓSTATA

<130> 536-40 PCTEPT1

10 <140>

<150> EP03090414.8

<151> 2003-12-01

<150> EP04090040.9

<151> 2004-02-10

15 <150> EP04090187.8

<151> 2004-05-10

<150> EP04090292.6

<151> 2004-07-21

<160> 1181

20 <210> 1 - 30

000

<210> 31

<211> 2553

<212> ADN

<213> Homo Sapiens

<400> 31

ctggccaccc	ctgccctccc	ttagacctcc	agagccccca	gtgtagccac	agaggatgct	60
gttggcttca	gccccaaagaa	gacgccgctt	cctccagagg	gctaagtaag	tgggaatccc	120
cctccctact	tgtcctgggc	tccaggcagg	gccccctggtg	taaggcctgg	ggctggaagc	180
cgacccacct	aggtccaggc	tctggggcag	aactgaaact	ccttggttac	tgtcggctgc	240
agcctgggag	caggccactg	ccaaagctgt	gggtccttcc	aggacagtct	ccccatgagg	300
ccggtcctcc	acctgctgtt	tcttcacacc	tgggtggccag	ggatgtggcc	ctgggtagaa	360
cgatgattct	ccactcctgt	cattatggaa	gccaccgctg	tctcccagcc	cagccagcca	420
cctgggctgc	agagcacccc	tttcatgccc	tccgggtgcc	tcccccttct	cctgccccag	480
cctggcctttg	tcctaccctg	ctctcaggga	ggggtaccct	ggagtggggc	cagggcatgg	540
ctctcccccg	agggagttcc	tctctggctg	tccccagggc	agctctgcac	agcctcagta	600
cctggcgcac	ctcccttgac	atccttctta	gggacagtca	ggcactctgt	gtggggcact	660
caagagagcc	aggcccgtca	gcctctagct	cctgccagaa	tgcaggcctg	aggggtgagg	720
ggcggggcag	gggcagggac	aggaactccg	gcgtgctctc	catccgcaaa	ggttcactga	780
ggccccgagc	cccagccact	gagccaccaa	gtcagcctgg	gccaggcctg	ggtgccctgt	840
ctgcaatgga	ggcagagacg	gggtctcggg	gcagttctga	ggatgctggg	tgcacagcgg	900
gggcctcgcc	ggcaggaatc	acttatgctc	tctcctgggc	caagctttgt	ggatgccccag	960
cctggggccg	cggggagctg	gcaggtcagt	ggcagacact	ggtgggcaga	cctagtgtct	1020
ggtagaacag	gcatcaagga	agtggtgacc	ggagggaagc	caagtgcact	caaaccctcg	1080
ggtgagtcac	caccgccggg	tctttcacag	ctgctgaaag	tgagcaacag	tgatgaaggt	1140
ttgtgagttt	ctgcgtgagc	gagtgaatgg	accagtagca	gtttccaggt	tgtggaagag	1200
cgttccctcc	ccgggatggg	gacacttggg	tacagcaatt	cctaataccc	caccaccca	1260
ccgcccactg	cagaggtatg	cgggggccct	gcttcctgca	ggcaggagtg	aggggcactc	1320
ctgtgatgtg	gcacccctgt	gaccgaggtc	atgtgtgatc	ggtgtaaggg	caggaagcga	1380
gtcattggtc	tgcaccaggc	gtgggggctt	ctgcgagggc	aggacccaaa	gtcggcctgg	1440
cctcccggct	gcagcactcc	tttccctttc	gaattagggt	agagccctgg	gacgggaggt	1500
gccctgtaga	ccacccccct	caccaacttc	cgtcctccgc	cccacccccg	cggtgatccg	1560
gtgaactgcc	ggccccctgc	tgtgcaccga	gtggggcagt	gaccctgacg	tggcgtctcc	1620
tgccgcccct	gccaccgcca	ccacctccgg	tggcccagcc	tccgcattcc	ccacccccat	1680
ggaggaatgc	accaggcctc	ccttcctgga	tgcacccctc	acccacatgc	ttccaaaacc	1740
tggcattttc	tgtccccct	ttactcccac	cccttcccct	aggctcccag	acaaagggga	1800
agtggctgga	tcctcttaaa	gggacagtgt	cccaccagct	tactgctgaa	ctcccctcct	1860
caacccagct	tccttagtta	cagttaatta	gcattagcag	acagcccctg	agtataccc	1920
atgcaggccc	caggctgtgg	agagtttcct	gggtaggaaa	cagcccttaa	ggtccctcat	1980
ctcatccagg	tcccagtcct	tcctacctgc	ctctctccta	gattgtggcc	ctttggagcc	2040
tggttcttct	gtccctgtgt	gaccgacaca	tagcacccaa	acagtggcag	agcgggacgg	2100
acccccctagc	ctgttctctg	tgtgggtctg	taccctgacc	cagacatgcc	ccccacagc	2160
aggacccagg	ggggcacatg	tgtgcctgcg	ggttcactgg	ggcacccgca	tttggtttat	2220
tttatTTTTT	agagagaggg	tcttgctgtg	tcacccagct	ggagtgcagt	ggtgtaatca	2280
tagcacactg	cagccttcaa	ctcctgggct	caagcgatcc	tccctcccca	gcctccctag	2340
tagctgggag	tacaggaccc	actgtatcct	ggctaatttt	ttaataattt	tttaagagat	2400
ggggctcttac	tgtgttgccc	aggctggcct	caaaccctctg	gcctcaagtg	atcctccac	2460
cttcgcctcc	tgaagtgtctg	agattacagg	catgagccac	catgcccatc	ccagactgac	2520
5 atttctatat	ttgttcatcc	tggctgggca	ggg			2553

<210> 32 - 119

000

<210> 120

<211> 2553

10 <212> ADN

<213> Secuencia artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 120

ttggttatttt	ttgtttttttt	ttagatttttt	agagtttttta	gtgtagttat	agaggatggt	60
gttggttttta	gttttaagaa	gacgtcgttt	tttttagagg	gttaagtaag	tgggaattttt	120
tttttttatt	tgttttgggt	tttaggtagg	gtttttgggtg	taaggtttgg	ggttggaagt	180
cgattttattt	aggtttaggt	tttggggtag	aattgaaatt	ttttggttat	tgtcggttgt	240
agtttgggag	taggttattg	ttaaagtgtg	gggtttttttt	aggatagttt	ttttatgagg	300
tcggttttttt	atttgttgtt	tttttatatt	tgggtggttag	ggatgtggtt	ttgggtagaa	360
cgatgattttt	ttattttttgt	tattatggaa	gttatcgttg	tttttttagtt	tagttagtta	420
tttgggttgt	agagtattttt	ttttatgttt	ttcgggtgtt	ttttttttttt	tttgtttttag	480
tttgggttttg	ttttatttttg	tttttaggga	ggggtattttt	ggagtgggtt	tagggtatgg	540
tttttttttcg	agggagttttt	tttttgggtg	tttttaggggt	agttttgtat	agtttttagta	600
tttggcgtat	ttttttttgat	attttttttta	gggatagttta	ggtattttgt	gtggggtatt	660
taagagagtt	aggttcgtta	gttttttagtt	tttgtttagaa	tgtaggtttg	aggggtgagg	720
ggcggggtag	gggtagggat	aggaattttcg	gcgtgtttttt	tattcgtaaa	ggttttattga	780
ggtttcgagt	tttagttatt	gagttattaa	gttagtttggt	gttaggtttg	gggtttttgt	840
ttgtaatgga	ggtagagacg	gggtttcggg	gtagttttga	ggatgttggg	tgtatagcgg	900
gggtttcgtc	ggtaggaatt	atttatgttt	ttttttgggt	taagttttgt	ggatgttttag	960
tttgggggtcg	cggggagttg	gtaggttagt	ggtagatatt	ggtgggtaga	tttagtgttt	1020
ggtagaatag	gtattaagga	agtgggtatc	ggagggaagt	taagtgtatt	taaatttttcg	1080
ggtgagttat	tatcgtcggg	tttttttatag	ttgttgaaag	tgagtaatag	tgatgaagggt	1140
ttgtgagttt	ttgcgtgagc	gagtgaatgg	attagtagta	gttttttaggt	tgtggaagag	1200
cgttttttttt	tcgggatggg	gatattttggt	tatagtaatt	tttaattttt	tattttattta	1260
tcgtttattg	tagaggtagt	cgggggtttt	gtttttttgta	ggtaggagtg	aggggtattt	1320
ttgtgatgtg	gtattttttgt	gatcgaggtt	atgtgtgatc	ggtgtaaggg	taggaagcga	1380
gttattgggt	tgtattaggc	gtgggggttt	ttgcgagggt	aggatttaaa	gtcggtttg	1440
tttttcggtt	gtagtatttt	ttttttttttc	gaattagggt	agagttttg	gacgggaggt	1500
gttttgtaga	ttattttttt	tattaatttt	cgttttttcgt	tttatttttcg	cgggtgattcg	1560
gtgaattgtc	ggttttttgt	tgtgtatcga	gtggggtagt	gattttgacg	tggcgttttt	1620
tgtcgttttt	gttatcgtta	ttatttttcgg	tggtttagtt	ttcgtatttt	ttatttttat	1680
ggaggaatgt	attaggtttt	tttttttggg	tgtatttttt	atttatatgt	ttttaaattt	1740
tgggtattttt	tgtttttttt	ttattttttat	ttttttttttt	aggtttttag	ataaagggga	1800
agtggttgga	ttttttttaaa	gggatagtgt	tttattagtt	tattgttgaa	ttttttttttt	1860
taatttttagt	tttttagtta	tagttaatta	gtatttagtag	atagtttatg	agtgatattt	1920
atgtaggttt	taggttgtgg	agagtttttt	gggtaggaaa	tagtttttaa	ggtttttttat	1980
tttatttagg	tttttagttt	ttttatttgt	tttttttttta	gattgtggtt	ttttggagtt	2040
tgggttttttt	gtttttgtgt	gatcgatata	tagtatttaa	atagtggtag	agcgggacgg	2100
atttttttagt	ttgtttttttg	tgtgggtttg	tatttttgatt	tagatatgtt	tttttatagt	2160
aggatttagg	gggttatatg	tgtgtttgcg	ggtttattgg	ggtattcgt	tttggtttat	2220
tttattttttt	agagagaggg	ttttgttgtg	ttatttagtt	ggagtgtagt	ggtgtaatta	2280
tagtatattg	tagtttttaa	tttttgggtt	taagcgattt	tttttttttta	gttttttttag	2340
tagttgggag	tataggattt	attgtatttt	ggttaatttt	ttaataattt	tttaagagat	2400
ggggtttttat	tgtgttggtt	aggttggttt	taaattttttg	gttttaagtg	atttttttat	2460
tttcgtttttt	tgaagtgttg	agattatagg	tatgagttat	tatgtttatt	ttagattgat	2520
attttttatat	ttgtttattt	tggttgggta	ggg			2553

<210> 121

5 <211> 2553

<212> ADN

<213> Secuencia artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

10 <400> 121

ttttgttttag	ttaggatgaa	taaatataga	aatgttagtt	tgggatgggt	atggtggttt	60
atgtttgtaa	ttttagtatt	ttaggaggcg	aaggtgggag	gattatttga	ggtagaggt	120
ttgaggttag	tttgggtaaat	atagtaagat	tttatttttt	aaaaaattat	taaaaaatta	180
gttaggatat	agtgggtttt	gtattttttag	ttattaggga	ggttggggag	ggaggatcgt	240
ttgagtttag	gagttgaagg	ttgtagtgtg	ttatgattat	attattgtat	tttagttggg	300
tgatatagta	agattttttt	tttaaaaaat	aaaataaatt	aaatgcgggt	gttttagtga	360
attcgttaggt	atatatgtgt	ttttttgggt	tttgttgtgg	gggggtatgt	ttgggttagg	420
gtatagattt	atatagagaa	taggttaggg	ggttcgtttc	gttttgttat	tgtttgggtg	480
ttatgtgtcg	gttatatagg	gatagaagaa	ttaggtttta	aagggttata	atttaggaga	540
gaggtaggt	ggaaagattg	ggatttggt	gagatgagg	attttaagg	ttgtttttta	600
tttaggaaat	tttttatagt	ttggggtttg	tatgggtatt	atttatgggt	tgtttgttaa	660
tgtaatttaa	ttgttaattag	ggaattgggg	ttgaggagg	gagtttagta	gtaagttggt	720
gggatattgt	ttttttaaga	ggatttagtt	attttttttt	tgtttgggag	tttaggggaa	780
gggggtgggag	taaaggggga	gtagaaaatg	ttagggtttg	gaagtatgtg	ggtgaggggt	840
gtattttagga	agggaggttt	ggtgtatttt	tttatggggg	tggggaatgc	ggaggttggg	900
ttatcgagg	tgggtggcgt	ggtagggcg	gtaggagacg	ttacgttagg	gttattgttt	960
tattcgggtg	atagtagggg	gtcggtagtt	tatcggatta	tcgcgggggt	ggggcgagg	1020
acggaagttg	gtgagggggg	tggtttatag	ggtatttttc	gttttaggg	tttaatttaa	1080
ttcgaaaggg	aaaggagtgt	tgtagtcggg	aggtaggtc	gattttgggt	tttgttttcg	1140
tagaagtttt	tacgtttggt	gtagattaat	gattcgtttt	ttgtttttat	atcgattata	1200
tatgatttcg	gttatagggg	tgttatatta	taggagtgtt	ttttattttt	gtttgttagga	1260
agtagggttt	tcgtatatatt	ttgtagtggg	cgggtgggtg	gtgggggatt	aggaattgtt	1320
gtaattaagt	gtttttattt	cggggaggga	acgttttttt	ataatttgga	aattgttatt	1380
ggtttatatta	ttcgtttacg	tagaaaatta	ttaaattttta	ttattgttgt	ttatttttag	1440
tagttgtgaa	agattcggcg	gtgatgattt	attcagggtt	ttgagtgtat	ttggtttttt	1500
ttcggttatt	atttttttga	tgtttgtttt	attagatatt	aggtttgttt	attagtgttt	1560
gttattgatt	tgtagttttt	tcgcggtttt	aggttgggt	tttataaagt	ttggttttag	1620
agagagtata	agtgattttt	gtcggcgagg	ttttcgttgt	gtatttagta	tttttagaat	1680
tgtttcgaga	tttcgttttt	gtttttattg	tagatagggt	atttaggttt	ggtttaggtt	1740
gatttggtgg	tttagtggtt	ggggttcggg	gttttagtga	atttttgcgg	atggagagta	1800
cgtcggagtt	tttgtttttg	tttttgtttc	gttttttatt	tttttaggtt	gtattttggt	1860
aggagttaga	ggttgacggg	tttggttttt	ttgagtgttt	tatatagagt	gtttgattgt	1920
ttttaagaag	gatgttaagg	gaggtgcgtt	aggatttag	gttgtgtaga	gttgttttgg	1980
ggatagttag	agaggaattt	tttcggggga	gagttatgtt	ttggttttat	tttaggggat	2040
ttttttttga	gagtagggta	ggataaaagt	aggttgggtt	aggagaagg	ggaggtattc	2100
ggagggtag	aaaggggtgt	tttgtagtgt	agggtggttg	ttgggttggg	agatagcggg	2160
ggttttttata	atgataaggag	tggagaatta	tcgtttttatt	tagggttata	tttttggtta	2220
ttaggtgtga	agaaatagta	ggtggaggat	cggttttatg	gggagattgt	tttggaagga	2280
tttatagttt	tggtagtggg	ttgttttttag	gttgtagtgc	atagtaatta	aggagtttta	2340
gttttgtttt	agagtttgga	tttaggtggg	tcggttttta	gttttaggtt	ttatattagg	2400
ggttttgttt	ggagtttagg	ataagtaggg	agggggattt	ttattttatt	agtttttttg	2460
aggaagcggc	gttttttttg	ggttgaagtt	aatagtattt	tttgtggtta	tattgggggt	2520
tttggaggtt	taaggggagg	taggggtggt	tag			2553

<210> 122 - 237

000

5 <210> 238

<211> 2553

<212> ADN

<213> Secuencia artificial

<220>

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

<400> 238

ttgggttattt	ttgtttttttt	ttagatttttt	agagtttttta	gtgtagttat	agaggatggt	60
gttggttttta	gttttaagaa	gatgttggtt	tttttagagg	gttaagtaag	tgggaattttt	120
ttttttttatt	tgttttgggt	tttaggtagg	gtttttgggtg	taagggttgg	gggttgaagt	180
tgattttattt	aggttttaggt	tttggggtag	aattgaaatt	ttttggttat	tgttgggtgt	240
agtttgggag	taggttattg	ttaaagtgt	gggtttttttt	aggatagttt	ttttatgagg	300
ttgggtttttt	atttggtgtt	tttttatatt	tgggtggttag	ggatgtggtt	ttgggtagaa	360
tgatgattttt	ttattttttgt	tattatggaa	gttattgttg	tttttttagtt	tagttagtta	420
tttgggttgt	agagtattttt	ttttatgttt	tttgggtgtt	ttttttttttt	tttgttttag	480
tttgggttttg	ttttatttttg	tttttaggga	gggttattttt	ggagtggggt	tagggtagtg	540
ttttttttttg	agggagttttt	tttttggttg	tttttagggt	agttttgtat	agtttttagta	600
tttgggtgtat	tttttttgat	atttttttta	gggatagtta	ggtattttgt	gtggggtatt	660
taagagaggtt	aggtttgtta	gttttttagtt	tttgttagaa	tgtaggtttg	aggggtgagg	720
gggtggggtag	gggttagggat	aggaatttttg	gtgtgtttttt	tatttgtaaa	ggttttattga	780
gggttttgagt	tttagttatt	gagttattaa	gttagtttggt	gttaggtttg	gggtttttgt	840
ttgtaatgga	ggtagagatg	gggttttggg	gtagttttga	ggatgttggg	tgtatagtgg	900
gggttttggt	ggtaggaatt	atttatgttt	ttttttgggt	taagttttgt	ggatgttttag	960
tttgggggttg	tggggagttg	gtaggttagt	ggtagatatt	ggtgggtaga	tttagtggtt	1020
ggtagaatag	gtattaagga	agtgtgatt	ggagggaagt	taagtgtatt	taaattttttg	1080
ggtgagttat	tattgttggt	tttttttatag	ttgttgaaa	tgagtaatag	tgatgaaggt	1140
ttgtgagttt	ttgtgtgagt	gagtgaatgg	attagtagta	gttttttaggt	tgtggaagag	1200
tgttttttttt	ttgggatggg	gatatttggg	tatagtaatt	tttaatttttt	tattttattta	1260
ttgttttattg	tagaggtatg	tgggggtttt	gttttttga	ggtaggagtg	aggggtattt	1320
ttgtgatgtg	gtatttttgt	gattgaggtt	atgtgtgatt	ggtgtaaggg	taggaagtga	1380
gttattgggt	tgtattaggt	gtgggggttt	ttgtgaggt	aggattttaa	gttgggttg	1440
ttttttgggt	gtagtatttt	ttttttttttt	gaattaggtt	agagtttttg	gatgggaggt	1500
gttttgtaga	ttatttttttt	tattaattttt	tgttttttgt	tttattttttg	tgggtgatttg	1560
gtgaattgtt	gggttttttgt	tgtgtattga	gtggggtagt	gatttttgatg	tgggtgttttt	1620
tgttggttttt	gttattgtta	ttatttttttg	tgggttagtt	tttgtattttt	ttattttttat	1680
ggaggaatgt	attaggtttt	tttttttgga	tgtatttttt	atttatatgt	ttttaaatttt	1740
tgggtattttt	tgttttttttt	ttattttttat	ttttttttttt	aggttttttag	ataaagggga	1800
agtgggttga	ttttttttaa	gggataggt	tttattagtt	tattgttgaa	ttttttttttt	1860
taatttttagt	tttttagtta	tagttaatta	gtatttagtag	atagtttatg	agtgatattt	1920
atgtaggttt	taggttgtgg	agagttttttt	gggtaggaaa	tagtttttta	gggttttttat	1980
tttattttagg	tttttagtttt	ttttattttgt	tttttttttta	gattgtggtt	ttttggagtt	2040
tgggttttttt	gtttttgtgt	gattgatata	tagtatttaa	atagtggtag	agtgggatgg	2100
atttttttagt	ttgtttttttg	tgtgggtttg	tatttttgatt	tagatatgtt	tttttatagt	2160
aggattttagg	ggggtatatg	tgtgtttgtg	gggtttattgg	ggtatttgta	tttgggttat	2220
tttattttttt	agagagaggg	ttttgttgtg	ttatttagtt	ggagtgtagt	gggtgaatta	2280
tagtatattg	tagtttttta	tttttggttt	taagtgattt	tttttttttta	gttttttttag	2340
tagttgggag	tataggattt	attgtattttt	gggttaatttt	tttaataattt	tttaagagat	2400
gggggttttat	tgtgttggtt	aggttggttt	taaattttttg	gttttaagtg	attttttttat	2460
ttttgtttttt	tgaagtgttg	agattatagg	tatgagttat	tatgtttatt	ttagattgat	2520
attttttatat	ttgttttattt	tgggtgggta	ggg			2553

<210> 239

<211> 2553

<212> ADN

5 <213> Secuencia artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 239

```

ttttgttttag ttaggatgaa taaatataga aatgttagtt tgggatgggt atgggtggttt 60
atgtttgtaa ttttagtatt ttaggaggtg aagggtgggag gattatttga ggtagaggt 120
ttgaggttag tttgggtaat atagtaagat tttattttttt aaaaaattat taaaaaatta 180
gttaggatat agtgggtttt gtatttttttag ttatttaggga gggtggggag ggaggattgt 240
ttgagtttag gagttgaagg ttgtagtgtg ttatgattat attattgtat tttagttggg 300
tgatatagta agattttttt tttaaaaaat aaaataaatt aaatgtgggt gtttttagtga 360
atgtttaggt atatatgtgt ttttttgggt tttgttgtgg ggggtatgt ttgggttagg 420
gtatagattt atatatagaa taggttaggg gggttgtttt gttttgttat tgtttgggtg 480
ttatgtgttg gttatatagg gatagaagaa ttaggtttta aagggttata atttaggaga 540
gaggtaggta ggaaagattg ggatttggat gagatgaggg attttaaggg ttgtttttta 600
tttaggaaat tttttatagt ttgggtttt tatgggtatt atttatgggt tgtttgttaa 660
tgtaaattaa ttgtaattag ggaattgggg ttgaggaggg gagtttagta gtaagtgggt 720
gggatattgt ttttttaaga ggatttagtt attttttttt tgtttgggag tttaggggaa 780
ggggtgggag taaaggggga gtagaaaatg ttagggtttg gaagtatgtg ggtgaggggt 840
gtatttagga agggaggttt ggtgtatttt tttatggggg tggggaatgt ggaggttggg 900
ttattggagg tgggtggtgt ggtaggggtg gtaggagatg ttatgttagg gttattgttt 960
tatttggtgt atagtagggg gttggtagt ttattggatta ttgtgggggt ggggtggagg 1020
atggaagttg gtgagggggg tgggttatag ggtattttttt gtttttaggt ttttaatttaa 1080
tttgaaaggg aaaggagtgt tgtagtggg aggttaggtt gattttgggt tttgtttttg 1140
tagaagtttt tatgtttggt gtagattaat gatttgtttt ttgtttttat attgattata 1200
tatgattttg gttatagggg tgttatatta taggagtgtt ttttattttt gttttagga 1260
agtagggttt ttgtatatatt ttgtagtggg tgggtgggtg gtgggggatt aggaattgtt 1320
gtaattaagt gtttttattt tggggaggga atgttttttt ataatttga aattgttatt 1380
ggtttattta tttgtttatg tagaaaattt taaattttta ttattgttgt ttttttttag 1440
tagttgtgaa agatttgggt gtgatgattt atttgagggt ttgagtgtat ttgggttttt 1500
tttggttatt atttttttga tgtttgtttt attagatatt aggtttgttt attagtgttt 1560
gttattgatt tgttagtttt ttgtggtttt aggttgggta tttataaagt ttgggttagg 1620
agagagtata agtgattttt gttggtgagg tttttgttgt gtatttagta tttttagaat 1680
tgttttgaga ttttgttttt gtttttattg tagatagggt atttaggttt gggttaggtt 1740
gatttggtgg tttagtgggt ggggtttggg gtttttagtga atttttgtgg atggagagta 1800
tgttggagtt tttgtttttg tttttgtttt gttttttatt ttttaggttt gtattttggt 1860
aggagttaga ggttgatggg tttggttttt ttgagtgttt tatatataggt gtttgattgt 1920
ttttaagaag gatgttaagg gaggtgtgtt aggtattgag gttgtgtaga gttgttttgg 1980
ggatagttag agaggaattt ttttggggga gagttatgtt ttgggtttat tttagggtat 2040
ttttttttga gagtagggta ggataaagtt aggttggggg aggagaaggg ggaggtattt 2100
ggagggtatg aaagggtgt tttgtagttt aggtggttgg ttgggttggg agatagtgggt 2160
ggtttttata atgataggag tggagaatta ttgttttatt tagggttata tttttggtta 2220
ttaggtgtga agaaatagta ggtggaggat tgggtttatg gggagattgt tttggaagga 2280
tttatagttt tggtagtggg ttgttttttag gttgtagtgt atagtaatta aggagtttta 2340
gttttgtttt agagtttggg tttaggtggg ttgggtttta gtttttaggt ttatattagg 2400
ggttttgttt ggagtttagg ataagtaggg agggggattt ttattttatt agttttttgg 2460
aggaagtggg gtttttttgg ggttgaagtt aatagtattt tttgtggtta tattgggggt 2520
tttgagggtt taagggaggg taggggtggt tag 2553

```

<210> 240 - 363

000

<400> 364

5 agggagggtt ggtgtattt 20

<210> 365

<211> 22

<212> ADN

<213> Secuencia artificial

10 <220>

<223> <223> Cebador de detección para DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1)

	<400> 365 aaaccctact tcctacaaac aa 21	22
	<210> 366 - 701 000	
5	<210> 702 <211> 16 <212> ADN <213> Secuencia artificial <220>	
10	<223> Oligonucleótido de detección para DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) <400> 702 tatcggatta tcgcgg 16 <210> 703 <211> 17	
15	<212> ADN <213> Secuencia artificial <220> <223> Oligonucleótido de detección para DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) <400> 703	
20	attggattat tgtgggg 17 <210> 704 <211> 18 <212> ADN <213> Secuencia artificial	
25	<220> <223> Oligonucleótido de detección para DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1) <400> 704 gtcggtagtt tatcggat 18 <210> 705	
30	<211> 18 <212> ADN <213> Secuencia artificial <220> <223> Oligonucleótido de detección para DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1)	
35	<400> 705 gttgtagtt tattggat 18	

<210> 706
 <211> 16
 <212> ADN
 <213> Secuencia artificial
 5 <220>
 <223> Oligonucleótido de detección para DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1)
 <400> 706
 taggagacgt tacgtt 16
 <210> 707
 10 <211> 17
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> Oligonucleótido de detección para DOMINIO QUINASA 1 LIM (EC 2.7.1.37) (LIMK-1)
 15 <400> 707
 agatgttatg ttaggt 17
 <210> 708 - 927
 000
 <210> 928
 20 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 25 <400> 928
 gggccgcggc 10
 <210> 929
 <211> 10
 <212> ADN
 30 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 929
 ccccgcgggg 10
 35 <210> 930
 <211> 10
 <212> ADN

<213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 930
 5 cgcgggggcg 10
 <210> 931
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 10 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 931
 gcgcgccgcg 10
 <210> 932
 15 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 20 <400> 932
 gcggggcggc 10
 <210> 933
 <211> 10
 <212> ADN
 25 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 933
 gcgccgacgt 10
 30 <210> 934
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 35 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 934
 cgggacgcga 10

<210> 935
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 5 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 935
 ccgcgatcgc 10
 <210> 936
 10 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 15 <400> 936
 tggccgccga 10
 <210> 937
 <211> 10
 <212> ADN
 20 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 937
 tgcgacgccg 10
 25 <210> 938
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 30 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 938
 atcccgcccg 10
 <210> 939
 <211> 10
 35 <212> ADN
 <213> Secuencia artificial
 <220>

<223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 939
 gcgcatgcgg 10
 <210> 940
 5 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 10 <400> 940
 gcgacgtgcg 10
 <210> 941
 <211> 10
 <212> ADN
 15 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <220>
 <221> inseguro
 20 <222> (7, 9)
 <223> base desconocida
 <400> 941
 gccgcgngng 10
 <210> 942
 25 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 30 <220>
 <221> inseguro
 <222> (8, 9)
 <223> base desconocida
 <400> 942
 35 gcccgcgngng 10
 <210> 943
 <211> 10

<212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 5 <400> 943
 agcggccgcg 10
 <210> 944
 <211> 10
 <212> ADN
 10 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 944
 ctcccacgcg 10
 15 <210> 945
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 20 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 945
 gaggtgcgcg 10
 <210> 946
 <211> 10
 25 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 946
 30 aggggacgcg 10
 <210> 947
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 35 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 947

gagaggcgcg 10
 <210> 948
 <211> 10
 <212> ADN
 5 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 948
 gccccgcga 10
 10 <210> 949
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 <220>
 15 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 949
 cggggcgcga 10
 <210> 950
 <211> 10
 20 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 950
 25 ggggacgcga 10
 <210> 951
 <211> 10
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 951
 accccaccgcg 10
 <210> 952
 35 <211> 24
 <212> ADN
 <213> Secuencia artificial

<220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 952
 agcactctcc agcctctcac cgac 24
 5 <210> 953
 <211> 12
 <212> ADN
 <213> Secuencia artificial
 <220>
 10 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 953
 ccgggtcggg ga 12
 <210> 954
 <211> 24
 15 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 954
 20 accgacgtcg actatccatg aacc 24
 <210> 955
 <211> 12
 <212> ADN
 <213> Secuencia artificial
 25 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 955
 ccgggggttca tg 12
 <210> 956
 30 <211> 24
 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 35 <400> 956
 aggcaactgt gctatccgag tgac 24
 <210> 957

<211> 12
 <212> ADN
 <213> Secuencia artificial
 <220>
 5 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 957
 ccgggtcact cg 12
 <210> 958 - 975
 000
 10 <210> 976
 <211> 21
 <212> ADN
 <213> Secuencia artificial
 <220>
 15 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 976
 cgtttttcgt ttattttcg c 21
 <210> 977
 <211> 18
 20 <212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 977
 25 gacaaaaaac gccacgtc 18
 <210> 978
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 30 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 978
 ccgacaattc accgaatcac cg 22
 <210> 979 - 1010
 35 000
 <210> 1011
 <211> 21

<212> ADN
 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 5 <400> 1011
 cgtttttcgt ttattttcg c 21
 <210> 1012
 <211> 18
 <212> ADN
 10 <213> Secuencia artificial
 <220>
 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 1012
 gacaaaaaac gccacgtc 18
 15 <210> 1013
 <211> 22
 <212> ADN
 <213> Secuencia artificial
 <220>
 20 <223> ADN genómico tratado químicamente (Homo sapiens)
 <400> 1013
 ccgacaattc accgaatcac cg 22
 <210> 1014 - 1181
 000
 25

REIVINDICACIONES

1. Un método para detectar cáncer de próstata y/o para distinguir entre o dentro de trastornos proliferativos de células de próstata en un sujeto, que comprende:

5 a) poner en contacto el ADN genómico derivado de una muestra biológica que se ha obtenido del sujeto, o un fragmento del mismo, con un reactivo o una pluralidad de reactivos para distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas dentro de al menos una secuencia diana del ADN genómico, o fragmento del mismo, en donde la secuencia diana comprende o hibrida en condiciones restrictivas con al menos 16 nucleótidos contiguos de la secuencia de acuerdo con la SEQ ID NO: 31 o un complemento de la misma, comprendiendo dichos nucleótidos contiguos al menos una secuencia de dinucleótidos CpG; y

10 b) determinar, con base al menos en parte en dicha distinción, el estado de metilación de al menos una secuencia de dinucleótidos CpG diana, o un promedio, o un valor que refleje un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG diana,

en donde la muestra biológica se selecciona del grupo que consiste en biopsia y fluidos corporales.

15 2. El método de la reivindicación 1, en donde distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas dentro de la secuencia diana comprende convertir bases de citosina no metiladas dentro de la secuencia diana en uracilo u otra base que sea detectablemente diferente a la citosina en términos de propiedades de hibridación.

20 3. El método de la reivindicación 1, en donde distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas dentro de la secuencia diana comprende el uso de al menos una molécula de ácido nucleico o molécula de ácido nucleico peptídico (PNA) que comprende, en cada caso, una secuencia contigua de al menos 9 nucleótidos de longitud que es complementaria a o hibrida en condiciones restrictivas o moderadamente restrictivas a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121 y sus complementos, preferiblemente en donde la secuencia contigua comprende al menos una secuencia de dinucleótidos CpG, TpG o CpA.

25 4. El método de la reivindicación 3, que comprende el uso de al menos dos de tales moléculas de ácido nucleico, o moléculas de ácido nucleico peptídico (PNA) como oligonucleótidos cebadores para la amplificación de secuencias seleccionadas del grupo que consiste en SEQ ID NO: 120, 121 y secuencias complementarias a las mismas, y regiones de las mismas que comprenden o hibridan en condiciones restrictivas con los cebadores.

5. Un método para detectar cáncer de próstata y/o para distinguir entre o en medio de trastornos proliferativos de células de próstata en un sujeto, que comprende:

30 a. extraer o aislar de otra manera el ADN genómico de una muestra biológica que se ha obtenido de un sujeto y comprende el ADN genómico del sujeto;

b. tratar el ADN genómico de a), o un fragmento del mismo, con uno o más reactivos para convertir bases de citosina que no están metiladas en la posición 5 del mismo a uracilo u otra base que sea detectablemente diferente a la citosina en términos de propiedades de hibridación;

35 c. poner en contacto el ADN genómico tratado, o el fragmento tratado del mismo, con una enzima de amplificación y al menos dos cebadores que comprenden, en cada caso, una secuencia contigua de al menos 9 nucleótidos que es complementaria o se hibrida en condiciones restrictivas o moderadamente restrictivas con una secuencia seleccionado del grupo que consiste en SEQ ID NO: 120, 121 y sus complementos, en donde el ADN genómico tratado o el fragmento del mismo se amplifica para producir al menos un amplificado, o no se amplifica; y

40 d. determinar, con base en la presencia o ausencia de, o en una propiedad de dicho amplificado, el estado de metilación de al menos un dinucleótido CpG de una secuencia de acuerdo con la SEQ ID NO: 31, o un promedio, o un valor que refleje un estado de metilación promedio de una pluralidad de dinucleótidos CpG de una secuencia de acuerdo con SEQ ID NO: 31, por lo que se proporciona al menos uno de detección, o detección y distinción entre trastornos proliferativos de células de próstata, al menos en parte.

45 6. El método de la reivindicación 5, en donde la amplificación comprende el uso de al menos un método seleccionado del grupo que consiste en el uso de una ADN polimerasa resistente al calor como enzima de amplificación; uso de una polimerasa que carece de actividad 5'-3'exonucleasa; uso de una reacción en cadena de la polimerasa (PCR); generación de una molécula de ácido nucleico amplificada que porta etiquetas detectables; y combinaciones de los mismos.

50 7. El método de la reivindicación 5, en donde al menos una de tales moléculas de ácido nucleico hibridante o molécula de ácido nucleico peptídico se une a una fase sólida, preferiblemente en donde una pluralidad de tales moléculas de ácido nucleico hibridante o moléculas de ácido nucleico peptídico se unen a una fase sólida en forma de un arreglo de ácidos nucleicos o ácidos nucleicos peptídicos seleccionado del grupo de arreglos que consiste en lineal o sustancialmente así, hexagonal o sustancialmente así, rectangular o sustancialmente así, y combinaciones de los

mismos, y en donde el método preferiblemente comprende además extender al menos una de tales moléculas de ácido nucleico hibridada en al menos una base de nucleótidos.

8. El método de la reivindicación 5, en donde determinar en d. comprende secuenciar el amplificado o poner en contacto en c. comprende el uso de cebadores específicos de metilación.

- 5 9. Un conjunto de oligómeros que comprende al menos dos oligonucleótidos, cada uno de los cuales comprende una secuencia de al menos 15 nucleótidos contiguos que es complementaria o se hibrida en condiciones restrictivas con una secuencia de ADN seleccionada del grupo que consiste en SEQ ID NO: 120, 121, 238, 239 y sus complementos, en donde los al menos dos oligonucleótidos son adecuados para el análisis de un patrón de metilación de citosina en la secuencia genómica de acuerdo con SEQ ID NO: 31, y en donde el conjunto se selecciona preferiblemente del grupo que consiste en SEQ ID NO: 364 y 365, SEQ ID NO: 976-978, y SEQ ID NO: 1011-1013.

- 10 10. Uso de una molécula de ácido nucleico que se hibrida en condiciones restrictivas con toda o una porción de al menos 9 nucleótidos de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121, 238, 239 y sus complementos, para la detección de y/o para la diferenciación entre trastornos proliferativos de células de próstata.

- 15 11. Uso de un kit que comprende: al menos uno de un reactivo de bisulfito, o una enzima de restricción sensible a la metilación; y al menos una molécula de ácido nucleico o una molécula de ácido nucleico peptídico que comprende, en cada caso, una secuencia contigua de al menos 9 nucleótidos que es complementaria a o se hibrida en condiciones restrictivas a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 120, 121 y sus complementos, para detectar o distinguir entre los trastornos proliferativos de células de próstata de un sujeto.

- 20 12. El uso de la reivindicación 11, en donde el kit comprende además reactivos estándar para realizar un ensayo de metilación seleccionado del grupo que consiste en MS-SNuPE, MSP, MethyLight, HeavyMethyl, secuenciación de ácidos nucleicos y combinaciones de los mismos.

13. Un método para detectar cáncer de próstata y/o para distinguir entre o en medio de trastornos proliferativos de células de próstata en un sujeto, que comprende:

a) extraer o aislar de otra manera el ADN genómico de una muestra biológica que se ha obtenido del sujeto;

- 25 b) poner en contacto el ADN genómico de a), o un fragmento del mismo, que comprende al menos 16 nucleótidos contiguos de una secuencia de SEQ ID NO: 31 y secuencias que se hibridan en condiciones restrictivas la mismo con una o más enzimas de restricción sensibles a la metilación, en donde el ADN genómico, con respecto a cada motivo de reconocimiento de escisión del mismo, se escinde de ese modo para producir fragmentos de escisión o no se escinde de ese modo; y

- 30 c) determinar, con base en la presencia o ausencia de, o en la propiedad de al menos uno de dichos fragmentos de escisión, el estado de metilación de al menos un dinucleótido CpG de una secuencia SEQ ID NO: 31, o un promedio, o un valor que refleje un estado de metilación promedio de una pluralidad de dinucleótidos CpG de SEQ ID NO: 31, en donde la muestra biológica se selecciona del grupo que consiste en biopsia y fluidos corporales.

- 35 14. El método de una cualquiera de las reivindicaciones 1-8 o 13, en donde el fluido corporal es eyaculado, orina o sangre.

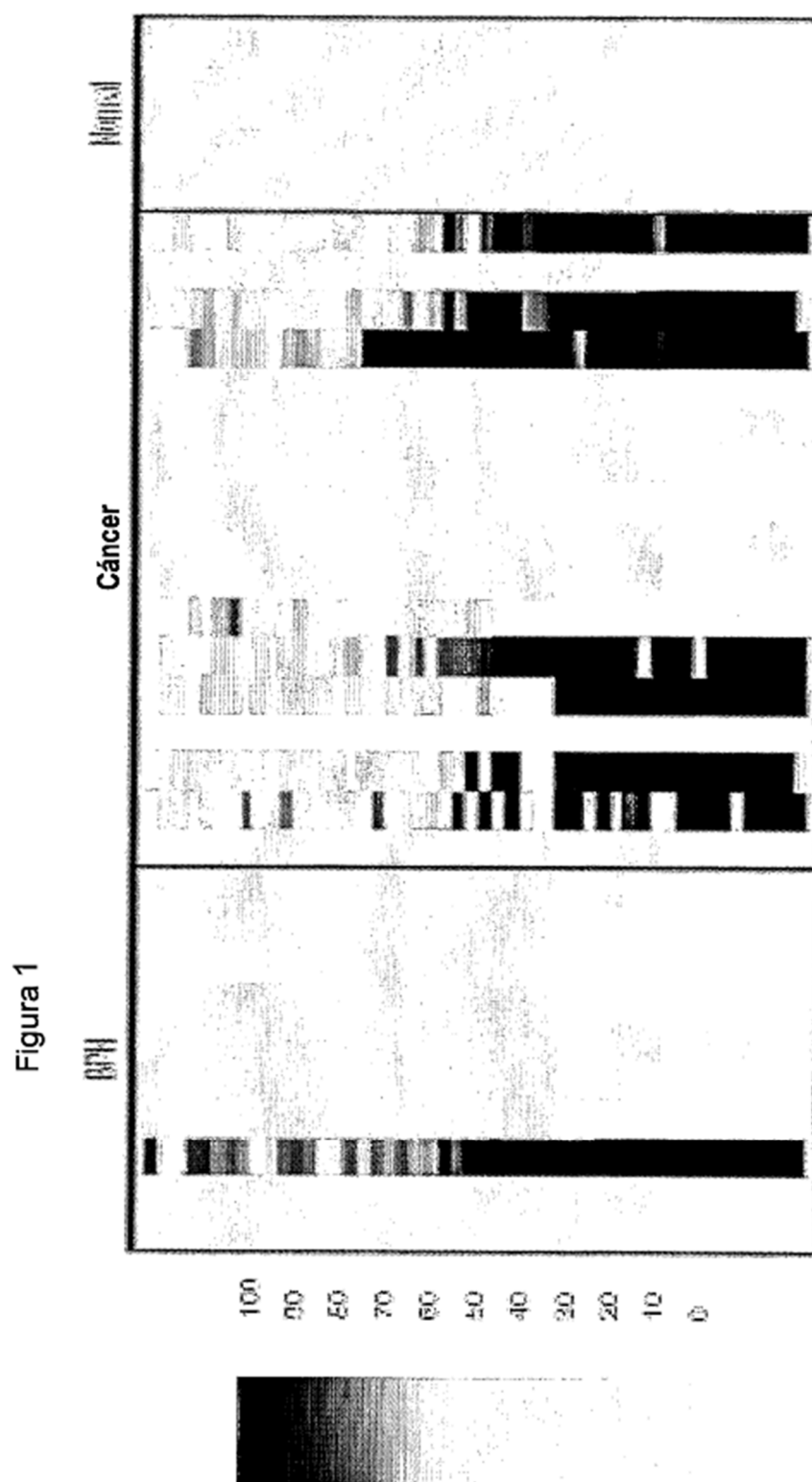


Figura 2

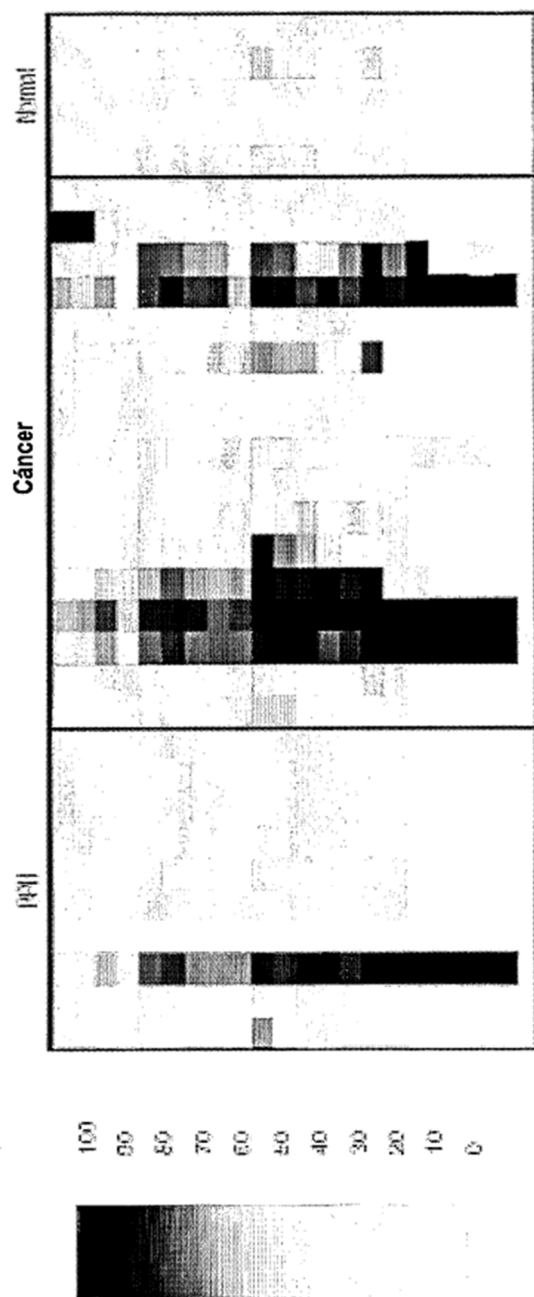


Figura 3

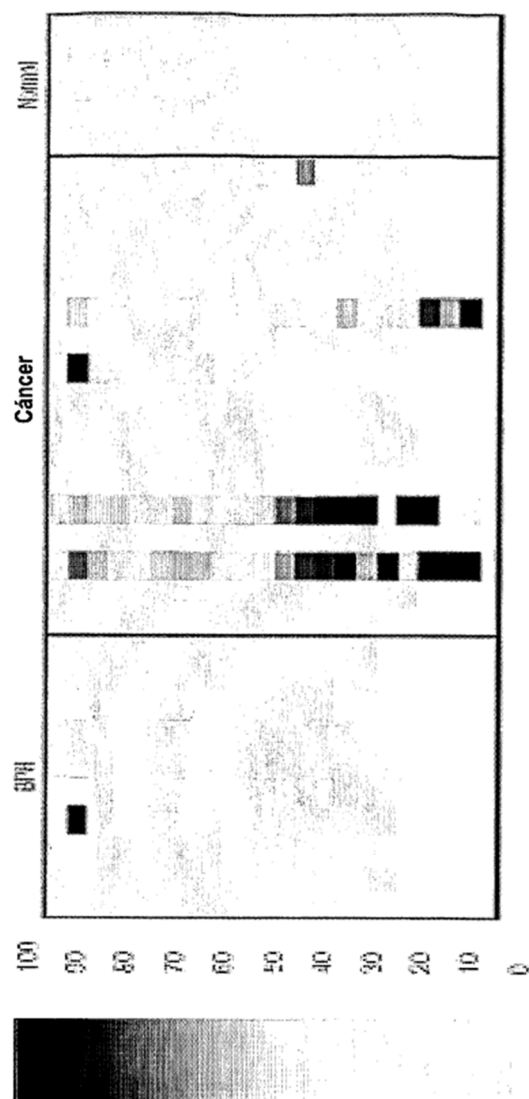


Figura 4

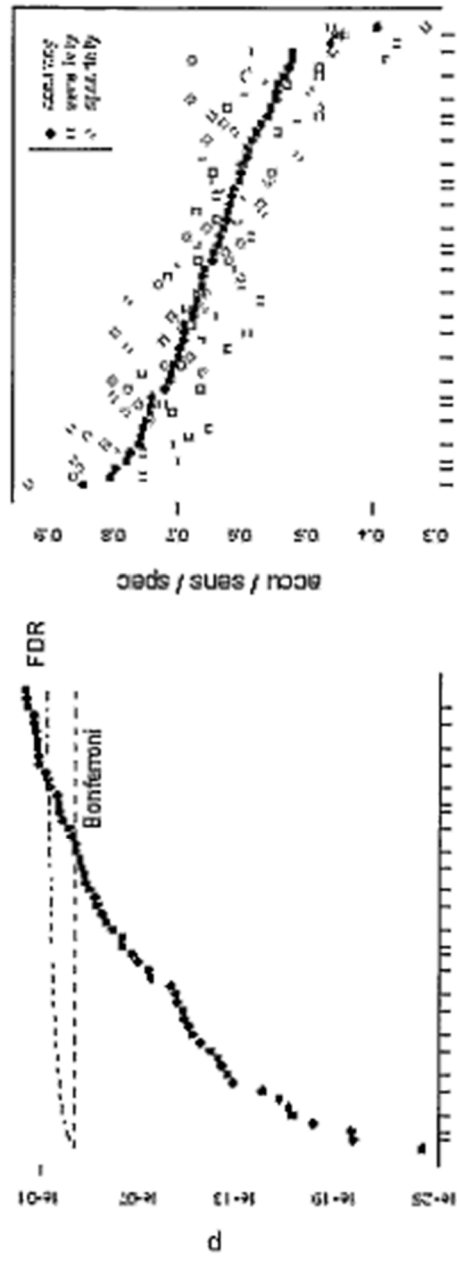


Figura 5

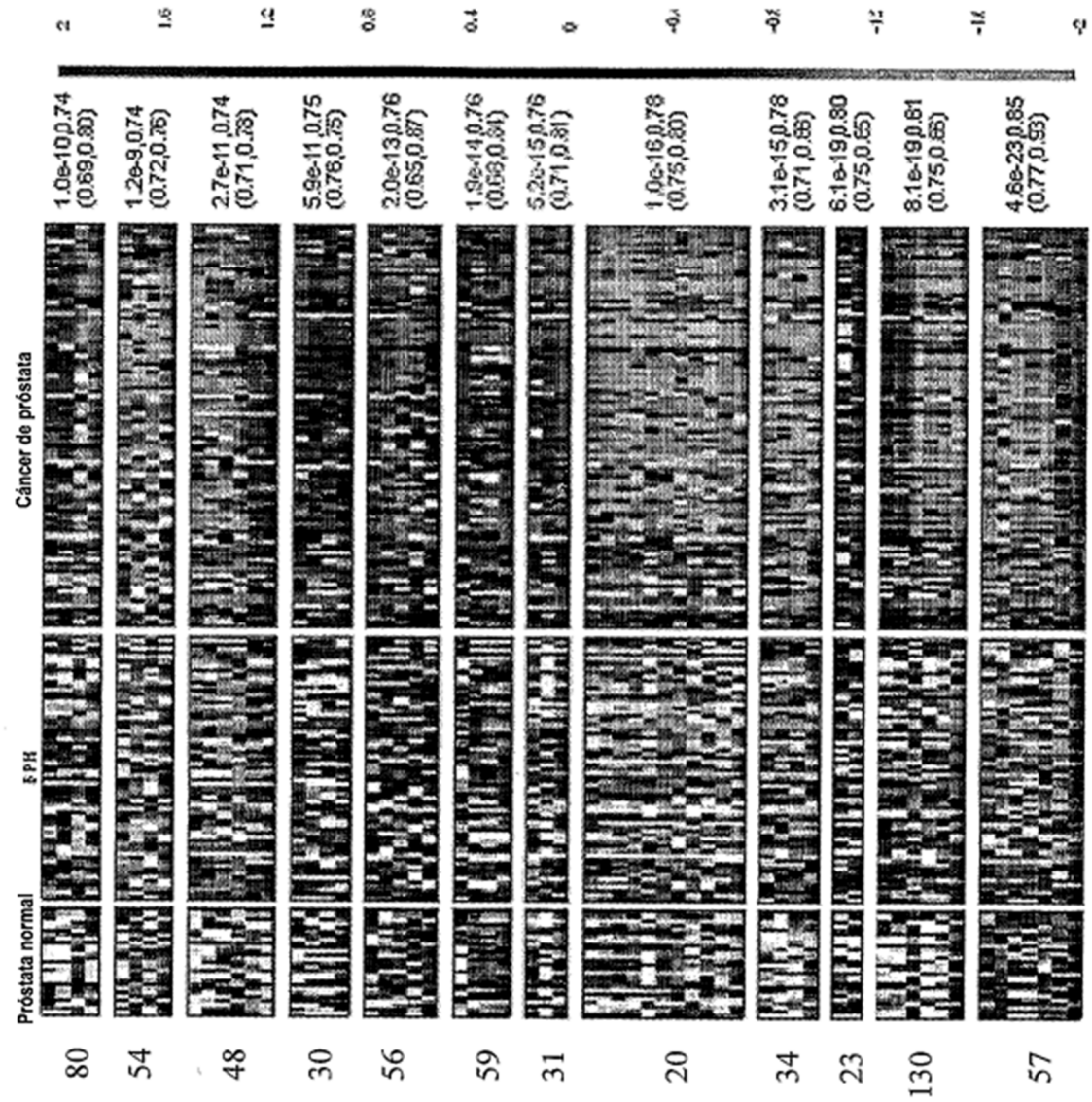


Figura 6

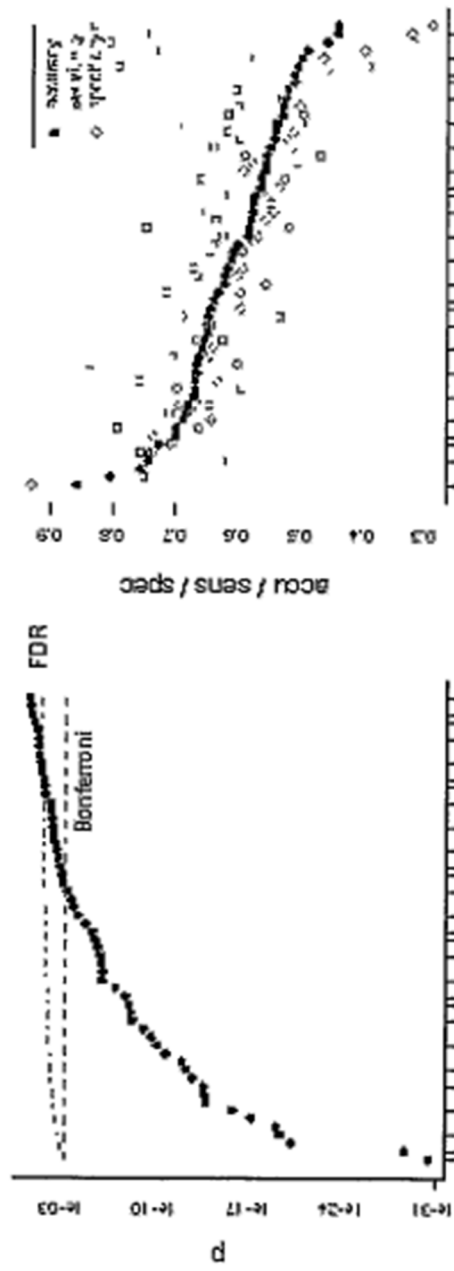


Figura 7

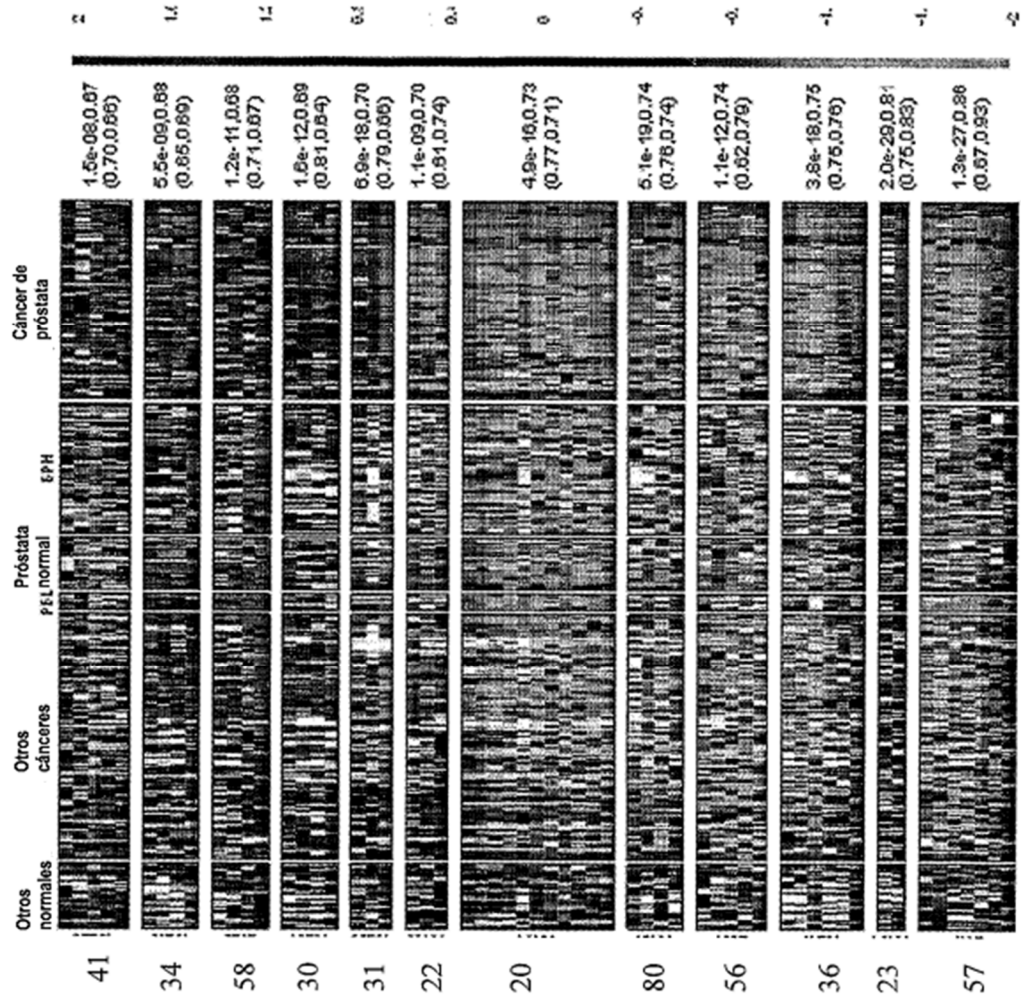


Figura 8

