

19



OFICINA ESPAÑOLA DE  
PATENTES Y MARCAS

ESPAÑA



11 Número de publicación: **2 615 354**

51 Int. Cl.:

**C12Q 1/68**

(2006.01)

12

TRADUCCIÓN DE PATENTE EUROPEA

T3

86 Fecha de presentación y número de la solicitud internacional: **26.11.2007 PCT/EP2007/010257**

87 Fecha y número de publicación internacional: **29.05.2008 WO08061801**

96 Fecha de presentación y número de la solicitud europea: **26.11.2007 E 07846823 (8)**

97 Fecha y número de publicación de la concesión europea: **16.11.2016 EP 2099938**

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

30 Prioridad:

**24.11.2006 EP 06124746**

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

**06.06.2017**

73 Titular/es:

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

72 Inventor/es:

**SLEDZIEWSKI, ANDREW;  
LOFTON-DAY, CATHERINE;  
TETZNER, REIMO;  
DISTLER, JÜRGEN;  
MODEL, FABIAN;  
PAYNE, SHANNON y  
DIETRICH, DIMO**

74 Agente/Representante:

**SÁEZ MAESO, Ana**

**ES 2 615 354 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 de genes asociados con el desarrollo de trastornos proliferativos de células de próstata

Campo de la invención

- 5 La presente invención se refiere a un método para detectar un trastorno proliferativo de células de próstata en un sujeto y al uso de oligonucleótidos, ácidos nucleicos o ácidos nucleicos de péptidos en el diagnóstico de un trastorno proliferativo de células de próstata.

Técnica anterior

- 10 El cáncer de próstata es el cáncer más común y la tercera causa principal de muerte en hombres estadounidenses (Jemal et al., 2006). Las tasas de incidencia y mortalidad de esta enfermedad aumentan considerablemente con la edad, con más del 65% de todos los casos de cáncer de próstata diagnosticados en hombres mayores de 65 años (Jemal et al., 2006). La etapa de la enfermedad en el diagnóstico también afecta a las tasas de supervivencia global. Debido al uso generalizado de la prueba de detección del antígeno prostático específico (PSA), casi el 90% de los nuevos pacientes son diagnosticados con enfermedad local o regional (Jemal et al., 2005). Los pacientes con enfermedad local o regional cuando se diagnostican tienen una tasa de supervivencia relativa de cinco años que se aproxima al 100% (Jemal et al., 2006).

- 15 Las directrices actuales para el cribado del cáncer de próstata, de acuerdo con la Sociedad Americana del Cáncer, aconsejan las pruebas de PSA elevado y el examen rectal digital anualmente a partir de los 50 años. Para los hombres con alto riesgo de desarrollar cáncer de próstata (afroamericanos hombres y hombres con uno o más parientes de primer grado diagnosticados a una edad temprana), el examen debe comenzar a los 45 años. Los hallazgos positivos en cualquiera de estos exámenes están conformados por biopsia de próstata. El advenimiento de la detección de PSA ha cambiado el panorama del diagnóstico de cáncer de próstata. Las tasas de incidencia en el cáncer de próstata han aumentado dramáticamente en los últimos 20 años, mientras que el diagnóstico en varones mayores de 65 años se ha estabilizado. Las pruebas de PSA sufren de dos desventajas. La primera es su baja especificidad puesto que el PSA se eleva en una serie de condiciones benignas aparte del cáncer de próstata. Esto se traduce en un gran número de biopsias de próstata que se llevan a cabo en hombres que no tienen cáncer de próstata. La segunda desventaja es que a pesar de la sensibilidad relativamente alta del PSA, hay hombres que albergan cáncer de próstata en ausencia de PSA elevado (>4 ng/ml) (Thompson et al., 2004). También se ha estimado que hasta el 10% de las biopsias de próstata bajo las directrices actuales son falsamente negativas, lo que resulta en una disminución de la sensibilidad incluso con la biopsia (Djavan et al., 2000, Gupta et al., 2005, Hanley Et al., 2006). Claramente se requieren pruebas mejoradas con mayor especificidad y sensibilidad.

- 20 Enfoque multifactorial. El diagnóstico de cáncer se ha basado tradicionalmente en la detección de marcadores moleculares individuales (por ejemplo, mutaciones génicas, niveles elevados de PSA). Desafortunadamente, el cáncer es un estado de enfermedad en el que los marcadores individuales típicamente no han detectado o diferenciado muchas formas de la enfermedad. Por tanto, se ha demostrado que los ensayos que reconocen sólo un único marcador tienen un valor predictivo limitado. Los diagnósticos de cáncer basados en metilación y el cribado, diagnóstico y monitorización terapéutica de tales enfermedades proporcionarán mejoras significativas sobre el estado de la técnica que utiliza análisis de marcador único mediante el uso de una selección de marcadores múltiples. El enfoque analítico multiplexado es particularmente adecuado para el diagnóstico del cáncer, ya que el cáncer no es una enfermedad simple, este enfoque de "panel" multifactorial es coherente con la naturaleza heterogénea del cáncer, tanto citológica como clínicamente. La clave para la implementación exitosa de un enfoque de panel a las pruebas de diagnóstico basadas en metilación es el diseño y desarrollo de paneles optimizados de marcadores que pueden caracterizar y distinguir estados de enfermedad. En el presente documento se describe una pluralidad de paneles de genes particularmente eficaces y únicos, el análisis de metilación de uno o una combinación de los miembros del panel que permite la detección de trastornos proliferativos de células prostáticas con una sensibilidad, especificidad y/o valor predictivo particularmente elevados.

- 25 Desarrollo de pruebas médicas. Dos medidas claves de evaluación de cualquier examen médico o prueba de diagnóstico son su sensibilidad y especificidad, que miden qué tan bien se desempeña la prueba para detectar con precisión todas las personas afectadas sin excepción, y sin incluir falsamente individuos que no tienen la enfermedad objetivo (valor predictivo). Históricamente, muchas pruebas diagnósticas han sido criticadas debido a su escasa sensibilidad y especificidad.

- 30 Un resultado positivo verdadero (TP) es aquel donde la prueba es positiva y la condición está presente. Un resultado positivo falso (FP) es donde la prueba es positiva pero la condición no está presente. Un resultado negativo verdadero (TN) es donde la prueba es negativa y la condición no está presente. Un resultado negativo falso (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 =  $TP/(TP+FP)$ .

5 La sensibilidad es una medida de la capacidad de una prueba para detectar correctamente la enfermedad objetivo en un individuo que está siendo probado. Una prueba que tiene una sensibilidad pobre produce una alta tasa de falsos negativos, es decir, individuos que tienen la enfermedad, pero que están falsamente identificados como libres de esa enfermedad en particular. El peligro potencial de un falso negativo es que el individuo enfermo permanecerá sin diagnosticar y sin tratamiento durante algún 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 eficaces. Un ejemplo de una prueba que tiene baja sensibilidad es un análisis de sangre basado en proteínas para el VIH. Este tipo de prueba presenta una sensibilidad pobre porque no detecta la presencia del virus hasta que la enfermedad está bien establecida y el virus ha invadido el torrente sanguíneo en cantidades sustanciales. Por el contrario, un ejemplo de una prueba que tiene alta sensibilidad es la detección de carga viral utilizando la reacción en cadena de la polimerasa (PCR). Se consigue 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 omitir un diagnóstico son altas.

15 La especificidad, por otra parte, es una medida de la capacidad de una prueba para identificar con precisión pacientes que están libres del estado de la enfermedad. Un ensayo que tiene una especificidad pobre produce una alta tasa de falsos positivos, es decir, individuos que son falsamente identificados como portadores de la enfermedad. Un inconveniente de los falsos positivos es que obligan a los pacientes a someterse a tratamientos de procedimientos médicos innecesarios con sus riesgos asociados, estrés emocional y financiero, y que podría tener efectos adversos sobre la salud del paciente. Una característica de las enfermedades que hace difícil desarrollar pruebas diagnósticas con alta especificidad es que los mecanismos de la enfermedad, particularmente en el cáncer, a menudo implican una pluralidad de genes y proteínas. Además, ciertas proteínas pueden ser elevadas por razones no relacionadas con un estado de la enfermedad, siendo un ejemplo de una prueba que tiene alta especificidad una prueba genética que puede detectar una mutación p53. La especificidad es importante cuando el coste o el riesgo asociado con otros procedimientos de diagnóstico o intervención médica adicional son muy altos.

25 Breve descripción de los dibujos

30 La figura 1 proporciona una visión general de la media logarítmica de metilación medida a través de el ensayo de HM según el Ejemplo 1. Para cada gen analizado (marcado a la izquierda de las figuras), se proporcionan dos gráficas, las gráficas del lado izquierdo proporcionan el análisis multiclasa, la sensibilidad se muestra en el eje Y, la metilación del ADN medida en ( $\log_{10}$  ng/mL) se muestra en el eje X. El diagrama de la derecha proporciona una ROC en la que la sensibilidad se muestra en el eje Y la especificidad 1 se muestra en el eje X.

La figura 2 proporciona una visión general de la media logarítmica de la metilación medida a través de ensayos de PCR en tiempo real HM de orina de postmasaje prostático de PCa y clase I negativa (individuos sanos) según el Ejemplo 1. Para cada gen analizado, (como se ha marcado encima de cada gráfica) la sensibilidad se muestra en el eje Y, la metilación del ADN medida en ( $\log_{10}$  ng/mL) se muestra en el eje X.

35 La Figura 3 proporciona una visión general de la media logarítmica de la metilación medida a través de ensayos de PCR en tiempo real de HM de orina postmasaje prostático de PCa y clase II negativa (individuos sanos más biopsias negativas) según el Ejemplo 1. Para cada gen analizado (marcado encima de cada gráfica), la sensibilidad se muestra en el eje Y, la metilación del ADN medida en ( $\log_{10}$  ng/mL) se muestra en el eje X.

40 La Figura 4 muestra la media logarítmica de la metilación medida a través de ensayos de PCR en tiempo real HM de plasma de PCa y clase I negativa (individuos sanos). Para cada gen analizado (tal como se ha marcado encima de cada gráfica), la sensibilidad se muestra en el eje Y, la metilación del ADN medida en ( $\log_{10}$  ng/mL) se muestra en el eje X.

45 La Figura 5 muestra la media logarítmica de la metilación medida a través de ensayos de PCR en tiempo real HM de plasma de PCa y clase II negativa (individuos sanos más biopsias negativas). Para cada gen analizado (tal como se marca encima de cada gráfica), la sensibilidad se muestra en el eje Y, la metilación del ADN medida en ( $\log_{10}$  ng/mL) se muestra en el eje X.

Resumen de la invención

El alcance de la presente invención se define en las reivindicaciones adjuntas.

50 La presente invención proporciona un método para detectar trastornos proliferativos de células de próstata, más preferiblemente carcinoma de próstata, en un sujeto, que comprende determinar los niveles de metilación del gen RASSF2A en una muestra biológica aislada de dicho sujeto, en donde dicha metilación se determina detectando la presencia o ausencia de metilación de CpG dentro de dicho gen, y donde la presencia de metilación de CpG es indicativa de la presencia de dicho trastorno. Preferiblemente se analizan una pluralidad de genes (en el presente

documento también denominados "panel de genes"). Preferiblemente se analizan 2, 3 o 4 genes. En una realización del método, dicho panel comprende RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E; HIST1H4K y GSTPi. Preferiblemente dicho grupo comprende el gen HIST1H4K.

5 Preferiblemente, dicho método comprende poner en contacto ADN genómico aislado de una muestra biológica (preferiblemente seleccionada del grupo que consiste en eyaculación, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre) obtenidas del sujeto con al menos un reactivo o serie de reactivos que distinguen entre dinucleótidos CpG metilados y no metilados dentro de al menos una y más preferiblemente una pluralidad de regiones objetivo del ADN genómico, en donde la secuencia nucleotídica de dicha región objetivo comprende al menos una secuencia dinucleotídica CpG.

10 Preferiblemente, la región objetivo comprende, o hibrida bajo condiciones restrictivas a una secuencia de al menos 16, 50, 100 o 500 nucleótidos contiguos de al menos una secuencia seleccionada del grupo que consiste en la SEQ ID NO: 1 a la SEQ ID NO: 4.

15 Preferiblemente, la fuente de la muestra de ensayo se selecciona del grupo que consiste en células o líneas celulares, diapositivas histológicas, biopsias, tejido embebido en parafina, fluidos corporales, eyaculación, orina, sangre y combinaciones de los mismos. Más preferiblemente, la fuente se selecciona del grupo que consiste en eyaculación, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre obtenida del sujeto.

20 También se describen nuevas secuencias de ácido nucleico genómicas y químicamente modificadas, así como oligonucleótidos y/o oligómeros de PNA para el análisis de patrones de metilación de citosina dentro de secuencias del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4.

Descripción detallada de la invención

Definiciones:

25 El término "Relación Observada/Esperada" ("relación O/E") se refiere a la frecuencia de dinucleótidos CpG dentro de una secuencia de ADN particular y corresponde al  $[\text{número de sitios CpG}/(\text{número de bases C} \times \text{número de bases G})]/\text{longitud de banda para cada fragmento}$ .

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 GC"  $>0.5$ . Las islas CpG tienen típicamente, pero no siempre, entre aproximadamente 0.2 a aproximadamente 1 KB, o aproximadamente 2 kb de longitud.

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

35 El término "hemi-metilación" o "hemimetilación" se refiere al estado de metilación de un ADN de doble cadena en el que sólo una cadena del mismo está metilada.

40 El término "AUC", tal como se utiliza aquí, es una abreviatura para el área bajo una curva. En particular se refiere al área bajo una curva de Característica de Operación del Receptor (ROC). La curva ROC es un gráfico de la tasa positiva real frente a la tasa de falsos positivos para los diferentes puntos de corte posibles de una prueba de diagnóstico. Muestra el equilibrio entre sensibilidad y especificidad dependiendo del punto de corte seleccionado (cualquier aumento de sensibilidad se acompañará de una disminución en la especificidad). El área bajo una curva ROC (AUC) es una medida para la exactitud de una prueba diagnóstica (cuanto mayor sea el área mejor, siendo 1 óptimo, una prueba aleatoria tendría una curva ROC situada en la diagonal con un área de 0.5; para referencia: JP Egan. Signal Detection Theory and ROC Analysis, Academic Press, New York, 1975).

45 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 encontrado en los dinucleótidos CpG correspondientes dentro de una muestra de ADN de control normal.

50 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 encontrado en los dinucleótidos CpG correspondientes dentro de una muestra de ADN de control normal.

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

5 Los "parámetros genéticos" son mutaciones y polimorfismos de genes y secuencias adicionales requeridos para su regulación. Para ser designados como mutaciones se encuentran, en particular, inserciones, eliminaciones, mutaciones puntuales, inversiones y polimorfismos y, particularmente preferidos, SNPs (polimorfismos de un solo nucleótido).

10 "Parámetros epigenéticos" son, en particular, la metilación de citosina. Otros parámetros epigenéticos incluyen, por ejemplo, la acetilación de histonas que, sin embargo, no pueden analizarse directamente utilizando 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 en el presente documento para distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas.

15 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 con cebado 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 enfocarse en las regiones que más probablemente contienen dinucleótidos de CpG y descrito por Gonzalzo et al., Cancer Research 57: 594-599, 1997.

20 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.

25 El término ensayo de "HeavyMethyl™", en la realización de la misma implementada en el presente documento, se refiere a un ensayo, en el que las sondas de bloqueo específicas de metilación (también denominadas aquí bloqueantes) que cubren posiciones CpG entre, o cubiertas por los cebadores de amplificación permiten la amplificación selectiva específica por metilación de una muestra de ácido nucleico.

El ensayo "HeavyMethyl™ MethyLight™", en su realización implementada en el presente documento, se refiere a un ensayo HeavyMethyl™ MethyLight™, que es una variación del ensayo MethyLight™, en el que el ensayo MethyLight™ se combina con sondas de bloqueo específicas de metilación que cubren posiciones CpG entre los cebadores de amplificación.

30 El término "Ms-SNuPE" (Extensión de cebador nucleótido individual sensible a la metilación) se refiere al ensayo reconocido en la técnica descrito por Gonzalzo & 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 Estados Unidos No. 5,786,146.

35 El término "COBRA" (Análisis de Restricción por Bisulfito Combinado) 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.

40 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.

45 Las "condiciones de hibridación restrictivas", como se definen aquí, implican la hibridación a 68°C en 5x SSC/solución de Denhardt 5x/ SDS al 1.0%, y lavado en SSC 0.2x/SDS al 0.1% a temperatura ambiente, o implican el equivalente reconocido en el arte (por ejemplo, condiciones en las que se lleva a cabo una hibridación a 60°C en 2.5 x regulador SSC, seguido de varias etapas de lavado a 37°C en una concentración de regulador baja y permanece estable). Las condiciones moderadamente restrictivas, como se definen aquí, implican incluir el lavado en 3x SSC a 42°C, o su equivalente reconocido en la técnica. Los parámetros de concentración de sal y temperatura se pueden variar para alcanzar el nivel óptimo de identidad entre la sonda y el ácido nucleico objetivo. La orientación con respecto a tales condiciones está disponible en la técnica, por ejemplo, en Sambrook et al., 1989, Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, N.Y.; y Ausubel et al. (Eds.), 1995, Current Protocols in Molecular Biology, (John Wiley & Sons, N.Y.) en la Unidad 2.10.

50

5 Los términos "enzimas de restricción específicas a la metilación" o "enzimas de restricción sensibles a la metilación" se entenderán como una enzima que digiere selectivamente un ácido nucleico dependiendo del estado de metilación de su sitio de reconocimiento. En el caso de tales enzimas de restricción que cortan específicamente si el sitio de reconocimiento no está metilado o hemimetilado, el corte no tendrá lugar, o lo hará con una eficiencia significativamente reducida, si el sitio de reconocimiento está metilado. En el caso de tales enzimas de restricción que cortan específicamente si el sitio de reconocimiento está metilado, el corte no tendrá lugar, o lo hará con una eficiencia significativamente reducida si el sitio de reconocimiento no está metilado. Se prefieren enzimas de restricción específicas de metilación, cuya secuencia de reconocimiento contiene un dinucleótido CG (por ejemplo cgcg o cccggg). Se prefieren además para algunas realizaciones enzimas de restricción que no cortan si la citosina en este dinucleótido está metilada en el átomo de carbono C5.

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

15 Haciendo referencia a secuencias de matrices mixtas, la frase "nucleótidos contiguos" se refiere a una región de secuencia contigua de cualquier secuencia contigua individual de la matriz compuesta, pero no incluye una región de la secuencia de matriz compuesta que incluye un "nodo" como se ha definido anteriormente en el presente documento.

20 Se entenderá que los términos "RASSF2A; TFAP2E; HIST1H4K y GSTPi" incluyen todas sus variantes de transcripción y todos sus elementos promotores y reguladores. Además, puesto que se conoce una pluralidad de SNPs dentro de dicho gen, se entenderá que el término incluye todas sus variantes de secuencia.

Visión de conjunto:

El alcance de la presente invención se define en las reivindicaciones adjuntas.

25 La presente invención proporciona un método para detectar carcinoma de próstata en un sujeto que comprende determinar los niveles de metilación del gen RASSF2A en una muestra biológica aislada de dicho sujeto, en la que la metilación de CpG es indicativa de la presencia o clase de dicho trastorno. Dicho marcador puede usarse para el diagnóstico de cáncer de próstata incluyendo detección temprana durante las etapas precancerosas de la enfermedad. El marcador de la presente invención es particularmente eficaz en la detección de trastornos proliferativos de células prostáticas malignas tales como carcinoma de próstata, proporcionando así medios mejorados para la detección, clasificación y tratamiento tempranos de dichos trastornos.

30 Además de las realizaciones anteriores, en las que se analiza el análisis de metilación del gen RASSF2A, la invención presenta otros paneles de genes que comprenden RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E; HIST1H4K y GSTPi con una novedosa utilidad para la detección de trastornos proliferativos celulares, en particular el cáncer de próstata. Preferiblemente se analizan una pluralidad de genes (en el presente documento también denominados "panel de genes"). Preferiblemente se analizan 2, 3 o 4 genes. En una realización del método, dicho panel comprende RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E; HIST1H4K y GSTPi. Preferiblemente, dicho grupo comprende el gen HIST1H4K.

40 La modificación de ADN con bisulfito es una herramienta reconocida en la técnica utilizada para evaluar el estatus de metilación 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 un componente de la información genética es de gran interés. Sin embargo, las posiciones de 5-metilcitosina no pueden ser identificadas por secuenciación, porque la 5-metilcitosina tiene el mismo comportamiento de apareamiento de bases que la citosina. Además, la información epigenética transportada por la 5-metilcitosina se pierde completamente durante, por ejemplo, la amplificación por PCR.

45 El método más utilizado para analizar el ADN para la presencia de 5-metilcitosina se basa en la reacción específica del bisulfito con citosina, por lo que, tras la hidrólisis alcalina subsiguiente, la citosina se convierte en uracilo que corresponde a timina en su comportamiento de emparejamiento de bases. Significativamente, sin embargo, 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 se podía distinguir de la citosina por su comportamiento de hibridación, puede ahora ser detectada como la única citosina restante usando técnicas de biología molecular estándar, reconocida por 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 pueden ser explotadas completamente. La técnica anterior, en términos de sensibilidad, se define por un método que comprende encerrar el ADN que va a ser analizado en una matriz de agarosa, impidiendo de este modo la difusión y renaturalización del ADN (el bisulfito sólo reacciona con el ADN de cadena sencilla) y reemplazando todas las

etapas de precipitación y de purificación con diálisis rápida (Olek A, et al., A modified and improved method for bisulfite based cytosine methylation analysis, *Nucleic Acids Res.* 24:5064-6, 1996). Por lo tanto, es posible analizar células individuales en cuanto al estado de metilación, ilustrando la utilidad y la sensibilidad del método. Un resumen de los métodos reconocidos en la técnica para detectar la 5-metilcitosina es proporcionado por Rein, T., et. al., *Nucleic Acids Res.*, 26: 2255, 1998.

La técnica de bisulfito, salvo pocas excepciones (por ejemplo, Zeschnigk M, et al., *Eur J Hum Genet.*, 5: 94-98, 1997), se utiliza actualmente sólo en la 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 y Walter, *Nat Genet.* 1997 17: 275-6, 1997), se someten a una o más reacciones de extensión del cebador (Gonzalzo y Jones, *Nucleic Acids Res.*, 25: 2529-31, 1997; WO 95/00669; Patente de Estados Unidos No. 6,251,594) para analizar posiciones de citosina individuales 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 y 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 el presente documento se describe el uso de la técnica de bisulfito, en combinación con uno o más ensayos de metilación, para la determinación del estado de metilación de las secuencias de dinucleótidos CpG dentro de SEQ ID NO: 1 a SEQ ID NO: 4. Los dinucleótidos CpG genómicos pueden ser metilados o no metilados (alternativamente conocidos como sobre y submetilados respectivamente). Sin embargo, los métodos descritos en el presente documento son adecuados para el análisis de muestras biológicas de naturaleza heterogénea, por ejemplo a baja concentración de células tumorales dentro de un fondo de sangre o eyaculado. Por consiguiente, al analizar el estado de metilación de una posición CpG dentro de dicha muestra, el experto en la técnica puede utilizar un ensayo cuantitativo para determinar el nivel (por ejemplo, porcentaje, fracción, relación, proporción o grado) de metilación en una posición CpG particular como opuesto a un estado de metilación. Por consiguiente, también debe considerarse que el término estatus de metilación o estado de metilación significa un valor que refleja el grado de metilación en una posición CpG. A menos que se indique específicamente los términos "hipermetilado" o "sobremetilado" se tomará para significar un nivel de metilación por encima de un punto de corte especificado, en el que dicho corte puede ser un valor que representa el nivel de metilación medio o mediano para una población dada, o es preferiblemente un nivel de corte optimizado. El "corte" también se refiere aquí como un "umbral". En el contexto de la presente invención, los términos "metilado", "hipermetilado" o "sobremetilado" se tomarán para incluir un nivel de metilación por encima del punto de corte de cero (0)% (o equivalentes) de metilación para todas las posiciones CpG dentro de y asociadas con (por ejemplo, en regiones promotoras o reguladoras) el gen RASSF2A.

Preferiblemente se analiza una pluralidad de genes (en el presente documento también denominados "panel de genes"). Preferiblemente se analizan 2, 3 o 4 genes. En una realización del método, dicho panel comprende RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E; HIST1H4K y GSTPi. Preferiblemente dicho grupo comprende el gen HIST1H4K.

La determinación del estado de metilación de secuencias de dinucleótidos de CpG dentro de SEQ ID NO: 1 a SEQ ID NO: 4 tiene utilidad en el diagnóstico de cáncer de próstata.

Procedimientos de ensayos de metilación. En la técnica se conocen diversos procedimientos de ensayo de metilación, y se pueden usar conjuntamente con la presente invención. Estos ensayos permiten la determinación del estado de metilación de uno o una pluralidad de dinucleótidos CpG (por ejemplo, islas CpG) dentro de una secuencia de ADN. Dichos ensayos implican, entre otras técnicas, la secuenciación de ADN de ADN tratado con bisulfito, PCR (para amplificación específica de secuencias), 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 la 5-metilcitosina mediante el uso de tratamiento con bisulfito (Frommer et al., *Proc. Nat. 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 en bisulfito, por ejemplo, el método descrito por Sadri & Hornsby (*Nucl. Acids Res.* 24: 5058-5059, 1996) o COBRA (Combined Bisulfite Restriction Analysis) Xiong & Laird, *Nucleic Acids Res.* 25: 2532-2534, 1997).

COBRA. El análisis COBRA<sup>TM</sup> es un ensayo de metilación cuantitativo útil para determinar los niveles de metilación del ADN en locus 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 metilación se introducen primero en el ADN genómico mediante tratamiento estándar con bisulfito de acuerdo con el procedimiento descrito por Frommer et al. (*Proc. Natl. Acad. Sci. USA* 89: 1827-1831, 1992). La amplificación por PCR del ADN convertido en bisulfito se realiza a continuación utilizando cebadores específicos para las islas CpG de interés, seguido por digestión con endonucleasas de restricción, electroforesis en

- 5 gel y detección usando sondas de hibridación específicas etiquetadas. 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 una manera linealmente cuantitativa a través de un amplio espectro de niveles de metilación de ADN. Además, esta técnica puede aplicarse confiablemente al ADN obtenido a partir de muestras de tejido embebidas en parafina microdisseccionadas.
- 10 Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico de COBRA™) para el análisis COBRA™ pueden incluir, pero no se limitan a: cebadores de PCR para gen específico (o secuencia de ADN tratada con bisulfito o isla CpG); enzima de restricción y regulador apropiado; oligonucleótido de hibridación génica; oligonucleótido de hibridación de control; kit de marcación de quinasa para sonda de oligonucleótido; y nucleótidos marcados. Adicionalmente, los reactivos de conversión de 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.
- 15 Preferiblemente, los ensayos tales como "MethyLight™" (una técnica de PCR en tiempo real basada en fluorescencia) (Eads et al., Cancer Res. 59: 2302-2306, 1999), reacciones Ms-SNuPE™ (Extensión de cebador de oligonucleótido sencillo sensible a la metilación) (Gonzalzo y Jones, Nucleic Acids Res. 25: 2529-2531, 1997), PCR de metilación específica ("MSP"; Herman et al., Proc. Natl. Acad. Sci. USA 93: 9821-9826, 1996; Patente de Estados Unidos No. 5,786,146), y la amplificación de isla CpG metilada ("MCA", Toyota et al., Cancer Res. 59: 2307-12, 1999) se usan solos o en combinación con otros de estos métodos.
- 20 La técnica de ensayo de "HeavyMethyl™" es un método cuantitativo para evaluar diferencias de metilación basadas en la amplificación específica de metilación de ADN tratado con bisulfito. Las sondas de bloqueo específicas de metilación (también denominadas aquí bloqueantes) que cubren posiciones CpG entre, o cubiertas por los cebadores de amplificación, permiten la amplificación selectiva específica de metilación de una muestra de ácido nucleico.
- 25 El término "HeavyMethyl™ MethyLight™", en su realización implementada aquí, se refiere a un ensayo HeavyMethyl™ MethyLight™, que es una variación del ensayo MethyLight™, en el que el ensayo MethyLight™ se combina con sondas de bloqueo específicas de metilación que cubren posiciones CpG entre los cebadores de amplificación. El ensayo de HeavyMethyl™ también puede usarse en combinación con cebadores de amplificación específicos de metilación.
- 30 Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico de MethyLight™) para análisis de HeavyMethyl™ pueden incluir, pero no se limitan a: cebadores de PCR para genes específicos (o secuencia de ADN tratada con bisulfito o isla CpG); oligonucleótidos de bloqueo; reguladores y desoxinucleótidos de PCR optimizados; y Taq polimerasa.
- 35 MSP. La MSP (PCR específica para metilación) permite evaluar el estado de metilación de virtualmente 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. 93: 9821-9826, 1996; Patente de los Estados Unidos No. 5,786,146). En resumen, el ADN se modifica mediante bisulfito de sodio convirtiendo todas las citosinas no metiladas, pero no las metiladas, en uracilo, y posteriormente se amplifica con cebadores específicos para ADN metilado frente a ADN no metilado. La MSP requiere sólo pequeñas cantidades de ADN, es sensible a los alelos metilados al 0.1% de un locus insular CpG dado, y puede realizarse sobre el ADN extraído de muestras embebidas en parafina. Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico basado en MSP) para análisis de MSP pueden incluir, pero no se limitan a: cebadores de PCR metilados y no metilados para gen específico (o secuencia de ADN tratada con bisulfito o isla CpG), reguladores y desoxinucleótidos, y sondas específicas.
- 40
- 45 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 manipulaciones adicionales 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, a un conjunto mixto de diferencias de secuencias dependientes de la metilación según procedimientos estándar (el proceso de bisulfito convierte residuos de citosina no metilados en uracilo). La PCR basada en fluorescencia se realiza entonces en una reacción "polarizada" (con cebadores de PCR que se superponen a dinucleótidos CpG conocidos). La discriminación de secuencia puede ocurrir tanto a nivel del proceso de amplificación como a nivel del proceso de detección de fluorescencia. El ensayo MethyLight™ puede usarse como una prueba cuantitativa para los patrones de metilación en la muestra de ADN genómico, en la que se produce la discriminación de la secuencia al nivel de la hibridación de la sonda. En esta versión cuantitativa, la reacción de PCR proporciona una amplificación específica de metilación en presencia de una sonda fluorescente que se superpone a un sitio de metilación putativo particular. Un control imparcial para la cantidad de ADN de entrada se proporciona mediante una reacción en la que ni los cebadores ni la sonda se encuentran encima de ningún dinucleótido CpG. Alternativamente, se logra una prueba cualitativa para la metilación genómica mediante el sondeo de la agrupación de PCR sesgada con
- 55

oligonucleótidos de control que no "cubren" sitios de metilación conocidos (una versión basada en fluorescencia de las técnicas de HeavyMethyl™ y MSP), o con oligonucleótidos que cubren posibles sitios de metilación.

5 El proceso MethyLight™ se puede usar con cualquier sonda adecuada, por ejemplo "TaqMan®", Lightcycler® etc. Por ejemplo, el ADN genómico de doble cadena 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 MSP y/o oligonucleótidos bloqueadores de HeavyMethyl y sonda TaqMan®. La sonda TaqMan® está etiquetada dos veces con moléculas fluorescentes "informador" y "detenedor", y está diseñada para ser específica para una región de contenido de GC relativamente alta de modo que se funde a una temperatura superior a 10°C en el ciclo de PCR que la de los cebadores de avance o retroceso. Esto permite que la sonda TaqMan® permanezca completamente hibridada durante la etapa de fusión/extensión de PCR. A medida que la polimerasa Taq sintetiza enzimáticamente una nueva cadena durante la PCR, eventualmente alcanzará la sonda TaqMan® fusionada. La actividad de la endonucleasa 5' a 3' de la polimerasa Taq desplazará entonces la sonda TaqMan® digiriéndola para liberar la molécula informadora fluorescente para la detección cuantitativa de su señal ahora no sometida utilizando un sistema de detección fluorescente en tiempo real.

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

20 El ensayo QM™ (metilación cuantitativa) es un ensayo cuantitativo alternativo para los patrones de metilación en muestras de ADN genómico, en el que se produce la discriminación de secuencia al nivel de hibridación de sonda. En esta versión cuantitativa, la reacción de PCR proporciona una amplificación no sesgada en presencia de una sonda fluorescente que se superpone a un sitio de metilación putativo particular. Un control imparcial para la cantidad de ADN de entrada se proporciona mediante 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 de la agrupación de PCR sesgada con oligonucleótidos de control que no "cubren" sitios de metilación conocidos (una versión basada en fluorescencia de las técnicas de HeavyMethyl™ y MSP) o con oligonucleótidos que cubren los sitios de metilación potenciales.

25 El procedimiento QM™ puede utilizarse con cualquier sonda adecuada, por ejemplo "TaqMan®", Lightcycler® etc... en el proceso de amplificación. Por ejemplo, el ADN genómico de doble cadena se trata con bisulfito de sodio y se somete a cebadores no sesgados y la sonda TaqMan®. La sonda TaqMan® está etiquetada dos veces con moléculas fluorescentes "informador" y "detenedor", y está diseñada para ser específica para una región de contenido de GC relativamente alta de modo que se funde a una temperatura superior en 10°C en el ciclo de PCR a la de los cebadores de avance o retroceso. Esto permite que la sonda TaqMan® permanezca completamente hibridada durante la etapa de fusión/extensión de PCR. A medida que la polimerasa Taq sintetiza enzimáticamente una nueva cadena durante la PCR, eventualmente alcanzará la sonda TaqMan® fusionada. La actividad de la endonucleasa 5' a 3' de la polimerasa Taq desplazará entonces la sonda TaqMan® digiriéndola para liberar la molécula informadora fluorescente para la detección cuantitativa de su señal ahora no sometida utilizando un sistema de detección fluorescente en tiempo real.

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

35 Ms-SNuPE. La técnica de Ms-SNuPE™ es un método cuantitativo para evaluar las diferencias de metilación en sitios específicos de CpG basados en el tratamiento con bisulfito de ADN, seguido por la extensión de un 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 mientras se deja la 5-metilcitosina sin cambios. La amplificación de la secuencia objetivo deseada se realiza a continuación usando cebadores de PCR específicos para el ADN convertido con bisulfito y el producto resultante se aísla y se usa como molde para el análisis de metilación en el sitio o lugares CpG de interés. Se pueden analizar pequeñas cantidades de ADN (por ejemplo, secciones de patología microdisseccionadas), y se evita la utilización de enzimas de restricción para determinar el estado de metilación en sitios CpG.

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

recuperación de ADN.

Se determinó que las secuencias genómicas de acuerdo con SEQ ID NO: 1 a SEQ ID NO: 4 y las variantes tratadas de forma no natural de las mismas de acuerdo con SEQ ID NO: 5 a SEQ ID NO: 20 tenían nueva utilidad para la detección precoz, clasificación y/o tratamiento de trastornos proliferativos celulares, en particular carcinoma de próstata.

En una realización, la invención del método comprende las siguientes etapas: i) poner en contacto ADN genómico (preferiblemente aislado de fluidos corporales) obtenido del sujeto con al menos un reactivo, o una serie de reactivos que distinguen entre dinucleótidos CpG metilados y no metilados dentro del gen RASSF2A; y ii) detectar trastornos proliferativos de las células de próstata, lo más preferiblemente, carcinoma de próstata.

Preferiblemente se analizan una pluralidad de genes (en el presente documento también denominados "panel de genes"). Preferiblemente se analizan 2, 3 o 4 genes. En una realización del método, dicho panel comprende RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E; HIST1H4K y GSTPi. Preferiblemente, dicho grupo comprende el gen HIST1H4K.

El ADN genómico puede aislarse por cualquier medio estándar en la técnica, incluyendo el uso de kits comercialmente disponibles. En resumen, cuando el ADN de interés está encapsulado en una membrana celular, la muestra biológica debe ser deshecha y sometida a lisis por medios enzimáticos, químicos o mecánicos. La solución de ADN puede entonces limpiarse de proteínas y otros contaminantes, por ejemplo, mediante digestión con proteinasa K. El ADN genómico se recupera entonces de la solución. Esto puede llevarse a cabo por medio de una diversidad de métodos que incluyen extracción con sal, extracción orgánica o unión del ADN a un soporte en fase sólida. La elección del método se verá afectada por una variedad de factores, incluyendo el tiempo, los gastos y la cantidad necesaria de ADN. Todos los tipos de muestras clínicas que comprenden materia neoplásica o materia preneoplásica son adecuados para su uso en el presente método, prefiriéndose líneas celulares, secciones histológicas, biopsias, tejido embebido en parafina, fluidos corporales, eyaculación, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre y combinaciones de las mismas. Los fluidos corporales son la fuente preferida del ADN; son particularmente preferidos la eyaculación, el plasma sanguíneo, el suero sanguíneo, la sangre entera, las células sanguíneas aisladas y las células aisladas de la sangre.

La muestra de ADN genómico se trata entonces con al menos un reactivo o serie de reactivos que distinguen entre dinucleótidos CpG metilados y no metilados dentro de al menos una y más preferiblemente una pluralidad de regiones objetivo del ADN genómico, en el que cada blanco comprende, o hibrida bajo condiciones restrictivas a una secuencia de al menos 16, 50, 100 o 500 nucleótidos contiguos de al menos una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4, respectivamente, en la que dichos nucleótidos contiguos comprenden al menos una secuencia de dinucleótido CpG.

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

En el que la muestra de ADN genómico se trata de tal manera que las bases de citosina que son no metiladas en la posición 5' se convierten en uracilo, timina u otra base que es diferente a citosina en términos de comportamiento de hibridación. Se prefiere que este tratamiento sea llevado a cabo con bisulfito (sulfito, disulfito de hidrógeno) y posterior hidrólisis alcalina. Tal tratamiento da como resultado la conversión de la SEQ ID NO: 1 a SEQ ID NO: 4 a SEQ ID NO: 5 a SEQ ID NO: 12 (véase la Tabla 1) en la que dichos dinucleótidos CpG están metilados o SEQ ID NO: 13 a SEQ ID NO: 20 en la que dichos dinucleótidos CpG no están metilados.

El ADN tratado se analiza a continuación para determinar el estado de metilación del gen RASSF2A antes del tratamiento.

Preferiblemente se analiza una pluralidad de genes (en el presente documento también denominados "panel de genes"). Preferiblemente se analizan 2, 3 o 4 genes. En una realización del método, dicho panel comprende RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E; HIST1H4K y GSTPi. Preferiblemente dicho grupo comprende el gen HIST1H4K.

Se prefiere particularmente que cada región objetivo comprenda, o hibride bajo condiciones restrictivas, al menos 16, 50, 100 o 500 nucleótidos contiguos de un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi. Se prefiere que la secuencia de dichos genes de acuerdo con SEQ ID NO: 1 a SEQ ID NO: 4 se analice como se proporciona en la Tabla 1 y la lista de secuencias adjunta. El método de análisis puede seleccionarse entre los conocidos en la técnica, incluyendo los enumerados en el presente documento. Son particularmente preferidos MethyLight™, MSP y el uso de oligonucleótidos bloqueantes (HeavyMethyl™) como se

- describe en el presente documento. Se prefiere además que cualquier oligonucleótido usado en dicho análisis (incluyendo cebadores, oligonucleótidos de bloqueo y sondas de detección) sea inversamente complementario, idéntico o hibride en condiciones restrictivas o altamente restrictivas a un segmento largo de al menos 16 pares de bases de las secuencias de bases de una o más de las SEQ ID NO: 5 a SEQ ID NO: 20 y secuencias complementarias a las mismas.
- 5 La metilación aberrante, más específicamente hipermetilación de RASSF2A; TFAP2E; HIST1H4K y GSTPi se asocia con la presencia de trastornos proliferativos de la próstata, en particular, el cáncer de próstata. De acuerdo con ello, en el que una muestra biológica se presenta dentro de cualquier grado de metilación, dicha muestra debe determinarse como un trastorno proliferativo celular, en particular el cáncer.
- 10 Realizaciones particulares de la presente invención proporcionan una aplicación novedosa del análisis de los niveles y/o patrones de metilación dentro de dichas secuencias que permite una detección, caracterización y/o tratamiento preciso del carcinoma de próstata. La detección precoz del cáncer está directamente relacionada con el pronóstico de la enfermedad y el método revelado permite al médico y al paciente tomar decisiones de tratamiento mejores y más informadas.
- 15 Mejoras adicionales
- Se divulgan aquí nuevos usos para las secuencias genómicas SEQ ID NO: 1 a SEQ ID NO: 4. Las realizaciones adicionales proporcionan variantes modificadas de SEQ ID NO: 1 a SEQ ID NO: 4, así como oligonucleótidos y/o PNA-oligómeros Para el análisis de patrones de metilación de citosina dentro de la SEQ ID NO: 1 a la SEQ ID NO: 4.
- 20 La presente divulgación comprende el análisis del estado de metilación de uno o más dinucleótidos CpG dentro de la SEQ ID NO: 1 a la SEQ ID NO: 4 y secuencias complementarias a la misma.
- 25 La presente divulgación proporciona ácidos nucleicos tratados, derivados de la SEQ ID NO: 1 a SEQ ID NO: 4, en la que el tratamiento es adecuado para convertir al menos una base citosina no metilada de la secuencia de ADN genómico en uracilo u otra base que es detectablemente diferente a la citosina en términos de hibridación. Las secuencias genómicas en cuestión pueden comprender una, o más posiciones CpG metiladas consecutivas. Dicho tratamiento comprende preferiblemente el uso de un reactivo seleccionado del grupo que consiste en bisulfito, sulfito, disulfito de hidrógeno y combinaciones de los mismos. En una realización preferida, la descripción proporciona un ácido nucleico modificado que no ocurre naturalmente que comprende una secuencia de al menos
- 30 16 bases de nucleótidos contiguas en longitud de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 5 a SEQ ID NO: 20. En otras realizaciones preferidas dicho ácido nucleico tiene al menos 50, 100, 150, 200, 250 o 500 pares de bases en longitud de un segmento de la secuencia de ácido nucleico descrita en la SEQ ID NO: 5 a la SEQ ID NO: 20. Se prefiere particularmente una molécula de ácido nucleico que no es idéntica o complementaria a la totalidad o una parte de las secuencias SEQ ID NO: 5 a SEQ ID NO: 20 pero no SEQ ID NO: 1 a SEQ ID NO: 4 u otro ADN natural.
- 35 Se prefiere que dicha secuencia comprenda al menos un dinucleótido CpG, TpA o CpA y secuencias complementarias a los mismos. Las secuencias de SEQ ID NO: 5 a SEQ ID NO: 20 proporcionan versiones modificadas que de origen no natural del ácido nucleico de acuerdo con SEQ ID NO: 1 a SEQ ID NO: 4, 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.
- 40 Para cada ADN genómico de cadena en sentido, por ejemplo, SEQ ID NO: 1, se describen cuatro versiones convertidas. Una primera versión en la que "C" se convierte en "T", pero "CpG" permanece como "CpG" (es decir, corresponde al caso en el que, para la secuencia genómica, todos los residuos "C" de las secuencias de dinucleótidos CpG están metilados y, por lo tanto, no se convierten); una segunda versión describe el complemento de la secuencia de ADN genómico descrita (es decir, la cadena antisentido), en la que "C" se convierte en "T", pero
- 45 "CpG" permanece "CpG" (es decir, los residuos de secuencias de dinucleótidos CpG se metilan y, por lo tanto, no se convierten).
- Las secuencias sobremetiladas convertidas de SEQ ID NO: 1 a SEQ ID NO: 4 corresponden a SEQ ID NO: 5 a SEQ ID NO: 12. Se proporciona una tercera versión químicamente convertida de cada secuencia genómica, en la que "C" se convierte a "T" para todos los residuos "C", incluyendo aquellos de secuencias de dinucleótidos "CpG" (es decir, corresponde al caso en el que, para las secuencias genómicas, todos los residuos "C" de las secuencias de dinucleótidos CpG son no metilados); una versión final químicamente convertida de cada secuencia, describe el complemento de la secuencia de ADN genómico descrita (es decir, cadena antisentido), en la que "C" se convierte en "T" para todos los residuos "C", incluyendo aquellos de secuencias de dinucleótidos CpG es decir, corresponde al
- 50 caso en el que, para el complemento (cadena antisentido) de cada secuencia genómica, todos los residuos "C" de secuencias de dinucleótidos CpG son no metilados). Las secuencias submetiladas convertidas de SEQ ID NO: 1 a
- 55

SEQ ID NO: 4 corresponden a SEQ ID NO: 13 a SEQ ID NO: 20. Véase Tabla 1 para más detalles.

Significativamente, hasta ahora, las secuencias de ácido nucleico y las moléculas de acuerdo con SEQ ID NO: 5 a SEQ ID NO: 20 no estaban implicadas o relacionadas con la detección, clasificación o tratamiento del cáncer .

5 En una realización preferida alternativa, la descripción proporciona además oligonucleótidos u oligómeros adecuados para uso en los métodos descritos en el presente documento para detectar el estado de metilación de la citosina dentro del ADN genómico o tratado (químicamente modificado), de acuerdo con SEQ ID NO: 1 a SEQ ID NO: 4 a la SEQ ID NO: 20. Dichos oligonucleótidos u oligómeros ácidos nucleicos proporcionan nuevos medios de diagnóstico. Dicho oligonucleótido u oligómero que comprende una secuencia de ácido nucleico que tiene una longitud de al menos nueve (9) nucleótidos que es idéntica a, híbrida, bajo condiciones moderadamente restrictivas o restrictivas (como se ha definido anteriormente) a una secuencia de ácido nucleico tratada de acuerdo con SEQ ID NO: 5 a SEQ ID NO: 20 y/o secuencias complementarias a la misma, o a una secuencia genómica según SEQ ID NO: 1 a SEQ ID NO: 4 y/o secuencias complementarias a la misma.

15 Por lo tanto, la presente descripción 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 a todas o una parte de las secuencias SEQ ID NO: 1 a SEQ ID NO: 4 a SEQ ID NO: 20 o a sus complementos. Se prefiere particularmente una molécula de ácido nucleico que se hibrida en condiciones de hibridación moderadamente restrictivas y/o restrictivas a todas o una parte de las secuencias SEQ ID NO: 5 a SEQ ID NO: 20 pero no SEQ ID NO: 1 a SEQ ID NO: 4 u otro ADN genómico humano.

20 La porción idéntica o que se hibrida de los ácidos nucleicos hibridantes tiene típicamente al menos 9, 16, 20, 25, 30 ó 35 nucleótidos de longitud. Sin embargo, las moléculas más largas tienen utilidad, y están por lo tanto dentro del alcance de la presente descripción.

25 Preferiblemente, la porción hibridante de los ácidos nucleicos hibridantes tiene al menos 95%, o al menos 98%, o 100% idéntica a la secuencia, o a una parte de la misma de la SEQ ID NO: 1 a la SEQ ID NO: 4 a la SEQ ID NO: 20, o a sus complementos. Los ácidos nucleicos hibridantes 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 pronóstico. Preferiblemente, la hibridación de la sonda oligonucleotídica con una muestra de ácido nucleico se realiza bajo condiciones restrictivas y la sonda es 100% idéntica a la secuencia objetivo. La estabilidad dúplex o híbrida de ácido nucleico se expresa como la temperatura de fusión o  $T_m$ , que es la temperatura a la que se disocia una sonda de un ADN objetivo. Esta temperatura de fusión se utiliza para definir las condiciones de restricción requeridas.

35 Para secuencias objetivo que están relacionadas y son sustancialmente idénticas a la secuencia correspondiente de SEQ ID NO: 1 a SEQ ID NO: 4 (tales como variantes alélicas y SNPs), en lugar de idénticas, es útil establecer primero la temperatura más baja a la que la hibridación homóloga solo se produce con una concentración particular de sal (por ejemplo, SSC o SSPE). Entonces, suponiendo que el desajuste del 1% da como resultado una disminución de 1°C en el  $T_m$ , la temperatura del lavado final en la reacción de hibridación se reduce en consecuencia (por ejemplo, si se buscan secuencias que tienen una identidad >95% con la sonda, la temperatura de lavado se reduce en 5°C). En la práctica, el cambio en  $T_m$  puede estar entre 0.5°C y 1.5°C por desajuste del 1%.

40 Ejemplos de oligonucleótidos de longitud X (en nucleótidos), como se indica mediante posiciones polinucleotídicas con referencia, por ejemplo, a la SEQ ID NO: 2, incluyen aquellos correspondientes a conjuntos (conjuntos en sentido y antisentido) de oligonucleótidos consecutivamente superpuestos de longitud X, en donde los oligonucleótidos dentro de cada conjunto consecutivamente superpuesto (correspondiente a un valor X dado) se definen como el conjunto finito de oligonucleótidos Z desde las posiciones de nucleótidos:

$N$  a  $(n + (X-1))$ ;

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

45 donde Y es igual a la longitud (nucleótidos o pares de bases) de SEQ ID NO: 2 (6096);

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 que se superponen consecutivamente); y

donde el número (Z) de oligómeros consecutivamente superpuestos de longitud X para una SEQ ID NO: 1 dada de longitud Y es igual a  $Y - (X-1)$ . Por ejemplo Z=6096

50 -19 = 6077 para conjuntos de sentido o antisentido de SEQ ID NO: 2, donde X=20.

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

5 Ejemplos de oligonucleótidos 20-meros incluyen el siguiente conjunto de 2.261 oligómeros (y el conjunto antisentido complementario de los mismos), indicado por posiciones polinucleotídicas con referencia a la SEQ ID NO: 1 a la SEQ ID NO: 4:

1-20, 2-21, 3-22, 4-23, 5-24, ..... y 6077-6096.

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

10 Del mismo modo, ejemplos de oligonucleótidos 25-meros incluyen el siguiente conjunto de 2.256 oligómeros (y el conjunto antisentido complementario de los mismos), indicado por posiciones polinucleotídicas con referencia a la SEQ ID NO: 2:

1-25, 2-26, 3-27, 4-28, 5-29, ... y 6072-6096.

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

15 La presente descripción abarca, para cada una de las SEQ ID NO: 1 a SEQ ID NO: 4 a SEQ ID NO: 20 (sentido y antisentido), múltiples conjuntos de oligonucleótidos consecutivamente superpuestos u oligonucleótidos modificados de longitud X, donde, por ejemplo, X=9, 10, 17, 20, 22, 23, 25, 27, 30 o 35 nucleótidos.

20 Los oligonucleótidos u oligómeros de acuerdo con la presente descripción constituyen herramientas eficaces útiles para determinar los parámetros genéticos y epigenéticos de las secuencias genómicas correspondientes a la SEQ ID NO: 1 a la SEQ ID NO: 4. Los conjuntos preferidos de tales oligonucleótidos u oligonucleótidos modificados de longitud X son los conjuntos de oligómeros consecutivamente superpuestos que corresponden a SEQ ID NO: 1 a SEQ ID NO: 4 a SEQ ID NO: 20 (y a los complementos de los mismos). Preferiblemente, dichos oligómeros comprenden al menos un dinucleótido CpG, TpG o CpA. Los oligonucleótidos u oligómeros particularmente preferidos de acuerdo con la presente descripción son aquellos en los que la citosina del dinucleótido CpG (o de las correspondientes secuencias de TpG o CpA dinucleótido convertidas) está dentro del tercio medio del oligonucleótido; es decir, cuando el oligonucleótido tiene, por ejemplo, 13 bases de longitud, el dinucleótido CpG, TpG o CpA se posiciona dentro del quinto a noveno nucleótido desde el extremo 5'.

30 Los oligonucleótidos de la descripción también pueden modificarse ligando químicamente el oligonucleótido a uno o más restos o conjugados para mejorar la actividad, estabilidad o detección del oligonucleótido. Tales restos o conjugados incluyen cromóforos, fluoróforos, lípidos tales como colesterol, ácido cólico, tioéter, cadenas alifáticas, fosfolípidos, poliaminas, polietilenglicol (PEG), restos de palmito y otros como se describe, por ejemplo, en las Patentes de los Estados Unidos números 5,514,758, 5,565,552, 5,567,810, 5,574,142, 5,585,481, 5,587,371, 5,597,696 y 5,958,773. Las sondas 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 añadidos tales como péptidos, y puede incluir agentes de escisión desencadenado por hibridación (Krol et al., BioTechniques 6: 958-976, 1988) o agentes intercalantes (Zon, Pharm. Res. 5: 539-549, 1988). Con este fin, el oligonucleótido puede conjugarse con otra molécula, por ejemplo, un cromóforo, un fluoróforo, un péptido, un agente de reticulación activado por hibridación, un agente de transporte, un agente de escisión provocado por hibridación, etc.

40 El oligonucleótido también puede comprender al menos una unidad estructural de azúcar y/o base modificada reconocida en la técnica, o puede comprender un esqueleto modificado o un enlace internucleósido no natural.

45 Los oligonucleótidos u oligómeros de acuerdo con realizaciones particulares de la presente descripción se usan típicamente en conjuntos que contienen al menos un oligómero para el análisis de cada uno de los dinucleótidos CpG de una secuencia genómica seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4 y secuencias complementarias a los mismos, o al correspondiente dinucleótido CpG, TpG o CpA dentro de una secuencia de los ácidos nucleicos tratados según SEQ ID NO: 5 a SEQ ID NO: 20 y secuencias complementarias a los mismos. Sin embargo, se anticipa que debido a factores económicos o de otro tipo puede ser preferible analizar una selección limitada de los dinucleótidos CpG dentro de dichas secuencias, y en consecuencia se altera el contenido del conjunto de oligonucleótidos.

50 Por lo tanto, en realizaciones particulares, la presente descripción proporciona un conjunto de al menos dos (2) (oligonucleótidos y/u oligómeros de PNA) útiles para detectar el estado de metilación de citosina en ADN genómico tratado (SEQ ID NO: 5 a SEQ ID NO: 20), o en ADN genómico (SEQ ID NO: 1 a SEQ ID NO: 4 y secuencias

complementarias a la misma). Estas sondas permiten el diagnóstico y la detección de trastornos proliferativos de células de próstata, lo más preferiblemente, carcinoma de próstata. El conjunto de oligómeros también se puede usar para detectar polimorfismos de nucleótido único (SNPs) en ADN genómico tratado (SEQ ID NO: 5 a SEQ ID NO: 20), o en ADN genómico (SEQ ID NO: 1 a SEQ ID NO: 4 y secuencias complementarias a los mismos).

- 5 En realizaciones preferidas, al menos uno, y más preferiblemente todos los miembros de un conjunto de oligonucleótidos están unidos a una fase sólida.

En otras realizaciones, la presente descripción proporciona un conjunto de al menos dos (2) oligonucleótidos que se usan como oligonucleótidos "cebadores" para amplificar secuencias de ADN de una de SEQ ID NO: 1 a SEQ ID NO: 4 a SEQ ID NO: 20 y secuencias complementarias a los mismos, o segmentos de los mismos.

- 10 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/o oligómeros de PNA unidos a una fase sólida). Una serie de oligonucleótidos y/o secuencias de oligómeros de PNA diferentes puede caracterizarse, por ejemplo, porque está dispuesta en la fase sólida en forma de red rectangular o hexagonal. La superficie de fase sólida puede estar compuesta de silicio, vidrio, poliestireno, aluminio, acero, hierro, cobre, níquel, plata u oro. También se pueden utilizar nitrocelulosa, así como plásticos como el nylon, que pueden existir en forma de pastillas o también como matrices de resina. Puede obtenerse una visión general de la técnica anterior en la fabricación de matrices de oligómeros a partir de una edición especial de Nature Genetics (Nature Genetics Supplement, volumen 21, enero de 1999, y de la bibliografía citada en la misma). Las sondas marcadas con fluorescencia se usan a menudo para el escaneo de matrices de ADN inmovilizadas. La fijación simple de los colorantes Cy3 y Cy5 al 5'-OH de la sonda específica son particularmente adecuados para los 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 colorantes Cy3 y Cy5, además de muchos otros, están disponibles comercialmente.

- 15 También se anticipa que los oligonucleótidos, o sus secuencias particulares, pueden constituir todo o parte de un "conjunto virtual" en el que los oligonucleótidos o sus secuencias particulares se usan, por ejemplo, como "especificadores" como parte 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 el documento US 2003/0013091 (Estados Unidos número de serie 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 estar unido de manera única por una etiqueta única y así detectarse (cada etiqueta se cuenta directamente, dando como resultado una lectura digital de cada especie molecular en la mezcla).

- 25 Se prefiere particularmente que los oligómeros de acuerdo con la descripción se utilicen para detectar, o para diagnosticar trastornos proliferativos de células de próstata, lo más preferiblemente, carcinoma de próstata.

- 35 En la realización más preferida del método, se determina la presencia o ausencia de trastornos proliferativos de células de próstata, más preferiblemente cáncer de próstata. Esto se logra mediante el análisis del estado de metilación de al menos una y más preferiblemente una pluralidad de secuencia(s) objetivo que comprende al menos una posición CpG, comprendiendo dicha secuencia o hibridando en condiciones restrictivas a al menos 16, 50, 100 o 500 nucleótidos contiguos de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4 y complementos de los mismos. Preferiblemente se analizan una pluralidad de regiones objetivo (en el presente documento también denominada "panel de genes"). Preferiblemente se analizan regiones objetivo de 2, 3 o 4 genes.
- 40 En una realización del método dicho panel comprende las regiones objetivo de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi y/o su promotor o regiones reguladoras. Preferiblemente dicho panel comprende una región objetivo del gen HIST1H4K.

Son particularmente preferidas las siguientes combinaciones de regiones objetivo de genes:

- HIST1H4K +RASSF
- 45 - HIST1H4K +GSTPi

- 50 La presente divulgación proporciona además un método para determinar parámetros genéticos y/o epigenéticos de las secuencias genómicas según SEQ ID NO: 1 a SEQ ID NO: 4 dentro de un sujeto analizando la metilación de citosina y polimorfismos de nucleótido individual. Dicho procedimiento comprende poner en contacto un ácido nucleico que comprende al menos una secuencia genómica seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4 en una muestra biológica obtenida a partir de dicho sujeto con al menos un reactivo o una serie de reactivos, en el que dicho reactivo o series de reactivos, distingue entre dinucleótidos CpG metilados y no metilados dentro del (de los) ácido(s) nucleico(s) objetivo.

En una realización preferida, dicho método comprende las siguientes etapas: En la primera etapa, se obtiene una

muestra del tejido que se va a analizar. La fuente puede ser cualquier fuente adecuada, tales como líneas celulares, diapositivas histológicas, biopsias, tejido parafínico, fluidos corporales, eyaculación, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre y todas las combinaciones posibles de los mismos. Se prefiere que dichas fuentes de ADN sean eyaculaciones o fluidos corporales seleccionados del grupo que consiste en eyaculación, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre.

A continuación, el ADN genómico se aísla de la muestra. El ADN genómico puede aislarse por cualquier medio estándar en la técnica, incluyendo el uso de kits disponibles comercialmente. En resumen, en el que 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 puede entonces eliminarse de proteínas y otros contaminantes, por ejemplo por digestión con proteinasa K. A continuación se recupera el ADN genómico de la solución. Esto puede llevarse a cabo por medio de una diversidad de métodos que incluyen extracción con sal, extracción orgánica o unión del ADN a un soporte en fase sólida. La elección del método se verá afectada por una variedad de factores, incluyendo el tiempo, los gastos y la cantidad necesaria de ADN.

Cuando el ADN de muestra no está encerrado en una membrana (por ejemplo, ADN circulante de una muestra de sangre) se pueden emplear métodos estándar en la técnica para el aislamiento y/o purificación de ADN. Tales métodos incluyen el uso de un reactivo de degeneración de proteínas, por ejemplo una sal caotrópica, por ejemplo clorhidrato de guanidina o urea; o un detergente, por ejemplo dodecil sulfato de sodio (SDS), bromuro de cianógeno. Métodos alternativos incluyen, pero no se limitan a, precipitación con etanol o precipitación con propanol, concentración al vacío entre otros mediante una centrifuga. El experto en la técnica también puede hacer uso de dispositivos tales como dispositivos de filtro, por ejemplo ultrafiltración, superficies o membranas de sílice, partículas magnéticas, partículas de poliestirol, superficies de poliestirol, superficies cargadas positivamente y membranas cargadas positivamente, membranas cargadas, superficies cargadas, membranas de conmutación cargadas, superficies conmutadas cargadas.

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

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

Esto se consigue preferiblemente por medio de un tratamiento con un reactivo de bisulfito. El término reactivo de bisulfito se refiere a un reactivo que comprende bisulfito, disulfito, sulfito de hidrógeno o combinaciones de los mismos, útil como se describe aquí para distinguir entre secuencias de dinucleótidos CpG metiladas y no metiladas. Los métodos de dicho tratamiento son conocidos en la técnica (por ejemplo PCT/EP2004/011715, que se incorpora como referencia en su totalidad). Se prefiere que el tratamiento con bisulfito se realice en presencia de disolventes desnaturizantes tales como, pero sin limitarse a, n-alquilenglicol, particularmente dietilenglicol dimetil éter (DME), o en presencia de derivados de dioxano o dioxano. En una realización preferida, los disolventes desnaturizantes se utilizan en concentraciones entre 1% y 35% (v/v). También se prefiere que la reacción de bisulfito se lleve a cabo en presencia de agentes eliminadores tales como, pero sin limitación, derivados de cromano, por ejemplo ácido 6-hidroxi-2,5,7,8, -tetrametilcromano 2-carboxílico o ácido trihidroxibenzoico y derivados, por ejemplo ácido gálico (véase: PCT/EP2004/011715, que se incorpora como referencia en su totalidad). La conversión de bisulfito se lleva a cabo preferiblemente a una temperatura de reacción entre 30°C y 70°C, con lo que la temperatura se incrementa a más de 85°C durante cortos períodos de tiempo durante la reacción (véase PCT/EP2004/011715 que se incorpora por referencia en su totalidad). El ADN tratado con bisulfito se purifica preferiblemente antes de la cuantificación. Esto puede realizarse por cualquier medio conocido en la técnica, tal como, pero sin limitarse a, ultrafiltración, preferiblemente llevada a cabo mediante columnas Microcon<sup>(TM)</sup> (fabricadas por Millipore<sup>(TM)</sup>). La purificación se lleva a cabo de acuerdo con un protocolo del fabricante modificado (véase PCT/EP2004/011715 que se incorpora como referencia en su totalidad).

En la tercera etapa del método, se amplifican fragmentos del ADN tratado, utilizando conjuntos de oligonucleótidos cebadores de acuerdo con la presente descripción, 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). Preferiblemente, dichos amplificadores tienen una longitud de 100 a 2.000 pares de bases. El conjunto de oligonucleótidos cebadores incluye al menos dos oligonucleótidos cuyas secuencias son cada una inversamente complementarias, idénticas o hibridan bajo condiciones restrictivas o altamente restrictivas para al menos un segmento largo de 16 pares de bases de las secuencias de bases de una de las SEQ ID NO: 5 a SEQ ID NO: 20 y secuencias complementarias a la misma.

En una realización alternativa del método, se puede detectar el estado de metilación de las posiciones CpG preseleccionadas dentro de al menos una secuencia genómica seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4, mediante el uso de un oligonucleótido de cebador específico para metilación. Esta técnica

(MSP) se ha descrito en la Patente de Estados Unidos No. 6,265,171 de Herman. El uso de cebadores específicos para el 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 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 el ADN no metilado que contienen una "T" en la posición de la posición C en el CpG. Preferiblemente, por lo tanto, la secuencia de bases de dichos cebadores debe comprender una secuencia que tiene una longitud de al menos 9 nucleótidos que se hibrida a una secuencia de ácido nucleico tratada de acuerdo con una de SEQ ID NO: 5 a SEQ ID NO: 20 y secuencias complementarias a la misma, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG. Una realización más preferida del método comprende el uso de oligonucleótidos bloqueadores (el ensayo de HeavyMethyl™). El uso de tales oligonucleótidos bloqueadores ha sido descrito por Yu et al., BioTechniques 23: 714-720, 1997. Los oligonucleótidos de 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 manera que la amplificación de un ácido nucleico se suprime cuando está presente la secuencia complementaria a la sonda de bloqueo. Las sondas pueden diseñarse para hibridar con el ácido nucleico tratado con bisulfito en una forma específica 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, a diferencia de un 'CpG' si se desea la supresión de la amplificación de ácidos nucleicos metilados.

Para métodos de PCR que utilizan oligonucleótidos bloqueadores, la interrupción eficaz de la amplificación mediada por polimerasa requiere que los oligonucleótidos bloqueadores no sean alargados por la polimerasa. Preferiblemente, esto se consigue mediante el uso de bloqueadores que son 3'-desoxi oligonucleótidos, u oligonucleótidos derivados en la posición 3' con otro grupo hidroxilo "libre". Por ejemplo, los oligonucleótidos 3'-O-acetilo son representativos de una clase preferida de molécula bloqueante.

Además, debe evitarse la descomposición mediada por la polimerasa de los oligonucleótidos bloqueantes. Preferiblemente, dicha exclusión comprende bien el uso de una polimerasa que carece de actividad exonucleasa 5'-3', o bien el uso de oligonucleótidos bloqueadores modificados que tienen, por ejemplo, puentes tioato en su extremo 5' que hacen que la molécula bloqueante sea resistente a la nucleasa. Las aplicaciones particulares pueden no requerir tales modificaciones en 5' del bloqueador. Por ejemplo, si los sitios de unión de bloqueador y cebador se solapan, impidiendo de este modo la unión del cebador (por ejemplo, con un bloqueador en exceso), la degradación del oligonucleótido bloqueante quedará sustancialmente excluida. Esto se debe a que la polimerasa no extiende 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-.

Una realización particularmente preferida de bloqueador/PCR, para los fines de la presente descripción y como se implementa en la presente, comprende el uso de oligómeros de ácidos nucleicos peptídicos (PNA) como oligonucleótidos de bloqueo. Tales oligómeros bloqueadores de PNA son ideales, porque no se descomponen ni se extienden por la polimerasa.

Preferiblemente, por lo tanto, se requiere que la secuencia de bases de dichos oligonucleótidos bloqueadores comprenda una secuencia que tenga una longitud de al menos 9 nucleótidos que se hibrida con una secuencia de ácido nucleico tratada de acuerdo con una de SEQ ID NO: 5 a SEQ ID NO: 20 y secuencias complementarias a la misma, en la que la secuencia de bases de dichos oligonucleótidos comprende al menos un dinucleótido CpG, TpG o CpA.

Los fragmentos obtenidos por medio de la amplificación pueden llevar un marcador detectable directa o indirectamente. Se prefieren marcadores en forma de marcadores de fluorescencia, radionúclidos o fragmentos de molécula separables 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 amplificadores marcados tengan una única carga neta positiva o negativa, lo que permite una mejor capacidad de detección en el espectrómetro de masas. La detección puede llevarse a cabo y visualizarse por medio de, por ejemplo, espectrometría de masas con desorción/ionización láser asistida por matriz (MALDI) o usando espectrometría de masas con aspersión de electrones (ESI).

La Espectrometría de Masas con Desorción/Ionización por Láser Asistida por Matriz (MALDI-TOF) es un desarrollo muy eficiente para el análisis de biomoléculas (Karas & Hillenkamp, Anal. Chem., 60: 2299-301, 1988). Un analito es incrustado en una matriz de absorción lumínica. La matriz se evapora mediante un pulso láser corto, transportando así la molécula de analito en la fase de vapor de una manera no fragmentada. El analito es ionizado 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 velocidades.

Los iones más pequeños llegan al detector más pronto que los más grandes. La espectrometría MALDI-TOF es 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 ácidos nucleicos es aproximadamente 100 veces menor que la de los péptidos, y disminuye desproporcionadamente con el aumento del tamaño del fragmento. Además, para los ácidos nucleicos que tienen una cadena principal de carga múltiple negativamente, 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 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 la sensibilidad entre los péptidos y los ácidos nucleicos no se ha reducido. Sin embargo, esta diferencia de sensibilidad puede reducirse modificando químicamente el ADN de tal manera que se vuelva más similar a un péptido. Por ejemplo, los ácidos nucleicos fosforotioato, en los que los fosfatos habituales de la cadena principal están sustituidos con tiofosfatos, pueden convertirse en un ADN neutro en carga utilizando una química de alquilación simple (Gut & Beck, Nucleic Acids Res. 23: 1367-73, 1995). El acoplamiento de una etiqueta de carga con este ADN modificado da como resultado un aumento en la sensibilidad de MALDI-TOF al mismo nivel que el encontrado para los péptidos. Otra ventaja del marcado de carga es la mayor estabilidad del análisis contra las impurezas, lo que hace la detección de sustratos no modificados considerablemente más difíciles.

En la cuarta etapa del procedimiento, los amplificadores obtenidos durante la tercera etapa del método se analizan con el fin de determinar el estado de metilación de los dinucleótidos CpG antes del tratamiento.

En las realizaciones en las que los amplificadores se obtuvieron por medio de la amplificación de MSP, la presencia o ausencia de un amplificador 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 tanto de PCR estándar como de PCR específica para metilación pueden analizarse adicionalmente por medio de métodos basados en las bases tales como, pero sin limitarse a, tecnología de matriz y tecnologías basadas en sondas, así como mediante técnicas tales como secuenciación y extensión dirigida a plantilla.

En una realización del método, los amplificadores sintetizados en la etapa tres se hibridan subsiguientemente a una matriz o un conjunto de oligonucleótidos y/o sondas PNA. En este contexto, la hibridación tiene lugar de la siguiente manera: el conjunto de sondas utilizado durante la hibridación está compuesto preferiblemente por al menos 2 oligonucleótidos u oligómeros de PNA; en el procedimiento, los amplificadores sirven como sondas que hibridan con oligonucleótidos unidos previamente a una fase sólida; los fragmentos no hibridados se eliminan subsiguientemente; dichos oligonucleótidos contienen al menos una secuencia de bases que tiene una longitud de al menos 9 nucleótidos que es complementaria inversa 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. La porción de hibridación de los ácidos nucleicos hibridantes tiene típicamente al menos 9, 15, 20, 25, 30 o 35 nucleótidos de longitud. Sin embargo, las moléculas más largas tienen utilidad, y están por lo tanto dentro del alcance de la presente descripción.

En una realización preferida, dicho dinucleótido está presente en el tercio central del oligómero. Por ejemplo, en el que el oligómero comprende un dinucleótido CpG, dicho dinucleótido es preferiblemente el quinto a noveno nucleótido del extremo 5' de un 13-mero. Existe un oligonucleótido para el análisis de cada dinucleótido CpG dentro de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4, y las posiciones equivalentes dentro de SEQ ID NO: 5 a SEQ ID NO: 20.

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

En aún otra realización del método, el estado de metilación genómica de las posiciones CpG puede determinarse por medio de sondas oligonucleotídicas (como se detalla anteriormente) que se hibridan con el ADN tratado con bisulfito simultáneamente con los cebadores de amplificación por PCR (donde dichos cebadores podrían ser de metilación específica o estándar).

Una realización particularmente preferida de este método es el uso de la PCR cuantitativa en tiempo real basada en fluorescencia (Heid et al., Genome Res. 6: 986-994, 1996, véase también la Patente de Estados Unidos No. 6,331,393) que emplea una sonda de oligonucleótidos fluorescente de doble marcación TaqMan™ PCR, usando un sistema de detección de secuencias 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, denominado sonda TaqMan™, que en realizaciones preferidas está diseñado para hibridarse con una secuencia rica en CpG situada entre los cebadores de amplificación directa e inversa. La sonda TaqMan™ comprende además una "unidad estructural informadora" fluorescente y una "unidad estructural detenedora" unida covalentemente a unidades estructurales enlazadoras (por ejemplo, fosforamiditas) unidas a los nucleótidos del oligonucleótido TaqMan™. Para

- 5 el análisis de metilación dentro de los ácidos nucleicos subsiguiente al tratamiento con bisulfito, se requiere que la sonda sea específica a la metilación, como se describe en la Patente de los Estados Unidos No. 6,331,393 (también incorporada por referencia en su totalidad) también conocida como ensayo MethyLight™. Las variaciones en la metodología de detección de TaqMan™ que son también adecuadas para uso con la presente descripción 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 usar con ADN tratado con bisulfito, y además para el análisis de metilación dentro de dinucleótidos CpG.
- 10 En una realización preferida adicional del método, la cuarta etapa del método comprende el uso de extensión de oligonucleótido dirigida a plantilla, tal como MS-SNuPE como se describe en Gonzalzo & Jones, Nucleic Acids Res. 25: 2529-2531, 1997.
- En otra realización adicional del método, la cuarta etapa del método comprende la secuenciación y el análisis subsiguiente de la secuencia del amplificado generado en la tercera etapa del método (Sanger F., et al., Proc Natl Acad Sci USA 74: 5463-5467, 1977).
- Mejor modo
- 15 En la realización más preferida del método, los ácidos nucleicos genómicos se aíslan y se tratan según las tres primeras etapas del método descrito anteriormente, a saber:
- a) obtener, de un sujeto, una muestra biológica que tenga ADN genómico sujeto;
- b) extraer o aislar de otro modo el ADN genómico;
- 20 c) tratar el ADN genómico de b), o un fragmento del mismo, con uno o más reactivos para convertir bases de citosina que están no metiladas en su posición 5 a uracilo u otra base que es detectablemente diferente a citosina en términos de propiedades de hibridación; y en donde
- d) la amplificación posterior al tratamiento en c) se realiza de una manera específica de metilación, concretamente mediante el uso de cebadores específicos de metilación u oligonucleótidos de bloqueo, y además en donde
- 25 e) la detección de los amplificadores se lleva a cabo por medio de una sonda de detección en tiempo real, como se ha descrito anteriormente.
- Preferiblemente, cuando la posterior amplificación de d) se lleva a cabo por medio de cebadores específicos de metilación, como se ha descrito 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 ácido nucleico tratada de acuerdo con uno de SEQ ID NO: 5 a SEQ ID NO: 20 y secuencias complementarias a los mismos, en donde la secuencia de bases de dichos oligómeros comprende al menos un dinucleótido CpG.
- 30 La etapa e) del método, es decir, la detección de los amplificadores específicos indicativos del estado de metilación de una o más posiciones CpG de al menos una secuencia genómica seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4 se lleva a cabo por medio de métodos de detección en tiempo real como se ha descrito anteriormente.
- 35 Realizaciones adicionales de la invención proporcionan un método para el análisis del estado de metilación del ADN genómico de acuerdo con la invención sin necesidad de conversión de bisulfito. Se conocen métodos en la técnica en los que un reactivo de enzima de restricción sensible a la metilación o una serie de reactivos de enzima de restricción que comprenden reactivos de enzimas de restricción sensibles a la metilación que distinguen entre dinucleótidos CpG metilados y no metilados dentro de una región objetivo se utilizan para determinar la metilación, por ejemplo pero no limitándose a DMH.
- 40 En la primera etapa de tales realizaciones adicionales, la muestra de ADN genómico se aísla de fuentes de tejido o celulares. El ADN genómico puede aislarse por cualquier medio estándar en la técnica, incluyendo el uso de kits comercialmente disponibles. En resumen, en el que 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 puede entonces eliminarse de las proteínas y otros contaminantes, por ejemplo, mediante digestión con proteinasa K. El ADN genómico se recupera después de la solución. Esto puede llevarse a cabo por medio de una diversidad de métodos que incluyen saturación con sal, extracción orgánica o unión del ADN a un soporte en fase sólida. La elección del método se verá afectada por varios factores, incluyendo el tiempo, el gasto y la cantidad necesaria de ADN. Todos los tipos de muestras clínicas que comprenden materias neoplásicas o potencialmente neoplásicas son adecuados para su uso en el presente método, preferentemente líneas celulares, diapositivas histológicas, biopsias, tejido embebido en parafina, fluidos corporales, eyaculación, orina, plasma sanguíneo, suero
- 45
- 50

sanguíneo, células sanguíneas aisladas, células aisladas de la sangre y combinaciones de las mismas. Los fluidos corporales son la fuente preferida del ADN; se prefieren particularmente el plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas y células aisladas de la sangre.

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

- 5 En una realización preferida, el ADN se puede escindir antes del tratamiento con enzimas de restricción sensibles a la metilación.

10 Tales métodos son conocidos en la técnica y pueden incluir tanto medios físicos como enzimáticos. Particularmente preferido es el uso de una o una pluralidad de enzimas de restricción que no son sensibles a la metilación, y cuyos sitios de reconocimiento son ricos en AT y no comprenden dinucleótidos CG. El uso de tales enzimas permite la conservación de islas CpG y regiones ricas en CpG en el ADN fragmentado. Las enzimas de restricción no específicas de metilación se seleccionan preferiblemente del grupo que consiste en MseI, BfaI, Csp6I, Tru1I, Tvu1I, Tru9I, Tvu9I, MaeI y XspI. Particularmente preferido es el uso de dos o tres enzimas de este tipo. Particularmente preferido es el uso de una combinación de MseI, BfaI y Csp6I.

15 El ADN fragmentado puede entonces ligarse a oligonucleótidos adaptadores con el fin de facilitar la posterior amplificación enzimática. La ligación de oligonucleótidos a fragmentos de ADN de extremo romo y adherente se conoce en la técnica y se lleva a cabo por medio de la desfosforilación de los extremos (por ejemplo, usando fosfatasa alcalina de ternera o camarón) y posterior ligación usando enzimas ligasa (por ejemplo, T4 ADN ligasa) en la presencia de dATPs. Los oligonucleótidos adaptadores tienen típicamente al menos 18 pares de bases de longitud.

20 En la tercera etapa, el ADN (o sus fragmentos) es digerida entonces 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 es informativa del estado de metilación de un dinucleótido CpG específico de RASSF2A.

25 Preferiblemente se analizan una pluralidad de genes (en el presente documento también denominados "panel de genes"). Preferiblemente se analizan 2, 3 o 4 genes. En una realización del método, dicho panel comprende RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E; HIST1H4K y GSTPI. Preferiblemente dicho grupo comprende el gen HIST1H4K.

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

35 En la cuarta etapa, 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 fluoróforos, radionúclidos y marcadores de masa. Particularmente preferida es la amplificación por medio de una enzima de amplificación y por lo menos dos cebadores que comprenden, en cada caso, una secuencia contigua de al menos 16 nucleótidos de longitud que es complementaria o hibrida en condiciones moderadamente restrictivas o restrictivas a una secuencia seleccionada del grupo que consiste SEQ ID NO: 1 a SEQ ID NO: 4, y sus complementos. Preferiblemente, dicha secuencia contigua tiene una longitud de al menos 16, 20 o 25 nucleótidos. En una realización alternativa, dichos cebadores pueden ser complementarios a cualquier adaptador unido a los fragmentos.

40 En la quinta etapa se detectan los amplificados. La detección puede ser por cualquier medio estándar en la técnica, por ejemplo, pero sin limitarse a, análisis de electroforesis en gel, análisis de hibridación, incorporación de marcas detectables dentro de los productos de PCR, análisis de matriz de ADN, análisis MALDI o ESI. Preferiblemente dicha detección se lleva a cabo mediante hibridación con al menos un ácido nucleico o ácido nucleico peptídico que comprende en cada caso una secuencia contigua de al menos 16 nucleótidos de longitud que es complementaria o hibrida en condiciones moderadamente restrictivas o restrictivas a una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4, y complementos de los mismos. Preferiblemente, dicha secuencia contigua tiene una longitud de al menos 16, 20 o 25 nucleótidos.

50 Después de la determinación del estado o nivel de metilación de los ácidos nucleicos genómicos, se deduce la presencia, ausencia de trastornos proliferativos de células de próstata, más preferiblemente, se deduce carcinoma de próstata basándose en el estado o nivel de metilación de al menos una secuencia de dinucleótidos CpG de SEQ ID NO: 1 a SEQ ID NO: 4, o un valor medio que refleja un estado de metilación promedio de una pluralidad de secuencias de dinucleótidos CpG de SEQ ID NO: 1 a SEQ ID NO: 4 en las que la metilación está asociada con la presencia de células prostáticas proliferativas, lo más preferiblemente, cáncer de próstata.

5 Cuando dicha metilación se determina por medios cuantitativos, el punto de corte para determinar dicha presencia de metilación es preferiblemente cero (es decir, cuando una muestra despliega cualquier grado de metilación, se determina que tiene un estado metilado en la posición CpG analizada). Sin embargo, se prevé que el experto en la técnica pueda desear ajustar dicho valor de corte para proporcionar un ensayo de una sensibilidad o especificidad particularmente preferida. Por consiguiente, dicho valor de corte puede aumentarse (aumentando así la especificidad), dicho valor de corte puede estar dentro de un intervalo seleccionado del grupo que consiste en 0%-5%, 5%-10%, 10%-15%, 15%-20%, 20%-30% y 30%-50%. Se prefieren particularmente los puntos de corte 10%, 15%, 25% y 30%.

#### Kits

10 También se divulga aquí un kit que comprende: un medio para determinar la metilación de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTP1. Los medios para determinar dicha metilación comprenden preferiblemente un reactivo que contiene bisulfito; uno o una pluralidad de oligonucleótidos consistentes en secuencias idénticas en cada caso, son complementarios o se hibridan bajo condiciones restrictivas o altamente restrictivas a un segmento de 9 o más preferentemente de 18 segmentos de base de una secuencia  
15 seleccionada de SEQ ID NO: 5 a SEQ ID NO: 20; y opcionalmente instrucciones para llevar a cabo y evaluar el método descrito de análisis de metilación. En una realización, la secuencia de bases de dichos oligonucleótidos comprende al menos un dinucleótido CpG, CpA o TpG.

20 En una realización adicional, dicho kit puede comprender además reactivos estándar para realizar un análisis de metilación específico de posición CpG, en el que dicho análisis comprende una o más de las siguientes técnicas: MS-SNuPE, MSP, Methy-Light™, HeavyMethyl, COBRA y secuenciación de ácidos nucleicos. Sin embargo, un kit como se describe en el presente documento también puede contener sólo parte de los componentes antes mencionados.

25 En una realización preferida, el kit puede comprender reactivos de conversión de bisulfito adicionales seleccionados del grupo que consiste en: 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.

30 En una realización alternativa adicional, el kit puede contener, empaquetados en recipientes separados, una polimerasa y un regulador de reacción optimizado para la extensión del cebador mediado por la polimerasa, tal como PCR. En otra realización, el kit comprende además medios para obtener una muestra biológica del paciente. Se  
35 prefiere un kit, que comprende además un recipiente adecuado para contener los medios para determinar la metilación de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTP1 en la muestra biológica del paciente, y más preferiblemente comprende además instrucciones para el uso e interpretación de los resultados del kit. En una realización preferida, el kit comprende: (a) un reactivo de bisulfito; (b) un recipiente adecuado para contener dicho reactivo de bisulfito y la muestra biológica del paciente; (c) al menos un conjunto de oligonucleótidos cebadores que contienen dos oligonucleótidos cuyas secuencias son idénticas en cada caso, son complementarios o se hibridan bajo condiciones restrictivas o altamente restrictivas a un segmento de 9 o más preferentemente 18 segmentos de base de una secuencia seleccionada de la SEQ ID NO: 5 a SEQ ID NO: 20; y opcionalmente (d) instrucciones para el uso e interpretación de los resultados del kit. En una realización preferida  
40 alternativa el kit comprende: (a) un reactivo de bisulfito; (b) un recipiente adecuado para contener dicho reactivo de bisulfito y la muestra biológica del paciente; (c) al menos un oligonucleótido y/o oligómero de PNA que tiene una longitud de al menos 9 o 16 nucleótidos que es idéntica o hibrida a una secuencia de ácido nucleico pretratada de acuerdo con una de SEQ ID NO: 5 a SEQ ID NO: 20 y sus secuencias complementarias; y opcionalmente (d) instrucciones para el uso e interpretación de los resultados del kit.

45 En una realización alternativa, el kit comprende: (a) un reactivo de bisulfito; (b) un recipiente adecuado para contener dicho reactivo de bisulfito y la muestra biológica del paciente; (c) al menos un conjunto de oligonucleótidos cebadores que contienen dos oligonucleótidos cuyas secuencias son idénticas en cada caso, son complementarios o se hibridan bajo condiciones restrictivas o altamente restrictivas a un segmento de 9 o más preferentemente 18 segmentos de base de una secuencia seleccionada de la SEQ ID NO: 5 a SEQ ID NO: 20; (d) al menos un oligonucleótido y/o oligómero de PNA que tiene una longitud de al menos 9 o 16 nucleótidos que es idéntica o  
50 hibrida a una secuencia de ácido nucleico pretratada de acuerdo con una de SEQ ID NO: 5 a SEQ ID NO: 20 y sus secuencias complementarias; y opcionalmente (e) instrucciones para el uso e interpretación de los resultados del kit.

El kit también puede contener otros componentes tales como reguladores o soluciones adecuadas para bloquear, lavar o recubrir, envasados en un recipiente separado.

55 También se describe un kit para uso en la determinación de la presencia y/o diagnóstico de trastornos proliferativos de células de próstata, más preferiblemente, carcinoma de próstata, comprendiendo dicho kit: un medio para medir el nivel de transcripción de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTP1 y un medio para determinar la metilación de al menos un gen seleccionado del grupo que consiste en

5 RASSF2A; TFAP2E; HIST1H4K y GSTPi. Los reactivos típicos (por ejemplo, como se puede encontrar en un kit típico basado en COBRA™) para el análisis COBRA™ pueden incluir, pero no se limitan a: cebadores de PCR para al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi; enzima de restricción y regulador apropiado; oligo de hibridación génica; oligo de hibridación de control; quinasa para la sonda oligo; y nucleótidos marcados. Los reactivos típicos (por ejemplo, como puede encontrarse en un kit típico de MethyLight™) para el análisis MethyLight™ pueden incluir, pero no se limitan a: cebadores de PCR para la secuencia convertida con bisulfito de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi; sondas específicas de bisulfito (por ejemplo TaqMan™ o Lightcycler™); reguladores y desoxinucleótidos de PCR optimizados; y Taq polimerasa. Los reactivos típicos (por ejemplo, como se puede 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 y desoxinucleótidos de PCR optimizados; kit de extracción de gel; cebadores de control positivo; cebadores Ms-SNuPE™ para la secuencia convertida con bisulfito de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi; regulador de reacción (para la reacción de Ms-SNuPE); y los nucleótidos marcados.

Los reactivos típicos (por ejemplo, como podría encontrarse en un kit basado en MSP típico) para análisis de MSP pueden incluir, pero no se limitan a: cebadores de PCR metilados y no metilados para la secuencia convertida con bisulfito de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi, amortiguadores y desoxinucleótidos de PCR optimizados, y sondas específicas.

20 También se divulga un kit alternativo que comprende un medio para determinar la metilación de al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi, en el que dichos medios comprenden preferiblemente al menos una enzima de restricción específica de la metilación; uno o una pluralidad de oligonucleótidos cebadores (preferiblemente uno o una pluralidad de pares de cebadores) adecuados para la amplificación de una secuencia que comprende al menos un dinucleótido CpG de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4; y opcionalmente instrucciones para llevar a cabo y evaluar el método descrito de análisis de metilación. En una realización, la secuencia de bases de dichos oligonucleótidos son idénticas, son complementarias o se hibridan en condiciones restrictivas o altamente restrictivas con un segmento de por lo menos 18 bases de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4.

30 En una realización adicional, dicho kit puede comprender una o una pluralidad de sondas de oligonucleótidos para el análisis de los fragmentos de digestión, preferiblemente dichos oligonucleótidos son idénticos, son complementarios o hibridan bajo condiciones restrictivas o altamente restrictivas a un segmento de por lo menos 16 bases de longitud de un seleccionad del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4.

35 En una realización preferida, el kit puede comprender reactivos adicionales seleccionados del grupo que consiste en: regulador (por ejemplo, enzima de restricción, PCR, almacenamiento o lavado de reguladores); reactivos o kits de recuperación de ADN (por ejemplo, precipitación, ultrafiltración, columna de afinidad) y componentes de recuperación de ADN.

40 En una realización alternativa adicional, el kit puede contener, empaquetado en recipientes separados, una polimerasa y un regulador de reacción optimizado para la extensión del cebador mediado por la polimerasa, tal como PCR. En otra realización, el kit comprende además medios para obtener una muestra biológica del paciente. En una realización preferida, el kit comprende: (a) un reactivo de enzima de restricción sensible a la metilación; (b) un recipiente adecuado para contener dicho reactivo y la muestra biológica del paciente; (c) al menos un conjunto de oligonucleótidos uno o varios ácidos nucleicos o ácidos nucleicos peptídicos que son idénticos, son complementarios o se hibridan bajo condiciones restrictivas o altamente restrictivas a un segmento de al menos 9 bases de por lo menos una secuencia seleccionada del grupo que consta de SEQ ID NO: 1 a SEQ ID NO: 4; y opcionalmente (d) instrucciones para el uso e interpretación de los resultados del kit.

50 En una realización preferida alternativa el kit comprende: (a) un reactivo de enzima de restricción sensible a la metilación; (b) un recipiente adecuado para contener dicho reactivo y la muestra biológica del paciente; (c) al menos un conjunto de oligonucleótidos cebadores adecuados para la amplificación de una secuencia que comprende al menos un dinucleótido CpG de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4; y opcionalmente (d) instrucciones para el uso e interpretación de los resultados del kit.

55 En una realización alternativa, el kit comprende: (a) un reactivo de enzima de restricción sensible a la metilación; (b) un recipiente adecuado para contener dicho reactivo y la muestra biológica del paciente; (c) al menos un conjunto de oligonucleótidos cebadores adecuados para la amplificación de una secuencia que comprende al menos un dinucleótido CpG de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4; (d) al menos un conjunto de oligonucleótidos uno o varios ácidos nucleicos o ácidos nucleicos peptídicos que son idénticos, son complementarios o se hibridan bajo condiciones restrictivas o altamente restrictivas a un segmento de al menos 9 bases de longitud de una secuencia seleccionada del grupo consistente en la SEQ ID NO: 1 a SEQ ID NO: 4 y opcionalmente (e) instrucciones para el uso e interpretación de los resultados del kit.

El kit también puede contener otros componentes tales como reguladores o soluciones adecuadas para bloquear, lavar o recubrir, envasados en un recipiente separado.

5 También se divulga un kit para uso en proporcionar un diagnóstico de la presencia de trastornos proliferativos de células de próstata, más preferiblemente carcinoma de próstata en un sujeto por medio de análisis de enzimas de restricción sensibles a la metilación. Dicho kit comprende un recipiente y un componente de microarreglo de ADN. Dicho componente de microarreglo de ADN es una superficie sobre la cual una pluralidad de oligonucleótidos se inmoviliza en posiciones designadas y donde el oligonucleótido comprende al menos un sitio de metilación de CpG. Al menos uno de dichos oligonucleótidos es específico para al menos un gen seleccionado del grupo que consiste en RASSF2A; TFAP2E; HIST1H4K y GSTPi y comprende una secuencia de al menos 15 pares de bases de longitud pero no más de 200 bp de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4. Preferiblemente dicha secuencia tiene al menos 15 pares de bases de longitud pero no más de 80 pb de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4. Se prefiere además que dicha secuencia sea al menos 20 pares de bases de longitud pero no más de 30 pb de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4. Dicho kit de prueba comprende además, de preferencia, un componente de enzima de restricción que comprende una o una pluralidad de enzimas de restricción sensibles a la metilación.

En una realización adicional, dicho kit de prueba se caracteriza además porque comprende al menos una enzima de restricción específica de metilación, y en el que los oligonucleótidos comprenden un sitio de restricción de dichas al menos una enzima de restricción específica de metilación.

20 El kit puede comprender además uno o varios de los siguientes componentes, que son conocidos en la técnica para el enriquecimiento de ADN: un componente proteico, uniéndose dicha proteína selectivamente al ADN metilado; un componente de ácido nucleico formador de tríplex, uno o una pluralidad de conectores, opcionalmente en una solución adecuada; sustancias o soluciones para llevar a cabo una ligación, por ejemplo ligasas, reguladores; sustancias o soluciones para realizar una cromatografía en columna; sustancias o soluciones para realizar un enriquecimiento basado en inmunología (por ejemplo inmunoprecipitación); sustancias o soluciones para llevar a cabo una amplificación de ácido nucleico, por ejemplo PCR; un colorante o varios colorantes, si es aplicable con un reactivo de acoplamiento, si es aplicable en una solución; sustancias o soluciones para llevar a cabo una hibridación; y/o sustancias o soluciones para llevar a cabo una etapa de lavado.

30 También se divulga una composición de materia útil para detectar, o para diagnosticar carcinoma de próstata. Comprendiendo dicha composición al menos un ácido nucleico de 18 pares de bases de longitud de un segmento de la secuencia de ácido nucleico descrita en la SEQ ID NO: 5 a la SEQ ID NO: 20 y una o más sustancias tomadas del grupo que comprende: 1-5 mM de cloruro de magnesio, dNTP 100-500 μM, 0.5-5 unidades de taq polimerasa, albúmina de suero bovino, un oligómero en particular un oligonucleótido u oligómero peptídico de ácido nucleico (PNA), comprendiendo dicho oligómero al menos una secuencia de bases que tiene una longitud de al menos 9 nucleótidos que es complementaria o hibrida en condiciones moderadamente restrictivas o restrictivas a un ADN genómico pretratado de acuerdo con una de las SEQ ID NO: 5 a SEQ ID NO: 20 y secuencias complementarias a las mismas. Se prefiere que dicha composición de materia comprenda una solución reguladora apropiada para la estabilización de dicho ácido nucleico en una solución acuosa y que permita reacciones basadas en polimerasa dentro de dicha solución. Los reguladores adecuados son conocidos en la técnica y están comercialmente disponibles.

Preferiblemente, dicho al menos un ácido nucleico tiene al menos 50, 100, 150, 200, 250 o 500 pares de bases de longitud de un segmento de la secuencia de ácido nucleico descrita en la SEQ ID NO: 5 a la SEQ ID NO: 20.

45 Aunque la presente invención se ha descrito con especificidad de acuerdo con algunas de sus realizaciones preferidas, los siguientes ejemplos sirven sólo para ilustrar la invención y no pretenden limitar la invención dentro de los principios y el alcance de las interpretaciones más amplias y configuraciones equivalentes de la misma.

#### Ejemplo 1

50 El objetivo del presente estudio era determinar la viabilidad de medir los marcadores de metilación del ADN para el cáncer de próstata (en lo sucesivo también denominado PCa) en fluidos corporales remotos. En este proceso se utilizó un flujo de procedimiento de alta calidad para la orina, se analizaron los marcadores candidatos mediante la tecnología de HeavyMethyl™ (HM) (Cottrell et al., Nucleic Acids Res., 2004 Jan 13; 32(1): e10) y se demostró que el PCa desprende ADN que puede ser detectado por medio del análisis de metilación tanto en plasma como en orina con alta sensibilidad. Se estableció así que los marcadores analizados eran adecuados para el desarrollo de una prueba de cribado para PCa basada en el análisis de metilación del ADN.

Objetivos del estudio

El propósito del presente estudio era realizar una investigación para determinar si los marcadores de metilación del ADN de PCa se pueden medir en un fluido corporal remoto. El estudio fue diseñado para identificar el analito óptimo para tal prueba y para generar especificidad y datos de rendimiento analítico para los candidatos marcadores.

5 Marcadores de candidato y ubicación de los ensayos

Los marcadores RASSF2 y TFAP2E se identificaron en base a su metilación en tejidos de cáncer de próstata, como se determinó en un estudio preliminar (no descrito en el presente documento). Los marcadores GSTPi y HIST1H4K habían sido previamente identificados en un estudio del solicitante publicado en la Solicitud de Patente WO 2005/054517.

10 El análisis de metilación se realizó por medio del HeavyMethyl™. El genómico aislado es tratado con bisulfito para convertir citosinas no metiladas en uracilo, en las que se conservan las citosinas metiladas. Los fragmentos del ADN tratado con bisulfito que comprenden dinucleótidos CpG potencialmente metilados se amplifican a continuación mediante PCR. Los cebadores no cubren ninguna posición de citosina potencialmente metilada (es decir, no se hibridan con dinucleótidos CpG genómicos). La amplificación de fragmentos que comprenden dinucleótidos CpG no metilados se suprime por medio de un oligonucleótido de bloqueo que se hibrida con los dinucleótidos TG. Por consiguiente, sólo el ADN que se metiló en la muestra genómica se amplifica. Los fragmentos amplificados se detectan por medio de sondas marcadas detectablemente adecuadas para su uso en reacciones de PCR tales como sondas de detección en tiempo real.

20 El cebador de ensayo y las sondas se proporcionan en el listado de secuencias adjunto como se indica en la Tabla 2.

GSTPi

Ubicación cromosómica: 11q13

Gen(es) cercano(s): El cebador directo GSTPi HM está justo corriente arriba del exón 1 y el cebador inverso está justo corriente abajo del exón 1 de GSTP1

25 RASSF2A

Ubicación cromosómica: 20pteros-p12.1

Gen(es) cercano(s): con/en la isla CpG del intrón 1 de la transcripción v. 1 de RASSF2

HIST1H4K

Ubicación cromosómica: 6p22-21.3

30 Gen(es) cercano(s): se superpone a HIST1H4K sin intrones

TFAP2E

Ubicación cromosómica: 1p34.3

Gene(s) cercano(s): con/en intrón 3 de TFAP2E (~11 kb corriente abajo de txn de inicio) y ~20 kb corriente arriba de KIAA0319L txn de inicio (PKD-1 como gen)

35 Estudio de tejidos

40 El ensayo fue probado inicialmente en tejidos normales, NAT y PCa. Los candidatos marcadores que se analizaron en la prueba de tejido HM fueron muy específicos para la sangre normal y el tejido prostático normal. En contraste con los estudios anteriores se observó que el ADN de la tumor adyacente normal de próstata (NAT) es casi tan metilado como el ADN del tumor de próstata. El NAT (que puede contener BPH) es claramente distinto del tejido de BPH que se ha derivado de pacientes que no son portadores de tumor de próstata sin PSA elevado (el origen de muchas muestras de tejido de BPH en la prueba de tejido de MSP). De hecho, hay evidencia en la literatura que GSTP1 en NAT está metilado (Hanson et al., 2006).

El rendimiento de los marcadores en +BPH normal en comparación con PCa se proporciona en la Tabla 3 y la figura

1.

Análisis remoto de analitos

Con el fin de maximizar el equivalente de analito en ensayos de PCR en tiempo real (equivalentes de 1.5 ml), el número máximo de ensayos en el estudio se limitó a cuatro, realizado cada ensayo en duplicado para cada muestra.

5 Recolección de muestra

Para este estudio, se recolectaron muestras de plasma y orina de un total de 191 hombres, incluyendo 91 varones con cáncer de próstata confirmado por biopsia, 51 varones sin cáncer detectados por biopsia (posteriormente diagnosticados con HPB) y 50 jóvenes varones sanos. En todos los análisis, la clase positiva se compone de las muestras PCa.

10 Al diseñar el presente estudio, la definición de la clase negativa fue un problema ya que no existe un método de detección que excluya la presencia de PCa con 100% de certeza. La biopsia tiene una tasa de diagnóstico negativo falso de al menos el 10% (Djavan et al., 2000; Ming et al., 2002; Gupta et al., 2005; Hanley et al., 2006), mientras que la medición del PSA es propensa tanto a los falsos negativos y falsos positivos. Debido a que el objetivo principal del estudio era demostrar la viabilidad de medir los marcadores metilados de PCa en un fluido corporal remoto, se centró en una clase negativa que minimizó la probabilidad de falsos positivos. En consecuencia, los jóvenes varones sanos fueron elegidos como la clase negativa "verdadera". Se argumentó que los varones jóvenes sanos sin antecedentes familiares de cáncer de próstata deberían ser verdaderamente negativos para la PCa.

15 Debido a que una realización de la prueba PCa es como un seguimiento de diagnóstico de PSA, también se incluyó una segunda clase negativa de biopsia negativa, muestras de HBP. Un factor potencialmente confuso en esta clase es la probable presencia de biopsias falsas negativas.

20 En cinco casos de PCa, sólo se recogió una muestra de plasma y en diez casos adicionales sólo se recogió una muestra de orina. Las muestras se recogieron en múltiples sitios. La orina se recogió después de un masaje prostático, las muestras tanto de plasma como de orina se obtuvieron antes de cualquier tratamiento para PCa. Los criterios de inclusión y exclusión fueron diseñados para asegurar que los pacientes analizados reflejen los pacientes potenciales que usarían las pruebas de detección de PCa.

25 Los siguientes criterios de inclusión y exclusión se aplicaron a los pacientes sometidos a biopsia: Criterios de inclusión:

- Indicación de biopsia (PSA elevado y/o DRE sospechoso)
- Biopsia programada dentro de 1 semana después de la recolección de la muestra
- 30 • Edad 40-80

Criterios de exclusión:

- Cualquier tratamiento previo para el cáncer de próstata
- Historial de cáncer o enfermedad grave en los últimos 5 años
- Síntomas de la infección del tracto urinario

35 Los siguientes criterios se aplicaron a los hombres sanos del grupo de control:

Criterios de inclusión:

- Masculino
- Edad 18-30

Criterios de exclusión:

40 Cualquier tratamiento previo o síntomas de cáncer de próstata o enfermedad de la próstata

Historial de cáncer o enfermedad grave en los últimos 5 años

Síntomas de infección del tracto urinario

Datos del paciente y características del tumor

5 La Tabla 4 muestra la puntuación de Gleason (cuando corresponda) de las muestras de pacientes. Los valores medios de PSA para las muestras de cáncer de próstata, HGPIN y biopsia negativa fueron  $18.2 \pm 33.1$ ,  $7.0 \pm 3.0$  y  $8.8 \pm 5.2$ , respectivamente. El cáncer de próstata, HGPIN y las clases negativas de biopsia se diagnosticaron después de la recolección de muestras a través de la biopsia de próstata. El número medio de núcleos de biopsia para todas las clases de muestra fue 8, aunque hubo alguna variación entre los proveedores.

10 La extracción de ADN y el tratamiento con bisulfito se llevaron a cabo de acuerdo con protocolos estandarizados. Para cada ensayo, se hicieron funcionar 1.5 ml de equivalente de analito por duplicado.

Rendimiento de los marcadores, consideraciones generales

15 El objetivo inicial del estudio era desarrollar un panel de marcadores dirigidos como seguimiento de diagnóstico a pruebas de PSA de 2.5 ng/ml o más para hombres mayores de 50 años de edad para discriminar el cáncer de próstata más específica que competiría con pruebas de PSA debido a un rendimiento superior. En el presente estudio analizamos los datos de dos maneras diferentes: (i) utilizamos cáncer de próstata y muestras de biopsia negativa para evaluar el desempeño de los marcadores en el seguimiento de la prueba de PSA (aplicación diagnóstica) y (ii) usamos cáncer de próstata y todas las muestras no cancerígenas (biopsias negativas y saludables) para medir el desempeño de los marcadores en la prueba de detección (aplicación de tamizaje).  
20 Presentamos el rendimiento de los marcadores para el plasma y la orina por separado, también proporcionamos análisis de datos para marcadores individuales y paneles de marcadores. Todos los datos se informan como valores de metilación brutos logarítmicos medios.

25 Como prueba de selección primaria, el panel marcador preferiblemente identificaría PCa en hombres mayores de 50 años con una especificidad mejorada con respecto a PSA. Todos los análisis de aplicaciones de cribado utilizan las muestras de PCa como la clase positiva. Para los fines del presente estudio, analizamos los datos para la aplicación de cribado con dos clases negativas alternativas. La primera clase negativa analizó a los 50 varones sanos jóvenes con mínima probabilidad de PCa no detectada. Aunque esta clase negativa representa un negativo de prueba "verdadera", no está adaptada a la edad para la población de cribado de PCa objetivo y no incluye ninguna clase probable de falsos positivos, por ejemplo BPH. Por lo tanto, realizamos un segundo análisis en el que se analizaron los 50 controles jóvenes sanos y los 51 controles negativos de biopsia como una clase negativa de 101 muestras.  
30

35 En promedio, se realizan aproximadamente 20.000.000 pruebas de PSA cada año en los EE.UU. con sólo aproximadamente 1.000.000 de casos que llegan a biopsia (de los cuales aproximadamente 750.000 biopsias son innecesarias). Por lo tanto, menos del 5% de los individuos que son cribados actualmente por PSA caen en la clase negativa que se representa por PSABPH positivo elevado, mientras que la gran mayoría de la población de cribado de objetivo cae en el PSA de baja clase negativa. Mientras que la clase negativa de solo varones jóvenes sanos puede representar una sobrestimación de la capacidad discriminatoria de nuestros marcadores, la clase negativa combinada de varones jóvenes sanos más varones negativos de biopsia igualada por edad puede representar una subestimación de la capacidad discriminatoria de nuestros marcadores.

40 La sensibilidad y la especificidad de marcadores individuales (sencillos) probados por PCR en tiempo real en orina postmasaje prostático de pacientes con cáncer de próstata versus pacientes con biopsia negativa y sujetos de control sanos se muestran en la Tabla 5. La figura 2 muestra los ensayos HM en tiempo real de PCR de orina postmasaje prostático de PCa y clase I negativa (individuos sanos). La figura 3 muestra los ensayos de PCR en tiempo real HM de orina postmasaje prostático de PCa y clase II negativa (individuos sanos más biopsias negativas).

45 La sensibilidad y especificidad de los marcadores individuales (sencillos) probados por PCR en tiempo real en plasma de pacientes de cáncer de próstata frente a pacientes con biopsia negativa y sujetos de control sanos se muestra en la Tabla 6. La figura 4 muestra los ensayos de PCR HM en tiempo real de plasma de PCa y clase I negativa (individuos sanos). La figura 5 muestra los ensayos de PCR en tiempo real HM de plasma de PCa y clase II negativa (individuos sanos más biopsias negativas).

50 Como se ilustra en la Tabla 7, en todas las comparaciones negativas de clase y para todos los marcadores, la orina fue el analito más sensible.

Correlación de los marcadores con la puntuación de Gleason

Las cantidades crecientes de ADN marcador metilado se correlacionaron con el aumento de la puntuación de Gleason para todos los marcadores en plasma. Esto fue cierto para muestras con grandes cantidades de ADN marcador metilado en la orina (véanse especialmente los marcadores TFAP2E y RASSF2A), pero en general la correlación fue menos fuerte en el ADN de la orina que en el ADN del plasma. El PSA como marcador de PCa en pacientes con PSA elevado (>4 ng/ml) también se correlacionó con el aumento de la puntuación de Gleason.

El rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase I negativa (varones sanos) en la orina se proporciona en la Tabla 8.

El rendimiento de los paneles marcadores de detección para distinguir PCa de la clase II negativa (varones sanos más biopsia negativa) en la orina, se proporciona en la Tabla 9.

El rendimiento de los paneles marcadores de selección para distinguir PCa de la clase I negativa (varones sanos) en plasma se proporciona en la Tabla 10.

El rendimiento de los paneles marcadores de selección para distinguir PCa de la clase II negativa (varones sanos más biopsia negativa) en plasma se proporciona en la Tabla 11.

Rendimiento del marcador en la aplicación de diagnóstico: Seguimiento al PSA

Como marcador de seguimiento de la prueba de PSA, los marcadores preferiblemente identificarían PCa en hombres mayores de 50 años que han sido clasificados como individuos de alto riesgo debido a PSA elevado (>2.5 ng/ml). Esta es una aplicación y un análisis distintos y requiere una mayor discriminación en comparación con la prueba de detección. Los falsos positivos en esta aplicación surgen de la elevada PSA, biopsia negativa clase BPH. Una vez más, las muestras PCa representan la clase positiva. Para los fines de una aplicación de seguimiento de diagnóstico, analizamos los datos utilizando una sola clase negativa compuesta por las 51 muestras negativas de biopsia. Los AUC de los marcadores probados por PCR en tiempo real en orina postmasaje prostático y plasma de pacientes con cáncer de próstata y pacientes con biopsia negativa se proporciona en la Tabla 12.

A partir de dicha tabla puede observarse que para todos los marcadores de metilación analizados, la orina fue el analito más sensible. Para el PSA total (tratado aquí como un marcador adicional para determinar si hay más información proporcionada después de la indicación >4 ng/ml para la biopsia), no hubo diferencias en la sensibilidad entre la orina y el plasma.

Rendimiento de los paneles marcadores

Con el fin de proporcionar una mejor precisión, las combinaciones de marcadores se evaluaron tanto cualitativa como cuantitativamente.

La Tabla 13 proporciona el rendimiento de los paneles marcadores de diagnóstico para distinguir la PCa de la biopsia negativa en la orina.

La Tabla 14 proporciona el rendimiento de los paneles marcadores de diagnóstico para distinguir PCa de biopsia negativa en plasma.

Discusión

El estudio se realizó sobre muestras de plasma y/u orina recogidas de 91 pacientes con PCa, 51 pacientes con biopsia negativa (diagnosticada con HBP) y 50 jóvenes sanos. Se utilizaron ensayos de PCR en tiempo real HM<sup>TM</sup> para medir la metilación del ADN de los marcadores candidatos. La cantidad de ADN marcador metilado se correlacionó con PCa tanto en plasma como en orina, con ADN de orina mostrando mayor sensibilidad. Como prueba de detección (discriminación del cáncer de PCa de controles sanos usando analito de orina), los candidatos marcadores de anclaje GSTPi, RASSF2A, HIST1H4K y TFAP2E tienen una sensibilidad del 63%, 74%, 69% y 47% con una especificidad del 96%, respectivamente. Como marcador de seguimiento de la prueba de PSA (discriminación de PCa de controles negativos de biopsia, todos con PSA elevado), los marcadores tienen una sensibilidad del 23%, 18%, 28% y 23% con especificidad del 95%, respectivamente. Un panel de selección cuantitativo de marcadores RASSF2A y HIST1H4K produjo un 94% de sensibilidad con una especificidad del 88% frente a individuos sanos. Un panel de diagnóstico cuantitativo de los marcadores GSTPi y PSA arrojó un 83% de sensibilidad con un 45% de especificidad. El desempeño de estos marcadores se compara bien con el desempeño de PSA (18% de sensibilidad al 98% de especificidad para hombres <60 años y 19% de sensibilidad al 94% de especificidad para hombres >60 años) en la población de cribado (Punglia et al.). La metilación de todos los marcadores se correlacionó bien con la puntuación de Gleason en el ADN plasmático, pero la correlación fue menos fuerte en el ADN de la orina.

Conclusiones

5 Al completarse la presente investigación, se demostró que los biomarcadores de cáncer de próstata basados en ADN metilado pueden medirse en plasma y orina, mostrando el ADN de orina una mayor sensibilidad que el plasma. Además, se identificaron los marcadores de metilación del ADN que discriminan a los pacientes con PCa de controles sanos y aquellos con hiperplasia benigna de próstata (HBP). Las principales conclusiones del presente estudio son las siguientes:

- Los marcadores metilados de cáncer de próstata se pueden medir tanto en plasma como en orina de pacientes con PCa.
- Identificación de marcadores que discriminan a pacientes con PCa de aquellos sin PCa.

10

Tabla 1: Genes y secuencias de acuerdo con la presente divulgación

Gen	SEQ ID NO: Genómica	Cadena en sentido metilada convertida con bisulfito	Cadena antisentido metilada convertida con bisulfito	Cadena en sentido no metilada convertida con bisulfito	Cadena antisentido no metilada convertida con bisulfito
RASSF2A	1	5	6	13	14
TFAP2E	2	7	8	15	16
HIST1H4K	3	9	10	17	18
GSTPi	4	11	12	19	20

Tabla 2: Componentes de ensayo según el Ejemplo 1

Gen	Cebador de avance	Cebador de retroceso	Bloqueador	Oligo de detección
GSTPi	21	22	23	24
HIST1H4K	25	26	27	28
RASSF2A	29	30	31	32
TFAP2E	33	34	35	36

Tabla 3: Análisis de rendimiento de los marcadores (normal más BPH y PCa) en la prueba en tejido de acuerdo con el Ejemplo 1.

Marcador	AUC	Sensibilidad	Especificidad
GSTPi	0.90	0.83	0.91
HIST1H4K	0.91	0.83	0.91
RASSF2A	0.93	0.75	0.91
TFAP2E	NA	NA	NA

Tabla 4: Muestras remotas según el Ejemplo 1.

Tipo de muestra	No. de muestras
Cáncer de Próstata - Puntuación Gleason	
4	1
5	5
6	33
7	39
8	8
9	4
No hay puntuación disponible	1
Cáncer de Próstata Total:	91
Biopsia negativa	51
Control sano	50

Tabla 5: Sensibilidad y especificidad de los marcadores individuales probados por PCR en tiempo real en orina postmasaje prostático de pacientes con cáncer de próstata, pacientes con biopsia negativa y sujetos de control sanos.

Marcador	Negativo Clase I: Sano			Negativo Clase II: Sano + Biopsia (-)		
		Sens/Espec	Valor de Wilcoxon p	AUC	Sens/Espec	Valor de Wilcoxon p
GSTPi	0.89	0.63/0.96	0	0.79	0.31/0.96	0
RASSF2A	0.90	0.74/0.96	0	0.79	0.24/0.96	0
HIST1H4K	0.91	0.69/0.96	0	0.77	0.36/0.96	0
TFAP2E	0.86	0.47/0.96	0	0.77	0.27/0.96	0

Tabla 6: Sensibilidad y especificidad de los marcadores individuales probados por PCR en tiempo real en plasma de pacientes con cáncer de próstata, pacientes con biopsia negativa y sujetos de control sanos.

Marcador	Negativo Clase I: Sano			Negativo Clase II: Sano + Biopsia (-)		
		Sens/Espec	Valor de Wilcoxon p	AUC	Sens/Espec	Valor de Wilcoxon p
GSTPi/GSTP1	0.61	0.17/0.96	0.0063	0.58	0.71/0.95	0.0183
RASSF2A	0.68	0.37/1.00	0	0.64	0.20/0.95	0

ES 2 615 354 T3

HIST1H4K	0.64	0.26/0.96	$5e^{-04}$	0.56	0.16/0.95	0.0572
TFAP2E	0.16	0.22/1.00	$4e^{-04}$	0.56	0.09/0.95	0.0128

Tabla 7

Marcador	Negativo Clase I: Salno		Negativo Clase II: Sano + Biopsia (-)	
	AUC de la orina	AUC del plasma	AUC de la orina	AUC del plasma
GSTPi/GSTP1	0.89	0.61	0.69	0.55
RASSF2A	0.90	0.68	0.66	0.60
HIST1H4K	0.91	0.64	0.64	0.50
TFAP2E	0.86	0.61	0.65	0.52

Tabla 8: Rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase I negativa (varones sanos) en la orina

Panel de marcadores	% Sens PCa	% Espec. Sano
Marcadores cuantitativos individuales:		
RASSF2A	74	96
HIST1H4K	69	96
GSTPi	63	96
TFAP2E	46	100
Paneles Cualitativos:		
GSTPi+HIST1H4K	79	98
RASSF2A+HIST1H4K	94	88
Paneles Cualitativos:		
RASSF2A+HIST1H4K	94	88
quadSVM (todos los marcadores, sin PSA)	79	98

Tabla 9

Panel de marcadores	% Sens PCa	% Espec. Sano + Biopsia (-)
Marcadores únicos cuantitativos:		
RASSF2A	74	76

ES 2 615 354 T3

HIST1H4K	69	68
GSTPi	63	80
TFAP2E	46	88
Paneles Cualitativos:		
GSTPi+HIST1H4K	79	72
RASSF2A+HIST1H4K	94	54
Paneles Cualitativos:		
RASSF2A+HIST1H4K	94	58
quadSVM (todos los marcadores, sin PSA)	79	76

Tabla 10: Rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase I negativa (varones sanos) en plasma

Panel de marcadores	% Sens PCa	% Espec. Sano
Marcadores únicos cuantitativos:		
RASSF2A	37	100
HIST1H4K	26	96
GSTPi	17	94
TFAP2E	22	100
Paneles Cualitativos:		
RASSF2A+HIST1H4K	41	98
Paneles Cualitativos:		
RASSF2A+TFAP2E (TFAP2E utilizado para normalizar)	32	100
quadSVM (todos los marcadores, sin PSA)	39	96

Tabla 11

Panel de marcadores	% Sens PCa	% Espec. Sano + Biopsia (-)
Marcadores cuantitativos individuales:		
RASSF2A	37	91
HIST1H4K	26	88
GSTPi	17	95

ES 2 615 354 T3

TFAP2E	22	92
Paneles cualitativos:		
RASSF2A+HIST1H4K	41	88
Paneles cualitativos:		
RASSF2A+TFAP2E (TFAP2E utilizado para normalizar)	32	92
quadSVM (todos los marcadores, sin PSA)	39	94

Tabla 12

Marcador	PCa vs. Biopsia (-)	
	AUC de la orina	AUC del plasma
GSTPi/ GSTP1	0.69	0.55
RASSF2 A	0.66	0.60
HIST1H4K	0.64	0.50
TFAP2E	0.65	0.52
***PSA	0.56	0.56
*** Prueba si el PSA contiene información adicional más allá de lo que fue aportado por la indicación de corte de >4 ng/ml para la biopsia de próstata.		

Tabla 13

Panel marcador	% Sens PCa	% Espec. Biopsia (-)
Marcadores cuantitativos individuales:		
RASSF2A	74	55
HIST1H4K	69	41
GSTPi	63	64
TFAP2E	46	77
Paneles cualitativos:		
GSTPi+HIST1H4K	79	46
RASSF2A+HIST1H4K	94	21
Paneles cualitativos:		
RASSF2A+HIST1H4K	94	27

GSTPi+PSA	83	45
quadSVM (todos marcadores, sin PSA)	79	55

Tabla 14

Panel marcador	% Sens PCa	% Espec. Biopsia (-)
Marcadores únicos cuantitativos:		
RASSF2A	37	82
HIST1H4K	26	79
GSTPi	17	96
TFAP2E	22	84
Paneles cualitativos:		
RASSF2A+HIST1H4K	41	79
Paneles cualitativos:		
RASSF2A+TFAP2E (TFAP2E usado para normalizar)	32	85
RASSF2A+TFAP2E+PSA (TFAP2E usado para normalizar)	94	22
quadSVM (todos los marcadores, sin PSA)	39	91
quadSVM (todos los marcadores+PSA)	48	87

Ejemplo 1

- 5 El objetivo del presente estudio fue determinar la viabilidad de medir los marcadores de metilación del ADN para el cáncer de próstata (en lo sucesivo también denominado PCa) en fluidos corporales remotos. En este proceso se utilizó un flujo de procedimiento de alta calidad para la orina, se analizaron los marcadores candidatos mediante la tecnología de HeavyMethyl™ (HM) (Cottrell et al., Nucleic Acids Res., 2004 Jan 13; 32(1): e10 y se demostró que el PCa desprende ADN que puede detectarse mediante análisis de metilación tanto en plasma como en orina con alta sensibilidad, estableciendo así que los marcadores analizados son adecuados para el desarrollo de una prueba de cribado para PCa basada en el análisis de metilación del ADN.
- 10

Objetivos del estudio

- 15 El propósito del presente estudio fue realizar una investigación para determinar si los marcadores de metilación del ADN de PCa se pueden medir en un fluido corporal remoto. El estudio fue diseñado para identificar el analito óptimo para tal prueba y para generar especificidad y datos de rendimiento analítico para los candidatos marcadores.

Marcadores de candidato y ubicación de los ensayos

- 20 Los marcadores RASSF2 y TFAP2E se identificaron en base a su metilación en tejidos de cáncer de próstata, como se determinó en un estudio preliminar (no descrito en el presente documento). Los marcadores GSTPi y HIST1H4K habían sido previamente identificados en un estudio del solicitante publicado en la Solicitud de Patente WO 2005/054517.

El análisis de metilación se realizó por medio del HeavyMethyl™. El genómico aislado es tratado con bisulfito para convertir citosinas no metiladas en uracilo, en las que se conservan las citosinas metiladas. Los fragmentos del ADN

- 5 tratado con bisulfito que comprenden dinucleótidos CpG potencialmente metilados se amplifican a continuación mediante PCR. Los cebadores no cubren ninguna posición de citosina potencialmente metilada (es decir, no se hibridan con dinucleótidos CpG genómicos). La amplificación de fragmentos que comprenden dinucleótidos CpG no metilados se suprime por medio de un oligonucleótido de bloqueo que se hibrida con los dinucleótidos TG. Por consiguiente, sólo el ADN que se metiló en la muestra genómica se amplifica. Los fragmentos amplificados se detectan por medio de sondas marcadas detectablemente adecuadas para su uso en reacciones de PCR tales como sondas de detección RealTime.
- El cebador de ensayo y las sondas se proporcionan en el listado de secuencias adjunto como se indica en la Tabla 2.
- 10 **GSTPi**
- Ubicación cromosómica: 11q13
- Gen(es) cercano(s): El cebador directo GSTPi HM está justo corriente arriba del exón 1 y el cebador inverso está justo corriente abajo del exón 1 de GSTP1
- RASSF2A
- 15 Ubicación cromosómica: 20pter-p12.1
- Gen(es) cercano(s): con/en la isla CpG del intrón 1 de la transcripción del v. 1 de RASSF2
- HIST1H4K
- Ubicación cromosómica: 6p22-21.3
- Gen(es) cercano(s): Se superpone a HIST1H4K sin intrones
- 20 **TFAP2E**
- Ubicación cromosómica: 1p34.3
- Gene(s) cercano(s): con/en intrón 3 de TFAP2E (~11 kb corriente abajo de txn de inicio) y ~20 kb corriente arriba de KIAA0319L txn de inicio (PKD-1 como gen)
- Estudio de tejidos
- 25 El ensayo se probó inicialmente en tejidos normales, NAT y PCa. Los candidatos marcadores que se analizaron en la prueba de tejido HM fueron muy específicos para la sangre normal y el tejido prostático normal. En contraste con los estudios anteriores se observó que el ADN de la próstata normal tumor adyacente (NAT) es casi tan metilado como el ADN del tumor de próstata. El NAT (que puede contener BPH) es claramente distinto del tejido de BPH que se ha derivado de pacientes que no son portadores de tumor de próstata sin PSA elevado (el origen de muchas muestras de tejido de BPH en la prueba de tejido de MSP). De hecho, hay evidencia en la literatura que GSTP1 en NAT está metilado (Hanson et al., 2006).
- 30 El rendimiento de los marcadores en normal + BPH en comparación con PCa se proporciona en la Tabla 3 y la figura 1.
- Análisis remoto de analitos
- 35 Con el fin de maximizar el equivalente de analito en ensayos de PCR en tiempo real (equivalentes de 1.5 ml), el número máximo de ensayos en el estudio se limitó a cuatro, realizándose cada ensayo en duplicado para cada muestra.
- Recolección de muestras
- 40 Para este estudio, se recogieron muestras de plasma y orina de un total de 191 hombres, incluyendo 91 hombres con cáncer de próstata confirmado por biopsia, 51 varones sin cáncer detectado por biopsia (posteriormente diagnosticados con HPB) y 50 jóvenes varones sanos. En todos los análisis, la clase positiva se compone de las muestras PCa.

- 5 Al diseñar el presente estudio, la definición de la clase negativa fue un problema ya que no existe un método de detección que excluya la presencia de PCa con 100% de certeza. La biopsia tiene una tasa de diagnóstico negativo falso de al menos el 10% (Djavan et al., 2000; Ming et al., 2002; Gupta et al., 2005; Hanley et al., 2006), mientras que la medición del PSA es propensa tanto a los falsos negativos como a los falsos positivos. Debido a que el objetivo principal del estudio era demostrar la viabilidad de medir los marcadores metilados de PCa en un fluido corporal remoto, se centró en una clase negativa que minimizó la probabilidad de falsos positivos. En consecuencia, los jóvenes varones sanos fueron elegidos como la clase negativa "verdadera". Se argumentó que los varones jóvenes sanos sin antecedentes familiares de cáncer de próstata deberían ser verdaderamente negativos para la PCa.
- 10 Debido a que una realización del ensayo de PCa es como un seguimiento de diagnóstico al PSA, también se incluyó una segunda clase negativa de muestras de BPH negativas de biopsia. Un factor potencialmente confuso en esta clase es la probable presencia de biopsias falsas negativas.
- 15 En cinco casos de PCa, sólo se recogió una muestra de plasma y en diez casos adicionales solo se recogió una muestra de orina. Las muestras se recogieron en múltiples sitios. La orina se recogió después de un masaje prostático, tanto plasma y muestras de orina se obtuvieron antes de cualquier tratamiento para PCa. Los criterios de inclusión y exclusión fueron diseñados para asegurar que los pacientes analizados reflejen los pacientes potenciales que usarían las pruebas de detección de PCa.
- Los siguientes criterios de inclusión y exclusión se aplicaron a los pacientes sometidos a biopsia:
- Criterios de inclusión:
- Indicación de biopsia (PSA elevado y/o DRE sospechoso)
- 20 • Biopsia programada dentro de 1 semana después de la recolección de la muestra
- Edad 40-80
- Criterios de exclusión:
- Cualquier tratamiento previo para el cáncer de próstata
- Historial de cáncer o enfermedad grave en los últimos 5 años
- 25 Síntomas de infección del tracto urinario
- Los siguientes criterios se aplicaron a los hombres sanos del grupo de control:
- Criterios de inclusión:
- Masculino
  - Edad 18-30
- 30 Criterios de exclusión:
- Cualquier tratamiento previo o síntomas de cáncer de próstata o enfermedad de la próstata
- Historial de cáncer o enfermedad grave en los últimos 5 años
- Síntomas de infección del tracto urinario
- Datos del paciente y características del tumor
- 35 La Tabla 4 muestra la puntuación de Gleason (cuando corresponda) de las muestras de pacientes. Los valores medios de PSA para las muestras de cáncer de próstata, HGPIN y biopsia negativa fueron  $18.2 \pm 33.1$ ,  $7.0 \pm 3.0$  y  $8.8 \pm 5.2$ , respectivamente. El cáncer de próstata, HGPIN y las clases negativas de biopsia se diagnosticaron después de la toma de muestras a través de la biopsia de próstata. El número medio de núcleos de biopsia para todas las clases de muestra fue 8, aunque hubo alguna variación entre los proveedores.
- 40 La extracción de ADN y el tratamiento con bisulfito se llevó a cabo de acuerdo con protocolos estandarizados. Para

cada ensayo, se hicieron funcionar 1.5 ml de equivalente de analito por duplicado.

#### Rendimiento del marcador, consideraciones generales

5 El objetivo inicial del estudio era desarrollar un panel de marcadores dirigidos como un seguimiento de diagnóstico a pruebas de PSA de 2.5 ng/ml o más para hombres mayores de 50 años de edad para discriminar el cáncer de próstata de condiciones no cancerosas. Tal prueba podría ampliarse como una prueba de detección de cáncer de próstata más específica que competiría con pruebas de PSA debido a un rendimiento superior. En el presente estudio analizamos los datos de dos maneras diferentes: (i) utilizamos muestras de cáncer de próstata y biopsia negativa para evaluar el desempeño de los marcadores en el seguimiento de la prueba de PSA (aplicación diagnóstica) y (ii) usamos muestras de cáncer de próstata y todas las no cancerígenas (biopsias negativas y saludables) para medir el desempeño de los marcadores en la prueba de detección (aplicación de tamizaje). Presentamos el rendimiento de los marcadores para el plasma y la orina por separado, también proporcionamos análisis de datos para marcadores individuales y paneles de marcadores. Todos los datos se informan como valores de metilación brutos medios logarítmicos.

15 Como prueba de selección primaria, el panel marcador preferiblemente identificaría PCa en hombres mayores de 50 años con una especificidad mejorada con relación al PSA. Todos los análisis de aplicaciones de cribado utilizan las muestras de PCa como la clase positiva. Para los fines del presente estudio, analizamos los datos para la aplicación de cribado con dos clases negativas alternativas. La primera clase negativa analizó a los 50 varones sanos jóvenes con mínima probabilidad de PCa no detectada. Aunque esta clase negativa representa un negativo de prueba "verdadera", no está adaptada a la edad para la población de cribado de PCa objetivo y no incluye ninguna clase probable de falsos positivos, por ejemplo BPH. Por lo tanto, realizamos un segundo análisis en el que se analizaron los 50 controles jóvenes sanos y los 51 controles negativos de biopsia como una clase negativa de 101 muestras.

25 En promedio, se realizan aproximadamente 20.000.000 de exámenes de PSA cada año en los EE.UU. con sólo aproximadamente 1.000.000 de casos que llegan a biopsia (de los cuales aproximadamente 750.000 biopsias son innecesarias). Por lo tanto, menos del 5% de los individuos que están actualmente seleccionados por PSA caen en la clase negativa que se representa por PSABPH positivo, mientras que la gran mayoría de la población con cribado de objetivo cae en la clase negativa de PSA baja. Mientras que la clase negativa de solo varones jóvenes sanos puede representar una sobrestimación de la capacidad discriminatoria de nuestros marcadores, la clase negativa combinada de varones jóvenes sanos más varones negativos de biopsia igualada por edad puede representar una subestimación de la capacidad discriminatoria de nuestros marcadores.

30 La sensibilidad y especificidad de marcadores individuales (sencillos) probados por PCR en tiempo real en orina postmasaje prostático de pacientes con cáncer de próstata frente a pacientes con biopsia negativa y sujetos de control sanos se muestra en la Tabla 5. La figura 2 muestra los ensayos de PCR en tiempo real HM de orina postmasaje prostático de PCa y clase I negativa (individuos sanos). La figura 3 muestra los ensayos de PCR en tiempo real HM de orina postmasaje prostático de PCa y clase II negativa (individuos sanos más biopsias negativas).

35 La sensibilidad y especificidad de marcadores individuales (sencillos) probados por PCR en tiempo real en plasma procedente de pacientes de cáncer de próstata frente a pacientes con biopsia negativa y sujetos de control sanos se muestra en la Tabla 6. La figura 4 muestra los ensayos de PCR en tiempo real HM de plasma de PCa y clase I negativa (individuos sanos). La figura 5 muestra los ensayos de PCR en tiempo real HM de plasma de PCa y clase II negativa (individuos sanos más biopsias negativas).

40 Como se ilustra en la Tabla 7, en todas las comparaciones negativas de clase y para todos los marcadores, la orina era el analito más sensible.

#### Correlación de los marcadores con la puntuación de Gleason.

45 Las cantidades crecientes de ADN marcador metilado se correlacionaron con el aumento de la puntuación de Gleason para todos los marcadores en plasma. Esto fue cierto para muestras con grandes cantidades de ADN marcador metilado en la orina (ver especialmente los marcadores TFAP2E y RASSF2A), pero en general la correlación fue menos fuerte en el ADN de la orina que en el ADN del plasma. El PSA como marcador de PCa en pacientes con PSA elevado (>4 ng/ml) también se correlacionó con el aumento de la puntuación de Gleason.

El rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase I negativa (varones sanos) en orina se proporciona en la Tabla 8.

50 El rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase negativa II (varones sanos más biopsia negativa) en orina se proporciona en la Tabla 9.

El rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase I negativa (varones sanos) en

plasma se proporciona en la Tabla 10.

El rendimiento de los paneles marcadores de detección para distinguir PCa de la clase II negativa (varones sanos más biopsia negativa) en el plasma se proporciona en la Tabla 11.

Rendimiento del marcador en la aplicación de diagnóstico: Seguimiento al PSA

5 Como marcador de seguimiento de la prueba de PSA, los marcadores preferiblemente identificarían PCa en hombres mayores de 50 años que han sido clasificados como individuos de alto riesgo debido a un PSA elevado (>2.5 ng/ml). Esta es una aplicación y un análisis distintos y requiere una mayor discriminación en comparación con la prueba de detección. Los falsos positivos en esta aplicación surgen de la biopsia negativa clase BPH de PSA elevada. Una vez más, las muestras PCa representan la clase positiva. Para los fines de una aplicación de  
10 seguimiento de diagnóstico, analizamos los datos utilizando una sola clase negativa compuesta por las 51 muestras negativas de biopsia. La AUC de los marcadores probados por PCR en tiempo real en orina postmasaje prostático y plasma de pacientes con cáncer de próstata y pacientes con biopsia negativa se proporciona en la Tabla 12.

15 A partir de dicha tabla puede observarse que para todos los marcadores de metilación analizados, la orina era el analito más sensible. Para el PSA total (tratado aquí como un marcador adicional para determinar si hay más información proporcionada después de la indicación >4 ng/ml para la biopsia), no hubo diferencias en la sensibilidad entre la orina y el plasma.

Rendimiento de los paneles marcadores

Con el fin de proporcionar una precisión mejorada, las combinaciones de marcadores se evaluaron tanto cualitativa como cuantitativamente.

20 La Tabla 13 proporciona el rendimiento de los paneles marcadores de diagnóstico para distinguir la PCa de la biopsia negativa en la orina.

La Tabla 14 proporciona el rendimiento de los paneles marcadores de diagnóstico para distinguir PCa de biopsia negativa en plasma.

Discusión

25 El estudio se realizó en muestras de plasma y/u orina recogidas de 91 pacientes con PCa, 51 pacientes con biopsia negativa (diagnosticada con HBP) y 50 jóvenes varones sanos. Se utilizaron ensayos de PCR en tiempo real HM<sup>TM</sup> para medir la metilación del ADN de los marcadores candidatos. La cantidad de ADN marcador metilado se correlacionó con PCa tanto en plasma como en orina, mostrando el ADN de orina mayor sensibilidad. Como prueba de detección (discriminación del cáncer de PCa de controles sanos usando analito de orina), los candidatos  
30 marcadores de anclaje GSTPi, RASSF2A, HIST1H4K y TFAP2E tienen una sensibilidad del 63%, 74%, 69% y 47% con una especificidad del 96%, respectivamente. Como marcador de seguimiento de la prueba de PSA (discriminación de PCa de controles negativos de biopsia, todos con PSA elevado), los marcadores tienen una sensibilidad del 23%, 18%, 28% y 23% con especificidad del 95%, respectivamente. Un panel de selección cuantitativo de marcadores RASSF2A y HIST1H4K produjo un 94% de sensibilidad con una especificidad del 88%  
35 frente a individuos sanos. Un panel de diagnóstico cuantitativo de los marcadores GSTPi y PSA arrojó un 83% de sensibilidad con un 45% de especificidad. El desempeño de estos marcadores se compara bien con el desempeño de PSA (18% de sensibilidad a 98% de especificidad para hombres <60 años y 19% de sensibilidad a 94% de especificidad para hombres >60 años) en la población de cribado (Punglia et al.). La metilación de todos los marcadores se correlacionó bien con la puntuación de Gleason en el ADN plasmático, pero la correlación fue menos  
40 fuerte en el ADN de la orina.

Conclusiones

45 Al completarse la presente investigación, se demostró que los biomarcadores de cáncer de próstata basados en ADN metilado pueden medirse en plasma y orina, mostrando el ADN de orina una mayor sensibilidad que el plasma. Además, se identificaron los marcadores de metilación del ADN que discriminan a los pacientes con PCa de controles sanos y aquellos con hiperplasia benigna de próstata (HBP). Las principales conclusiones del presente estudio son las siguientes:

- Los marcadores metilados de cáncer de próstata se pueden medir tanto en plasma como en orina de pacientes con PCa.
- Identificación de marcadores que discriminan a pacientes con PCa de aquellos sin PCa.

Tabla 1: Genes y secuencias de acuerdo con la presente invención

Gen	Genómica SEQ ID NO:	Cadena en sentido metilada convertida con bisulfito	Cadena antisentido metilada convertida con bisulfito	Cadena en sentido no metilada convertida con bisulfito	Cadena antisentido metilada convertida con bisulfito
RASSF2A	1	5	6	13	14
TFAP2E	2	7	8	15	16
HIST1H4K	3	9	10	17	18
GSTPi 4 11	4	11	12	19	20

Tabla 2: Componentes de ensayo según el ejemplo 1

Gen	Cebador avance		Bloqueador	Oligo de detección
GSTPi 21	21	22	23	24
HIST1H4K	25	26	27	28
RASSF2A	29	30	31	32
TFAP2E	33	34	35	36

Tabla 3: Análisis de rendimiento de los marcadores (normal más BPH y PCa) en la prueba de tejido de acuerdo con el Ejemplo 1.

Marcador	AUC	Sensibilidad	Especificidad
GSTPi	0.90	0.83	0.91
HIST1H4K	0.91	0.83	0.91
RASSF2A	0.93	0.75	0.91
TFAP2E	NA	NA	NA

Tabla 4: Muestras remotas según el Ejemplo 1.

Tipo de muestra	No. de muestras
Cáncer de Próstata - Puntuación Gleason	
4	1
5	5
6	33

7	39
8	8
9	4
No hay puntuación disponible	1
Cáncer de Próstata Total:	91
Biopsia negativa	51
Control sano	50

Tabla 5: Sensibilidad y especificidad de los marcadores individuales probados por PCR en tiempo real en orina postmasaje prostático de pacientes con cáncer de próstata, pacientes con biopsia negativa y sujetos de control sanos.

Marcador	Negativo Clase I: Sano			Negativo Clase II: Sano + Biopsia (-)		
	AUC	Sens/Espec	Valor de Wilcoxon p	AUC	Sens/Espec	Valor de Wilcoxon p
GSTPi	0.89	0.63/0.96	0	0.79	0.31/0.96	0
RASSF2 A	0.90	0.74/0.96	0	0.79	0.24/0.96	0
HIST1H4 K	0.91	0.69/0.96	0	0.77	0.36/0.96	0
TFAP2E	0.86	0.47/0.96	0	0.77	0.27/0.96	0

Tabla 6: Sensibilidad y especificidad de los marcadores individuales probados por PCR en tiempo real en plasma de pacientes con cáncer de próstata, pacientes con biopsia negativa y sujetos de control sanos.

Marcador	Negativo Clase I: Sano			Negativo Clase II: Sano + Biopsia (-)		
	AUC	Sens/Espec	Valor de Wilcoxon p	AUC	Sens/Espec	Valor de Wilcoxon p
GSTPi/GSTP1	0.61	0.17/0.96	0.0063	0.58	0.17/0.95	0.0183
RASSF2 A	0.68	0.37/1.00	0	0.64	0.20/0.95	0
HIST1H4 K	0.64	0.26/0.96	5e <sup>-04</sup>	0.56	0.16/0.95	0.0572
TFAP2E	0.61	0.22/1.00	4e <sup>-04</sup>	0.56	0.09/0.95	0.0128

Tabla 7

Marcador	Negativo Clase I: Sano		Negativo Clase II: Sano + Biopsia (-)	
	AUC de la orina	AUC del plasma	AUC de la orina	AUC del plasma

ES 2 615 354 T3

GSTPi/ GSTP1	0.89	0.61	0.69	0.55
RASSF2 A	0.90	0.68	0.66	0.60
HIST1H4 K	0.91	0.64	0.64	0.50
TFAP2E	0.86	0.61	0.65	0.52

Tabla 8: Rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase I negativa (varones sanos) en la orina

Panel de marcadores	% Sens PCa	% Espec. Saludable
Marcadores únicos cuantitativos:		
RASSF2A	74	96
HIST1H4K	69	96
GSTPi	63	96
TFAP2E	46	100
Paneles cualitativos:		
GSTPi+HIST1H4K	79	98
RASSF2A+HIST1H4K	94	88
Paneles cualitativos:		
RASSF2A+HIST1H4K	94	88
quadSVM (todos los marcadores, sin PSA)	79	98

Tabla 9

Panel de marcadores	% Sens PCa	% Espec. Saludable + Biopsia (-)
Marcadores únicos cuantitativos:		
RASSF2A	74	76
HIST1H4K	69	68
GSTPi	63	80
TFAP2E	46	88
Paneles cualitativos:		
GSTPi+HIST1H4K	79	72
RASSF2A+HIST1H4K	94	54

ES 2 615 354 T3

Paneles cualitativos:		
RASSF2A+HIST1H4K	94	58
quadSVM (todos los marcadores, sin PSA)	79	76

Tabla 10: Rendimiento de los paneles marcadores de cribado para distinguir PCa de la clase I negativa (varones sanos) en plasma

Panel de marcadores	% Sens PCa	% Espec. Saludable
Marcadores únicos cuantitativos:		
RASSF2A	37	100
HIST1H4K	26	96
GSTPi	17	94
TFAP2E	22	100
Paneles cualitativos:		
RASSF2A+HIST1H4K	41	98
Paneles cualitativos:		
RASSF2A+TFAP2E (TFAP2E utilizado para normalizar)	32	100
quadSVM (todos los marcadores, sin PSA)	39	96

Tabla 11

Panel de marcadores	% Sens PCa	% Espec. Sano + Biopsia (-)
Marcadores únicos cuantitativos:		
RASSF2A	37	91
HIST1H4K	26	88
GSTPi	17	95
TFAP2E	22	92
Paneles cualitativos:		
RASSF2A+HIST1H4K	41	88
Paneles cualitativos:		
RASSF2A+TFAP2E (TFAP2E utilizado para normalizar)	32	92
quadSVM (todos los marcadores, sin PSA)	39	94

Tabla 12

Marcador	PCa vs. Biopsia (-)	
	AUC de la orina	AUC del plasma
GSTPi/ GSTP1	0.69	0.55
RASSF2 A	0.66	0.60
HIST1H4 K	0.64	0.50
TFAP2E	0.65	0.52
*** PSA	0.56	0.56
*** Prueba si el PSA contiene información adicional más allá de lo que fue aportado por la indicación de corte de >4 ng/ml para la biopsia de próstata.		

Tabla 13

Panel de marcadores	% Sens PCa	% Espec. Biopsia (-)
Marcadores únicos cuantitativos:		
RASSF2A	74	55
HIST1H4K	69	41
GSTPi	63	64
TFAP2E	46	77
Paneles cualitativos:		
GSTPi+HIST1H4K	79	46
RASSF2A+HIST1H4K	94	21
Paneles cualitativos:		
RASSF2A+HIST1H4K	94	27
GSTPi+PSA	83	45
quadSVM (todos los marcadores, sin PSA)	79	55

Tabla 14

Panel de marcadores	% Sens PCa	% Espec. Biopsia (-)
Marcadores únicos cuantitativos:		
RASSF2A	37	82

## ES 2 615 354 T3

HIST1H4K	26	79
GSTPi	17	96
TFAP2E	22	84
Paneles cualitativos:		
RASSF2A+HIST1H4K	41	79
Paneles cualitativos:		
RASSF2A+TFAP2E (TFAP2E utilizado para normalizar)	32	85
RASSF2A+TFAP2E+PSA (TFAP2E utilizado para normalizar)	94	22
quadSVM (todos los marcadores, sin PSA)	39	91
quadSVM (todos los marcadores+PSA)	48	87

Listado de secuencias

<110> Epigenomics AG

5 <120> Métodos y ácidos nucleicos para el análisis de la expresión génica asociados con el desarrollo de trastornos proliferativos de la próstata

<130> E30990PCT

<160> 36

<210> 1

<211> 1920

10 <212> ADN

<213> Homo Sapiens

<400> 1

ES 2 615 354 T3

ccaggctgcc	gtagacacag	cctttgctct	cccgaaaaac	acgttctagg	cgccgggatt	60
ccagatacct	gggaaataga	gtgcacgcag	ctggtgagag	gcctcgcgct	tggcttctcc	120
tatcactgag	gcgcagaggt	gctgtggaca	gccagacccc	acacggcgcc	cgaggtgaaa	180
cagaaccctc	agtctcccta	tgaggccact	ggcactctcg	gctgtcccca	gagctctccg	240
acttagagct	gaatgcaaag	taagcgctcg	aaatgcagaa	gtagccgggg	ccgcccacgg	300
cacctgcctc	gctcggggcg	agagaagacg	ccaggctgag	gtcccagcga	cctcaggcac	360
cagctccgaa	ggagggcggg	gagaccgcaa	aggggaagtg	cccggagggc	caacggcccc	420
cgcgcaccct	gcgccccctc	gaagcgcgcc	gcctccccgc	gccggggact	gggacctgcc	480
tctggggaat	ccgcctagaa	gacggcgggc	gactggggtc	gggcactctc	cagggtctgc	540
aggccctccc	cagccctgca	cctgccgcgc	cgccccacct	cgccaggaag	tctcagagac	600
cccggggatg	gggtgggagc	gccttcccat	cgcgggctca	aaaagaagga	aggacgcccc	660
caggggtcgt	agaaggagga	ctagctccaa	gccacaactt	tcttcggacc	caaggcaggc	720
cggctggggc	tccgcgccta	cacggccccct	ggcgggggtc	cgcgcgcccc	gggagccccg	780
cggctcgggg	aggaaagagg	agacaagaga	caggcgagga	ttacggggct	gaccacagccg	840
gggtagggac	catcgtggaa	aaactttggc	gaggtggggg	gacgcggaaa	gagagcggcc	900
cgcgccttgc	accttgcgcc	gggcatcccc	cgccagtgcc	tcgctcccag	tgccccgcgc	960
cccgcgcccc	gcgccttgcc	ttcacccccg	gccagctgca	tcgcgccccg	gccgcaggaa	1020
ccgtggagtt	ggaaagtggg	ggcgcccgcg	ctggggggct	gcttcagctg	cgcctcggcc	1080
agcgatcggc	gggcccgggt	caaatccagc	caggctgggc	aggcgggtggc	cgcgcgactg	1140
gggaccgggc	gccccgcctc	cctcgcctcc	ctcctccttc	ctctccctcc	ctccagcccc	1200
ttggcctttt	tcagccccta	ccgatcttgc	tcgtccgctg	tctctctttt	tctctcgtct	1260
ttcatatcac	tctccacccc	ttcgcttgc	cttcgccttt	cttctctccc	ttgtctcctg	1320
ccccctctct	ttctccccct	ccctctaggg	gcgagcttct	tcccctccct	cccagacaat	1380
gctgtggctg	cgtccccctc	cccgccagct	cgtccaggct	cccgcgcgca	gcgattcttc	1440
cgggctgggg	gtggggaggt	ggggggggag	tgcagggttg	gggaggatga	gctggctccc	1500
ctcacctcct	tgctgctgcc	ctctccaaga	gggatggaga	cttggcccaa	gctcctcggg	1560
tcaccgagag	ctgtgacagc	cactcccagg	gaacagtcac	gctgccctac	caagcccacc	1620
tccagcggcc	tggattcccc	aggcagaggt	tgtgggattt	tgttttttct	aacatcccag	1680
cttattccca	aaagggtttg	agccggacag	gggctaaaca	ggccccttcg	acttggcggg	1740
ccggccagac	gtgacagcaa	tgccaaggag	gccaaagtct	tttgtccatt	tctcacctcc	1800
cccttttcca	tccctggacc	tcctggcgcc	cccagtacac	agaggccctt	gagcagcccc	1860
gctgcaggtt	ccctatctac	tcagagtctt	ccccctcacg	tgcttatccc	caaccctgca	1920

<210> 2

<211> 6096

<212> ADN

5 <213> Homo Sapiens

<400> 2

taccagtgta	agattcaaaa	ttcccttttt	gcactgcaca	gtgagatgcc	cagggtccca	60
gctcagtgcc	tggacatagc	gattcctggg	cctgcccgtc	gccgccccaa	gcgaagctgg	120
tgcgccttgg	gcgagagcaga	cagagaccct	gggtggcagg	ggcttgggaa	gacatggggc	180
gctagggcct	tatgcgcctt	caccgctgcc	ctctgctatt	tgcaggcaat	ggacgagccc	240

ES 2 615 354 T3

ggaatgagcc	tcctagacca	gtccgtgac	aagaaaggta	aggaatggtc	tgtcagggca	300
gagcccggcg	agatggtgca	ggcccttgg	gcacagatcc	atcttcttca	ccggccgtgc	360
ctcctgtgtg	tcgccaggct	gggtgtccac	caggcactct	tcctggccca	gccagatggt	420
aggcagacgt	gcgggcttgg	tgagtttgc	cagcacctg	tggcctggg	tgggcctcag	480
cggatcagca	ttcactgggc	tgacgactg	ggagcctggc	ctctccccgc	cgagggggag	540
ggcactcttg	tggatctgga	gttgatttgc	agaacgagtt	aaaccacttc	cctgtttccc	600
taagagatgg	gaatggaagt	gctgttcca	cggagtggg	gaaatgattt	tcactttaca	660
gtgccttagc	atctcgggtg	ctggcgggca	cttctcttct	cttcttcca	ggcagggcct	720
tggagggcctc	tgggggaatt	ttcttctgt	gggagtctct	tgcggcattt	agacttaggg	780
gagcttgtgt	tgagtgactg	tgtgttaggc	tgtgtgcacc	tgagtcaggg	cccacctgct	840
cctgggtgtc	tgtgtccatg	tgagttcagg	gtcctgtgca	tgtctgaaat	gttcccttca	900
tgggtgtctt	agtatttctt	ggagtgtgag	tgtgtctgtt	tctgtgaatg	tgtttgtgag	960
gtgtgtctct	gtatgttgg	gtgcatttct	ctgcatttgg	gggatgtaca	catttctcaa	1020
tatgtacagt	atctctgttg	tgtcctgcac	tttgttcttt	ggtatctgag	gatttccaag	1080
catgcgcggg	ccctctctgt	gtatatatag	gagtatattat	gtgactcctg	gcattagtaa	1140
aatccagggga	cacgggatcc	accttttctg	gcctgaggac	caagtactgg	ccatgacagg	1200
ggaaggtgag	agacgacaaa	aacagagaga	cagccagaga	ggagcagaga	gtcagagggg	1260
cccagcattg	gggtagcagc	ctctttacat	ttggggcagg	tgcccgaag	aattcagagg	1320
tgacacatgag	cctgaggtgc	cccaggcagg	cactgtccc	acagggtttg	gcttgagttg	1380
tttttcaaac	gagtgaattc	aagcctgggc	tctatgttgc	ctccacttgt	tctcagggga	1440
ggccaaggtg	gaagtgggtg	tagcagggct	ggggctggac	ttccaggagc	tggggctgag	1500
ttaccaggag	ctgggggttg	ggtggatgac	ttggagtgtg	tagcagggaa	gatgaggcaa	1560
cagggcagga	agtgggtggg	gggaggtgga	attggggctg	tgtcctgtgt	cgcttggaac	1620
tgggagtgtg	gaaagacac	taggaacctg	ggtgcagcgc	agctctgctg	gtggggcttg	1680
gttggcttac	tgtacagagc	ctttcttgac	ccctgaagaa	agagatccgt	ctgcagtggg	1740
caaaagcctg	cctggacttc	ctggccacca	gaaatatgag	catggtggtg	gtccccagtt	1800
ccctattcat	gcttgggctc	aagagactgg	gagtctaggt	tcactgactc	cctgagaaa	1860
actaagacc	tgcatthtag	aaagagttt	gggatctct	gcctgctgca	agggtagaag	1920
gatcagctgt	tcctctgagc	accttaacc	ggaaccccgg	tccgaagccg	agacaggaga	1980
ctggatgcga	ggcctcccca	gagctggttt	ctctcaaaaca	acttccaaaa	ctcctagatc	2040
ctaggggtac	gccgaaatcc	cccaaagcag	tccaaagaac	acaacgagag	tcctaaccatc	2100
ccaggtggcg	gcgcgctggc	tccttgagc	ggggcgggac	gcggccgcgc	ggactcacgt	2160
gcacaaccgc	gcgggacggg	gccacgcgga	ctcacgtgca	caaccgcggg	accccagcgc	2220
cagcgggacc	ccagcgcag	cgggaccca	gcgccagcgc	gaccccagcg	ccagcgggac	2280
cccagcgcga	gcgggacccc	agcgcagcgc	ggaccccagc	gccagcgggt	ctgtggccca	2340
gtggagcag	tggagcgtg	gcgacctgag	cggagactgc	gccctggagc	ccccagccta	2400
gacgtcaagt	tacagcccgc	gcagcagcag	caaaggggaa	ggggcaggag	ccgggcacag	2460
ttggatccgg	aggtcgtgac	ccaggggaaa	gcgtggcgg	tcgacccagg	gcagctgcgg	2520
cggcgaggca	gttgggctcc	ttgctcctg	gagccgccc	tcccacacc	tgccctcggc	2580
gccccagca	gttttcaact	tggccctcgc	cggctactgc	gggattcggc	gttgccgcca	2640
gcccagtg	gagtgaatta	gcgcccctct	tcgtcctcgg	cccttccgac	ggcacgagga	2700
actcctgtcc	tgccccacag	accttcggcc	tcgcccaggt	gcgggtactgg	agcctgcccc	2760
gccagggccc	tggaaatcaga	gaaagtcgct	ctttggccac	ctgaagcgtc	ggatccctac	2820
agtgcctccc	agcctgggcg	ggagcggcgc	ctgcgtcgtc	gaaggttggg	gtccttgggtg	2880
cgaaagggag	gcagctgcag	cctcagcccc	accccagaag	cggccttcgc	atcgtgcgg	2940
tgggcgttct	cgggcttcga	cttcgccagc	gccgcggggc	agaggcacct	ggagctcgca	3000
gggcccagac	ctgggttggg	aaagcttcgc	tgactgcagg	caagcgtccg	ggagggggcg	3060
ccagggcaag	ccccggcgtc	ttaccacaca	cttccgggtc	ccatgccagt	tgcatccgcg	3120
gtattgggca	ggaaatggca	gggctgaggc	cgaccctagg	agtataagg	agccctccat	3180
ttcctgcccc	catttgtcac	ctccagtttt	gcaacctatc	ccagacacac	agaaagcaag	3240
caggactggt	ggggagacgg	agcttaacag	gaatattttc	cagcagtgag	caggggctgt	3300
atgggacgcg	ggaggagctc	agaggaggcg	cggagagtgc	ccgaggttgg	gtgagtgcct	3360
agaggggaga	tagtgaacc	gggttcaaga	ggtgcttagt	gggtgtttgt	tgaatgaatg	3420
agtgatgggc	tttgaagtct	gagtgcattg	aaagaggggg	tgtgtaaaaa	gggctccttt	3480
catcacacag	gacacagcat	atgcaaatcc	tctccctgtg	gaaaagccag	acaggttaaa	3540
aaggttacia	acaaattagc	cgggcatggt	ggtgcgcgtc	tgtagtccca	gctactaggg	3600
aggctgagcc	aggggaatcg	cttgaacccg	ggagccggag	attgcagtga	gccaagatcg	3660
cgccactgca	ctccagcctg	gaaacagagc	gagactccgt	ctcggaaaaa	aaaaaaaaaa	3720
gttacaacacc	gtgtgtgggt	ttcaggttat	acaatcagag	ctggagggga	gtgggtcaagg	3780
atgagaactg	agatggatcc	ctcgttccct	ctggaggaga	gtgggtgggt	gcctacttgg	3840
gggtggggaa	tcctctcca	cgggctcagc	tgtccaatct	caggggatct	ctaggacagg	3900
agctgatgta	aacagtcgcc	ctattccttg	ctgtcttttg	ccctggagaa	ggaggagggg	3960
gctggggagg	gtctccactt	cccagacaat	ctctaagcag	ccaggacatg	ggtgagatga	4020

ES 2 615 354 T3

gtgagatact	gacttctggg	acagaatttg	agagggtgcc	aaaaaactca	gtaatcaaga	4080
taaataggcc	gggcgcagtg	gctcacgtct	gtaatcccag	cactttggga	ggccggatca	4140
cttgagggtca	agagttcgag	accagcctgg	ccaagatggg	gaaaccccat	ctctactaaa	4200
aatacaaaaa	ttagcccagt	gtggtggcgc	tagcctgtaa	tcccagccac	tcaggaggct	4260
gaggcaagag	aattgcttga	cccaggaggc	agaggttgca	gtgagccgag	atcatgccac	4320
tgtactccag	cctggacaac	agagggagac	tatctcaaaa	aaaaaaaaaa	aaaaaaaaaa	4380
aaaaaagagg	cggggcggcg	gtggctcaca	ccatgtgatc	ccagcacttt	gggaggccga	4440
ggcgggtgga	tcacctgagg	tctggagttc	gagaccagcc	tggccaatat	ggtgaaacct	4500
cgtttctact	aaaaatacaa	aaattagctg	ggtgggggtg	caggcacctg	taatcccagc	4560
tactccggag	gctgaggcag	gagaatccct	tgaacctgcg	gggaggagg	tgcagtgaac	4620
caagatcaca	ccattgcact	ccagcctgga	caacaacagc	aaaactctgt	ctcaaaaaaa	4680
aaaaaatct	tttttttcga	gacacagttt	tactctctcg	cccaggttgg	ggtgcagcac	4740
cacgatctca	gctcactgca	acctctgctt	ctcagattct	cgtacctcag	cctcccaagt	4800
agctgggatt	acaggtacct	gtcaccacgc	ccagctaatt	tttgtatttt	tagtaggggc	4860
gtggtttcac	catgttggcc	aggctggctt	tgaactcctg	acctcaagt	acctgcccgc	4920
ttcagccacc	caaagtgctg	ggattacagg	cgtgagccac	cacgcttggc	cttttttaat	4980
gaaaatagtg	caaaaatcca	cgataaacia	aatatcaaaa	atttactgaa	cttgcacttc	5040
cacaaccctt	tctcacctgc	ctcccaggct	actctctgcc	ccagaaagca	acttaaaaaa	5100
tgtgcagatg	gagtttggac	tttacctgaa	aatgggtgga	gctatggaaa	accttgggagc	5160
aggggagtg	aggatagaaa	ttatatgtaa	aagaaaccct	gggccggggc	cagtggctta	5220
tgcctgtaat	cccagcactt	tgggaggccg	aggcaggtgg	attacctgag	gtcaggagat	5280
tgagaccagc	ctgaccaaca	tggtgaaatg	tcctctctac	taaaaataca	aaaaaaatta	5340
gccaggcatg	gtggtgcacg	cctgtagtcc	cagctactcc	ggaggctgag	acaggaaaat	5400
cgcttgaacc	cgggaggcgg	aggttgcagt	gagccaagat	tgtgccattg	cactccagcc	5460
tgggcaacia	gagcaaaact	ccatcttaaa	aaaaaagaaa	gaaagaaacc	ctctggcagt	5520
tgatgagaag	gaaacttaat	cggcaggtcc	cagcagggga	gatgaggaga	ctctagggag	5580
ggcatttgca	catgctgtgc	cccagtggtg	gccagggagc	aggctactac	tcctcccgtc	5640
taccttcctc	ttgctccaac	cccttcaagc	tttggaccag	tggtagccta	agtgtagtcc	5700
aaggaaccac	atgcatcagg	acccccaggg	ggtgcttgtt	aaaaatgcaa	atthttggcca	5760
ggtgcagtgg	ctcacacctg	taatcccagc	actttggggag	gccgaggcgg	gtggatcacg	5820
aggtcaggag	atcgagacca	tcctggcaaa	cacggtgaaa	ccccatctct	actaaaaaaa	5880
caaaaacaaa	caaaaaaaaa	cattagctgg	gcgtggtggc	gggcgcctgt	agtcccagct	5940
actcgggagg	ctgaggcagg	agaatggcgt	gaacccggga	ggcggagctt	gcagtgagct	6000
gagattgcgc	cactgcaccc	cagcctgggc	gacagagcga	gactctgcct	caaaaaaaaa	6060
aaagcaaatt	tcttgggcac	cacccccat	tgactg			6096

<210> 3

<211> 2501

<212> ADN

5 <213> Homo Sapiens

<400> 3

ES 2 615 354 T3

tttgcaaatg	gagacatctt	cattattcct	atagtatcat	atgtttttaa	agtttgtact	60
cacactttgg	gtgataaatg	aaggacaaga	tccttcccta	tccttgtgag	gatgactaca	120
gcatgactgg	atgggcttgc	tatgattttt	atctttccct	gtgttctcac	taccgtttta	180
ttaatctcag	ttctttttca	cagggtagca	cagaatttaa	ctagcagaaa	gagatccagc	240
catgtagacc	agagatttgt	ctaagtgacg	gcatgtaaga	atcaggaagg	aaagtttttt	300
gtttaaatac	caacaggttc	cttccttaaa	gcaattatta	tttttcaa	ctaaccaca	360
aggtgatagt	atccttaaac	caattaaatc	agaatctcgg	gttggataac	ctcaaatatg	420
acttattage	acttcccatt	aatcactggg	ccttcaggcc	tttaagttta	cttactagga	480
atctcacttt	taataccatc	ttatcaactt	cagttgtaaa	taagagaaca	ctcaaaggct	540
gaggaattct	cagcggtaaa	gctctgcca	cgtttaagtaa	caaaggataa	gttagtcttt	600
gttgtgatca	ctttgttgta	ctgataagct	acgtatttct	actcaaggat	tcaaattctc	660
acctttctca	agaattgggc	caaaaccgat	aaactaaact	tatttacggg	ccactgatta	720
aaggttgttg	cataataagt	tcttgctatg	ttcagcagtt	ggattcacag	cgccagaaac	780
ctataactgc	ttgactttcc	tccccactac	actgcgaaaa	ttgcccctta	aatgtaacta	840
accctaaaac	ctcaacagta	tcgtggccag	gcgtgggtgg	tcactactgt	aataccaaca	900
ttaggcatag	gcgaggggat	tgaggccagg	atatcgaaac	tagcctggga	aacacacgga	960
gaccgggtct	ttggaaaaat	aattagcctt	gcgtgggtgg	gggcgcgagg	ttccggctaa	1020
tcgggaggct	acagtgagcc	atgatgacac	tgcaactacag	tctgcgcgac	ggcccagctc	1080
agtaagctct	ggagcacctg	aaacaagttg	tgttgggtat	tttatttact	ggagagcgat	1140
tagtgactga	tgccacttta	cagcgcactag	agacgcacgc	tccgatagca	gcacaaaactc	1200
agcaggcgcg	aacaaatggt	aaagagaaac	tgggcaaaca	agcatcacgg	ctcctcagct	1260
gagaaagtgg	gggccctaaa	aagggccttt	tgttgataga	aagggacgct	caaccaccga	1320
aaccgtagag	ggtgcggccc	tggcgcttga	gcgcgtagac	cacatccatg	gcgggtgaccg	1380
tcttgcgctt	ggcgtgctct	gtataggtca	cggcgtcccg	gatcacgttc	tccaggaaca	1440
ccttcagcac	cccgcgagtc	tcctcgtaga	tgaggccgga	gatgcgcttc	acgccgccgc	1500
ggcgagcaag	gocccggatg	gocggccttg	tgatgcctcg	gatattgtcg	cgcagtaact	1560
tacggtggcg	cttagcgcgg	cctttgccaa	gacccttccc	gcctttgccc	cgccagaca	1620
tgacgagcaa	gaggagtctc	acccaacgct	ttgtgaggac	tctggcctga	ggcagcgctt	1680
ttatacgaca	gttggcggac	cgaactgaga	acctgaaaga	agtcggcggg	aagtcccgcc	1740
ccggtggggg	aggggaaatc	taaagggcca	aaccgaaata	gggggaaaaa	aaaagcgagc	1800
ttcttgtttc	cgtgttctga	atthttgtaac	gtgcatagta	ttttgttacc	acgttatgag	1860
gctttaaana	attgcttttg	aacgcagaag	atatacatca	atactgtggg	aaatacaaga	1920
aaggacaaga	aattaagaaa	ctacaatggt	atcccatcac	acaggctagt	taatcatgta	1980
ttttgcagag	cagttgcaca	tatttttcca	agaaaatgta	tacagtgttg	tatatggagt	2040
tttgtaacct	ccttatattg	attataatth	aaccaatthc	tattaaagag	ataaaagtga	2100
tgttttgggtg	tctatgtttc	ttaggaatta	tcaatagtta	taatcagttc	cccagcaatt	2160
ttttaatcgg	ctgtatttta	aaaataatgt	ttccacatt	caacataaat	gtactttttc	2220
tctataacttg	ggaccaatat	tgaatthtat	gattthtatta	caccaaatt	taaatthtat	2280
tacattaata	tttaaaattg	tattagaggt	ctcatgattt	ggtactacgg	gtctccgcat	2340
tatttccttt	ccaaatthcc	taatctgttt	caccaaggtt	tctggacaac	tttagagacc	2400
ttttgtgaag	tttgaataaa	atctcttcca	gattthtgata	attgcattag	ctthtaggact	2460
taattggaat	agaatthaaa	tcctthaaaac	aagctthtat	a		2501

<210> 4

<211> 2501

5 <212> ADN

<213> Homo Sapiens

<400> 4

ES 2 615 354 T3

ttgttgtaca	gaatatttca	tcacccaggt	attatgccga	gtacccaata	gttctctttt	60
ctgctcctct	ccttctctcc	atcctgcacc	ctggagtcaa	ccacagtgtc	tgttgtttcc	120
ttgtttgtgt	tataagttct	catcatttag	ctcccactta	caagtgagaa	catccagtat	180
ttggatttct	gttcctgcat	tagtttgcta	aggataatag	cctctagctc	catccatggt	240
cccacaaaag	acatgatcta	gttcttttta	atggctgcat	taaataaggt	tttaaagata	300
caacataaac	accaacctct	tccccaccac	aaaaatccct	tgctgaattt	gattacactt	360
aaattaacga	gttttgtttc	atgaaagact	ccttgacaaa	acttgacagt	tgatggaata	420
ggagaagctg	tctgtcatgt	ctaaagccaa	caagagatca	atatctagaa	taaataggaga	480
tctgcaaate	aacagaaaag	aggcagcaaa	gccaaagaaa	atagcctaag	gcacagccac	540
taaaaggaac	gtgatcatgt	cctttgcagg	gacatgggtg	gagctggaag	ccgttagcct	600
cagcaaacctc	acacaggaac	agaaaaccag	cgagaccgca	tggtctcact	tataagtggg	660
agctgaacaa	tgagaacaca	tggtcacatg	gcggcgatca	acacacactg	gtgcctggtg	720
agcgggggtgc	tggggagggg	gagtaccagg	aagaatagct	aagggatact	gggcttaata	780
cctgggtgat	gggatgatct	gtacagcaaa	ccatcatggc	gcacacacct	atgtaacaaa	840
cctgcacatc	ctctacatgt	accccagaac	ttcaaataaa	agttggacgg	ccaggcgtgg	900
tggctcacgc	ctgtaatccc	agcactttgg	gaagccgagg	cgctgcagatc	acctaaggtc	960
aggagttcga	gaccagcccg	gccaacatgg	tgaaaccccc	tctctactaa	aaatacaaaa	1020
atcagccaga	tgtggcacgc	acctataatt	ccacctactc	gggaggctga	agcagaattg	1080
cttgaacccg	agaggcggag	gttgcagtga	gccgcggaga	tgcgcacct	gcactccagc	1140
ctggggccaca	gcgtgagact	acgtcataaa	ataaaataaa	ataacacaaa	ataaaataaa	1200
ataaaataaa	ataaaataaa	ataaaataaa	ataaaataaa	ataaaaaaat	aaaataaaat	1260
aaaataaaat	aaagcaattt	cctttcctct	aagcggcctc	caccctctc	ccctgcctctg	1320
tgaagcgggt	gtgcaagctc	cgggatcgca	gcggtcttag	ggaatttccc	cccgcgatgt	1380
cccggcgcgc	cagttcgctg	cgcacacttc	gctgcgggtc	tcttctgct	gtctgtttac	1440
tccctagggc	ccgtggggga	cctgggaaag	agggaaaggc	ttccccggcc	agctgcgcgg	1500
cgactccggg	gactccaggg	cgccccctctg	cggccgacgc	ccgggggtgca	gcggccgcgc	1560
gggctggggc	cggcgggagt	ccgcgggacc	ctccagaaga	gcggccggcg	ccgtgactca	1620
gcactggggc	ggagcggggc	gggaccaccc	ttataaggct	cggaggccgc	gaggccttcg	1680
ctggagtttc	gccgcgcag	tcttcgccac	cagtgagtac	gcgcggcccc	cgctccccggg	1740
gatggggctc	agagctccca	gcatggggcc	aaccgcgacg	atcaggcccc	ggctccccggc	1800
agggtcctc	gccacctcg	agaccggga	cgggggccta	ggggaccag	gacgtcccca	1860
gtgccgttag	cggttttcag	ggggcccgga	gcgcctcggg	gagggatggg	accccggggg	1920
cggggagggg	gggcagactg	cgctcaccgc	gccttggcat	cctccccgg	gctccagcaa	1980
acttttcttt	gttcgctgca	gtgccgcct	acaccgtgg	ctatttccca	gttcgaggta	2040
ggagcatgtg	tctggcaggg	aaggaggca	ggggctgggg	ctgcagccca	cagcccctcg	2100
cccaccggga	gagatccgaa	cccccttatc	cctccgtcgt	gtggctttta	ccccgggct	2160
ccttctgtt	ccccgcctct	cccgccatgc	ctgctccccg	cccagtggt	gtgtgaaatc	2220
tctggaggaa	cctgtttccc	tgttccctcc	ctgcactcct	gaccctccc	cggttgctg	2280
cgaggcggag	tcgcccggt	ccccacatct	cgtacttctc	cctccccgca	ggccgctgcg	2340
cggccctgcg	catgctgctg	gcagatcagg	gccagagctg	gaaggaggag	gtggtgaccg	2400
tggagacgtg	gcaggagggc	tactcaaag	cctcctcgt	aagtgacct	gcccgggcaa	2460
ggggaggggg	tgctgggct	tagggggctg	tgactaggat	c		2501

<210> 5

<211> 1920

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

ES 2 615 354 T3

<400> 5

ttaggttgtc	gtagatatag	tttttgttt	ttcgaaaaat	acgttttagg	cgtcgggatt	60
ttagatattt	gggaaataga	gtgtacgtag	ttggtgagag	gtttcgcggt	tggttttttt	120
tattattgag	gcgtagaggt	gttgtggata	gtttagattt	atacggcggt	cgaggtgaaa	180
tagaattttt	agttttttta	tgaggttatt	ggtattttcg	gttgttttta	gagtttttcg	240
atntagagtt	gaatgtaaag	taagcgttcg	aatgtagaa	gtagtcgggg	tcgtttacgg	300
tatttgtttc	gttcggggcg	agagaagacg	ttaggttgag	gttttagcga	ttttaggtat	360
tagtttcgaa	ggagggcggg	gagatcgtaa	aggggaagtg	ttcggagggt	taacggtttt	420
cgcgtatttt	gcggtttttt	gaagcgcgct	gttttttcgc	gtcggggatt	gggatttggt	480
tttggggaat	tcgtttagaa	gacggcggcg	gattggggtc	gggtattttt	tagggttggt	540
aggttttttt	tagttttgta	tttgcgcgct	cgttttattt	cgtaggaag	ttttagagat	600
ttcggggatg	gggtgggagc	gtttttttat	cgcgggttta	aaaagaagga	aggacgtttt	660
taggggtcgt	agaaggagga	ttagttttaa	gttataattt	ttttcggatt	taaggtaggt	720
cggttggggg	ttcgcgttta	tacggttttt	ggcggggggt	cgcgcgtttc	gggagtttcg	780
cggttcgggg	aggaaagagg	agataagaga	taggcgagga	ttacgggggt	gatttagtcg	840
gggtagggat	tatcgtggaa	aaattttggc	gaggtggggg	gacgcggaaa	gagagcgggt	900
cgcgttttgt	attttgcgct	gggtatttcg	cgttagtgtt	tcgtttttag	tgtttcgcgt	960
ttcgcgtttc	gcggtttggt	tttatttcgg	gttagttgta	tcgcgttcgc	gtcgtaggaa	1020
tcgtggagtt	ggaaagtggg	ggcgtcgcgg	ttgggggggt	gttttagttg	cgtttcgggt	1080
agcgatcggc	gggtcgggtt	taaatttagt	taggttgggt	aggcggtggt	cgcgcgattg	1140
gggatcgggc	gtttcgtttt	tttcgttttt	tttttttttt	tttttttttt	tttttagttt	1200
ttggtttttt	ttagttttta	tcggatttgt	tcgttcgttg	tttttttttt	tttttcgttt	1260
tttatattat	tttttatttt	ttcgttttgt	tttcgttttt	tttttttttt	ttgttttttg	1320
tttttttttt	tttttttttt	tttttttaggg	gcgaggtttt	tttttttttt	tttagataat	1380
gttgtgggtg	cgtttttttt	ttcgttagtt	cgtttaggtt	ttcgtcgtta	gcgatttttt	1440
cgggttgggg	gtggggaggt	gggggggggag	tgtagggttg	gggaggatga	gttggttttt	1500
tttatttttt	tggtgttggt	ttttttaaga	gggatggaga	tttggtttaa	gtttttcggg	1560
ttattcggag	ttgtgatagt	tatttttagg	gaatagttac	gttgttttat	taagtattat	1620
tttagcgggt	tggatttttt	aggtagaggt	tgtgggattt	tgtttttttt	aatatttttag	1680
tttattttta	aaagggtttg	agtcggatag	gggttaaata	ggttttttcg	atttggcggg	1740
tcggttagac	gtgatagtaa	tgtaaggag	gttaagtttt	tttgtttatt	ttttattttt	1800
ttttttttta	tttttggatt	ttttggcgtt	tttagtata	agaggttttt	gagtagttcg	1860
gtttagggtt	ttttatttat	tttagagttt	tttttttacg	tgtttatttt	taattttgta	1920

<210> 6

<211> 1920

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 6

ES 2 615 354 T3

tgtagggttg	gggataggta	cgtgaggggg	agaatthttga	gtagataggg	aatthttagt	60
cgggttgtht	aagggtthtt	gtgtattggg	ggcgttagga	ggtthaggga	tggaaaaggg	120
ggaggtgaga	aatggataaa	gaaatthtgg	thttthtggt	ttgttggtac	gthttggtcgg	180
ttcgttaagt	cgaaggggtt	tgtthtagtt	ttgttcggtt	taaaththtt	tgggaataag	240
ttgggatggt	agaaaaata	aaatthttata	atthttgtht	ggggaattht	ggtcgttgga	300
ggtgggttht	gtagggtagc	gtgattgtht	thttggagtg	gthgttatag	thttcgggtga	360
atcgaggagt	ttgggttaag	thtttattht	thttggagag	gtagtagta	aggaggtgag	420
gggagttagt	thathththt	taaththtga	thttththtt	atthththtat	thtttagttcg	480
gaagaatcgt	tggcggcggg	agthttggacg	agthttggcggg	gaaggggacg	tagttatagt	540
atthtttggt	agggagggga	gaagthttcgt	thtttagaggg	gaggggagaa	gaggaggggg	600
taggagataa	ggggaggaag	aaaggcgaag	gtaaggcga	ggggtggaga	gtgatatgaa	660
gagcgagaga	aaagagagga	tagcggacga	gtagattcgg	taggggttga	aaaaggttaa	720
ggggttgtag	ggagggagag	gaaggaggag	gggagcgagg	aggcggggc	gthtcggttht	780
tagtcgcgcg	gthtatcgttht	gthttagtht	gthttgattt	agthtcggtt	gthtcgacgtt	840
ggtcagggcg	tagttgaagt	agthththtag	thcggcggtt	ththththth	aatthttacgg	900
thththcggc	gcgggcgcga	tgtagttggt	thcgggtgaa	ggtaaaggcg	ggggcgcggg	960
gcgcggggta	ttgggagcga	ggtattggcg	cgggatgtht	ggcgtaaagg	gtagggcgcg	1020
ggtcgtthth	ththtcgcgtt	thththattt	gththaaagtht	thththacgat	gthththattt	1080
cggthtggtt	agththcgtaa	thththcgttht	ththththth	ththththth	thththcagtcg	1140
cggggtthth	ggggcgcgcg	gathththcgtt	aggggtcgtg	taggcgcgga	gthththtagtcg	1200
gththththt	ggththcgaaga	aagththtggt	thggagthtag	ththththth	acgathththt	1260
ggggcgttht	ththththth	tgagththcgcg	atgggaaggc	gthththattt	tathththcggg	1320
gthththgaga	ththththggcg	aggthggggcg	gcgcggttagg	tgtagggtht	gggagggtht	1380
gatagththt	gagagthgtt	gathththtagt	cgtcgtcgtt	thththaggcgg	atthththtaga	1440
ggtaggttht	agthththcggc	gcggggaggc	ggcgcgttht	agagggggt	agggtgcgcg	1500
ggggtcgtth	gthththcggg	tathththth	thcgggttht	thcgtththth	thcggagtht	1560
gtgththgag	thcgtthggat	thththththt	cgtthththth	cgtthththth	gaggtaggtg	1620
thcgtthggcg	thththcgttat	thththththt	cgcgcgttht	thththththt	agththththt	1680
cggagagtht	tgggatagat	cgcgcgtgtht	agththththt	tagggagatt	gaggtththt	1740
ththththth	ggcgtcgtg	gggtthththt	tgtthththt	atthththth	thththththt	1800
ggagaagtht	agcgcgaggt	thththththt	thththththt	thththththt	aggthththt	1860
aatththth	thththththt	gththththt	agagthththt	thththththt	ggtagththt	1920

<210> 7

<211> 6096

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 7

ES 2 615 354 T3

tattagtgta	agatttaaaa	tttttttttt	gtattgtata	gtgagatgtt	tagggtttta	60
gtttagtgtt	tggatatagc	gatttttggg	tttgttcgtc	gtcgttttaa	gcgaagtgg	120
tgcgttttgg	gcggagtaga	tagagatttt	gggtggtagg	ggtttgggaa	gatatgggcg	180
gttagggttt	tatgcgtttt	tatcgttggt	ttttgttatt	tgtaggtaat	ggacgagtcg	240
ggaatgagtt	ttttagatta	gttcgtgatt	aagaaaggta	aggaatgggt	tgttagggta	300
gagttcggcg	agatgggtga	ggtttttggg	gtatagattt	atttttttta	tcggtcgtgt	360
tttttgtgtg	tcgttagggt	gggtgtttat	taggtatttt	ttttggttta	gttagatggt	420
aggtagacgt	gcgggtttgg	tgagtttggt	tagtattttg	tggtttgggg	tgggttttag	480
cggattagta	tttattgggt	tgtagtattg	ggagtttggg	tttttttcgt	cgagggggag	540
ggtatttttg	tggatttggg	gttgatttgt	agaacgagtt	aaattatttt	tttgtttttt	600
taagagatgg	gaatggaagt	gttgttttta	cggagtgggg	gaaatgattt	ttattttata	660
gtgttttagt	atttcgggtg	ttggcgggta	tttttttttt	ttttttttta	ggtaggggtt	720
tggaggtttt	tgggggaatt	ttttttttgt	gggagttttt	tgcggtattt	agatttaggg	780
gagtttgtgt	gtgagtattg	tgtgttaggt	tgtgtgtatt	tgagttaggg	tttatttgtt	840
tttgggtggt	tgtgtttatg	tgagtttagg	gttttgtgta	tgtttgaat	gtttttttta	900
tgggtgtttt	agtatttttt	ggagtgtgag	tgtgtttggt	tttgtgaatg	tgtttgtgag	960
gtgtgttttt	gtatgttggg	gtgtattttt	ttgtatttgg	gggatgtata	tattttttta	1020
tatgtatagt	atttttgttg	tgttttgtat	tttgtttttt	ggtatttgag	gattttttaag	1080
tatgcgcggg	ttttttttgt	gtatatatag	gagtatttat	gtgatttttg	gtattagtaa	1140
aatttaggga	tacgggattt	attttttttg	gtttgaggat	taagtattgg	ttatgatagg	1200

ES 2 615 354 T3

ggaaggtgag	agacgataaa	aatagagaga	tagttagaga	ggagtagaga	gtagagggg	1260
tttaggtatt	gggtagtagt	ttttttatat	ttgggtagg	tgttcgaag	aatttagagg	1320
tgtatatgag	tttgaggtgt	tttaggtagg	tattgttttt	atagggttt	gtttgagttg	1380
ttttttaaac	gagtgaat	aagtttgggt	tttatttgtt	ttttatttgt	ttttagggga	1440
ggtaaggtg	gaagtgggtg	tagtaggggt	ggggttggat	ttttaggagt	tggggtgag	1500
ttattaggag	ttgggggttg	ggtggatgat	ttggagtgtg	tagtagggaa	gatgaggtaa	1560
tagggtagga	agtgggtggg	gggaggtgga	attggggttg	tgttttgtgt	cgtttggaa	1620
tgggagtgtg	ggaaagatat	taggaatttg	gttgtagcgt	agttttgttg	gtggggttg	1680
gttggtttat	tgtatagagt	ttttttgat	ttttgaagaa	agagattcgt	ttgtagtggg	1740
taaaagtttg	tttgattttt	ttggttatta	gaaatatgag	tatggtggtg	gtttttagtt	1800
ttttatttat	gtttgggttt	aagagattgg	gagtttaggt	ttattgattt	tttgagaaag	1860
attaagattt	tgtatttttag	aaagaggttt	ggggattttt	gttttgcgta	agggtagaag	1920
gattagttgt	ttttttgagt	attttaattc	ggaatttcgg	ttcgaagtcg	agataggaga	1980
ttggatgoga	ggttttttta	gagttgggtt	tttttaata	atttttaaaa	tttttagatt	2040
ttaggggtac	gtcgaat	tttaagtag	tttaagaat	ataacgagag	ttttaatatt	2100
ttaggtggcg	gcgcttgggt	tttttggagc	ggggcgggac	gcggtcgcgc	ggatttacgt	2160
gtataatcgc	gcgggacggg	gttacgcgga	tttactgtga	taatcgcggg	attttagcgt	2220
tagcgggatt	ttagcgttag	cgggatttta	gcgttagcgg	gatttttagc	ttagcgggat	2280
tttagcgtta	gcgggatttt	agcgttagc	ggatttttagc	gttagcgggt	ttgtggttta	2340
gtggagcag	tggagcgttg	gcgatttgag	cggagattgc	gttttggacg	ttttagttta	2400
gacgttaagt	tatagttcgc	gtagtagtag	taaaggggaa	gggtaggag	tcgggtatag	2460
ttggattcgg	aggctcgtgat	ttaggggaaa	gcgtgggcgg	tcgatttagg	gtagttgcgg	2520
cggcgaggt	ggtgggtttt	ttgttttttg	gagtcgtttt	tttttatatt	gttttcggc	2580
gttttttagta	gtttttattt	tggtttttcg	cggttattgc	gggattcggc	gttgcgtta	2640
gttttagtggg	gagtgaatta	gcgttttttt	tcgttttcgg	ttttttcgac	ggtacgagga	2700
atttttgttt	tgttttatag	attttcggtt	ttcgtcagat	gcggtattgg	agtttgtttc	2760
gttagggttt	tgaatttaga	gaaagtcggt	ttttggttat	ttgaagcgtc	ggatttttat	2820
agtgtttttt	agtttgggcg	ggagcggcgg	ttgcgtcgtt	gaaggttggg	gtttttggtg	2880
cgaaggggag	gtagttgtag	tttttagttt	attttagaag	cggttttcgt	atcgttgcgg	2940
tgggctttt	cgggtttcga	tttcgttagc	gtcgcggggt	agaggtattt	ggagttcgta	3000
gggttttagat	ttgggttggg	aaagtttcgt	tgattgtagg	taagcgttcg	ggagggcgg	3060
ttaggcgaag	tttcggcgtt	ttattatata	ttttcggggt	ttatgtagt	tgtattcgcg	3120
gtattgggta	ggaaatggta	gggttgaggt	cgatttttagg	agtataaggg	agttttttat	3180
tttttgttta	tatttgttat	tttttagttt	gtaatttatt	ttagatata	agaaagtaag	3240
taggattggg	ggggagacgg	agtttaatag	gaatattttt	tagtagtgag	taggggttgt	3300
atgggacgcg	ggaggagttt	agagggagcg	cggagaggtg	tcgaggttgg	gtgaggtttt	3360
agaggggaga	tagttgaatc	gggtttaaga	ggtgtttagt	gggtgtttgt	tgaatgaatg	3420
agtgatgggt	tttgaagttt	gagtgatttg	aaagaggggg	tgtgtaaaaa	gggttttttt	3480
tattatatag	gatatagtat	atgtaaattt	tttttttgtg	gaaaagttag	ataggttaa	3540
aaggttataa	ataaatttagt	cgggtatggg	ggtgcgcggt	tgtagtttta	gttattaggg	3600
aggttgagtt	aggggaatcg	tttgaattcg	ggagggcggg	attgtagtga	gttaagatcg	3660
cgttattgta	tttttagttt	gaaatagagc	gagatttcgt	ttcggaaaaa	aaaaaaaaaa	3720
gttataaatc	gtgtgtgggt	tttaggttat	ataattagag	ttggagggga	gtggttaagg	3780
atgagaattg	agatggattt	ttcgtttttt	ttggaggaga	gtgggtggtt	gtttatttgg	3840
gggtggggaa	ttttttttta	cgggtttagt	tgtttaattt	taggggattt	ttagatagg	3900
agttgatgta	taattgcggt	ttattttttg	ttgtttttgg	ttttggagaa	ggaggaggga	3960
gttggggagg	gtttttattt	tttagataat	ttttaagtag	ttaggatatg	ggtgagatga	4020
gtgagatatt	gatttttggg	atagaatttg	agaggggtgt	aaaaaattta	gtaattaaga	4080
taaataggtc	gggcgttagt	gtttacgttt	gtaatttttag	tattttggga	ggtcggatta	4140
tttgaggtta	agagttcag	attagtttgg	ttaagatggg	gaaattttat	ttttattaaa	4200
aatataaaaa	ttagtttagt	gtggtggcgt	tagtttgtaa	tttttagttat	ttaggagggt	4260
gaggtaaagag	aattgttttg	tttaggaggt	agaggttgta	gtgagtcgag	attatgttat	4320
tgtatttttag	tttggataat	agagggagat	tattttaaaa	aaaaaaaaaa	aaaaaaaaaa	4380
aaaaaagagg	tcgggcggcg	gtggtttata	ttatgtgatt	ttagtatttt	gggaggtcga	4440
ggcgggtgga	ttatttgagg	tttggagttc	gagatttagt	tggttaatat	ggtgaaat	4500
cgtttttatt	aaaaatataa	aaattagttg	ggtgggggtg	taggtatttg	taatttttagt	4560
tatttcggag	gttgaggttag	gagaattttt	tgaatttgcg	ggcggagggt	tgtagtgaat	4620
taagattata	ttattgtatt	ttagtttggg	taataatagt	aaaattttgt	tttaaaaaaa	4680
aaaaaaat	tttttttcga	gatatagttt	tattttttcg	tttagggtgg	ggtgtagtat	4740
tacgatttta	gtttattgta	atttttgttt	tttagat	cgtatttttag	tttttaagt	4800
agttgggatt	ataggtattt	gttattacgt	ttagtttaatt	tttgtatttt	tagtaggggc	4860
gtggttttat	tatgttgggt	aggttgggtt	tgaatttttg	attttaagtg	atttgttcgt	4920
tttagttatt	taaagtgttg	ggattatagg	cgtgagttat	tacgtttggt	tttttaaat	4980

ES 2 615 354 T3

gaaaatagtg	taaaaattta	cgataaataa	aatattaata	atattattgaa	tttgtatttt	5040
tataattttt	ttttatttgt	tttttagggt	atatttttgt	ttagaaagta	atttaaaaaa	5100
tgtgtagatg	gagtttggat	tttatttgaa	aatgggtgga	gttatggaaa	attttggagt	5160
aggggagtga	aggatagaaa	ttatatgtaa	aagaaatttt	gggtcgggcg	tagtggttta	5220
tgtttgtaat	tttagtattt	tgggaggtcg	aggtaggtgg	attatttgag	gttaggagat	5280
tgagattagt	ttgattaata	tggtgaaatg	ttatttttat	taaaaatata	aaaaaaatta	5340
gttaggtatg	gtggtgtacg	tttgtagttt	tagttatttc	ggaggttgag	ataggaaaat	5400
cgtttgaatt	cgggaggcgg	aggttgtagt	gagttaagat	tgtgttattg	tatttttagtt	5460
tgggtaataa	gagtaaaatt	ttatttttaa	aaaaaagaaa	gaaagaaatt	ttttggtagt	5520
tgatgagaag	gaaatttaat	cggtaggttt	tagtagggga	gatgaggaga	ttttagggag	5580
ggtatttgta	tatgttgtgt	tttagtgtgg	gttagggagt	aggttattat	ttttttcgtt	5640
tatttttttt	ttgttttaat	ttttttaagt	tttggattag	tggatattta	agtgtagttt	5700
aaggaattat	atgtatttag	atatttaggg	ggtgtttggt	aaaaatgtaa	attttggtta	5760
ggtgtagtgg	tttatatttg	taatttttagt	atatttgggag	gtcgaggcgg	gtggattacg	5820
aggttaggag	atcgagatta	ttttggtaaa	tacggtgaaa	ttttattttt	attaataaaa	5880
taaaaataaa	taaaaaaaa	tattagttgg	gcgtggtggc	gggcgtttgt	agtttttagtt	5940
atcgggagg	ttgaggtagg	agaatggcgt	gaattcggga	ggcggagttt	gtagtgagtt	6000
gagattgcgt	tattgtattt	tagtttgggc	gatagagcga	gattttgttt	taaaaaaa	6060
aaagtaaatt	ttttgggtat	tattttatat	tgattg			6096

<210> 8

<211> 6096

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 8

ES 2 615 354 T3

tagttaatgt	ggggtggtgt	ttaagaaatt	tgtttttttt	tttttgaggt	agagtttcgt	60
tttgcgttt	aggttgggt	gtagtggcgt	aatttttagt	tattgtaagt	ttcgtttttc	120
gggtttacgt	tatttttttg	tttttagttt	tcgagtagtt	gggattatag	gcgttcgtta	180
ttacgtttag	ttaatgtttt	ttttggtttg	tttttgtttt	tttagtagag	atgggggttt	240
atcgtgtttg	ttaggatggg	ttcgaatttt	tgatttcgtg	atttattcgt	ttcggttttt	300
taaagtgttg	ggattatagg	tgtgagttat	tgtatttggt	taaaatttgt	atttttaata	360
agtatttttt	gggggttttg	atgtatgtgg	ttttttggat	tatatttagg	gtattattgg	420
tttaaagttt	gaaggggttg	gagtaagagg	aaggtagacg	ggaggagtag	tgatttgttt	480
tttggtttat	attgggggat	agtatgtgta	aatgtttttt	ttagagtttt	tttatttttt	540
ttgttgggat	ttgtcgatta	agtttttttt	ttattaattg	ttagagggtt	tttttttttt	600
ttttttttta	agatggagtt	ttgtttttgt	tgtttagggt	ggagtgtaat	ggtataattt	660
tggtttattg	taattttcgt	ttttcgggtt	taagcgattt	ttttgtttta	gttttcggag	720
tagttgggat	tataggcgtg	tattattatg	tttggttaat	tttttttgta	tttttagtag	780
agatgatatt	ttattatggt	ggttaggttg	gttttaattt	tttgatttta	ggtaatttat	840
ttgtttcggg	tttttaaagt	gttgggatta	taggtataag	ttattgcgtt	cggtttaggg	900
ttttttttat	atataatttt	tattttttat	ttttttgttt	taaggttttt	tatagttttt	960
attattttta	ggtaaagttt	aaattttatt	tgtatatttt	ttaagttggt	ttttggggta	1020
gagagtagtt	tgggaggtag	gtgagaaagg	gttgtggaag	tgtaagttta	gtaaattttt	1080
gatattttgt	ttatcgtgga	tttttgtatt	atttttattt	aaaaaggtta	agcgtggtgg	1140
tttacgtttg	taatttttagt	attttgggtg	gttgaagcgg	gtaggttatt	tgaggttagg	1200
agtttaagat	tagtttggtt	aatatggtga	aattacgttt	ttattaaana	tataanaatt	1260
agttgggcgt	ggtgataggt	atttgtaatt	ttagttattt	gggaggttga	ggtacgagaa	1320
tttgagaggt	agaggttgta	gtgagttgag	atcgtgggtg	tgtattttaa	tttgggcgag	1380
agagtaaaat	tgtgtttcga	aaaaaaagat	tttttttttt	tttgagatag	agttttggtg	1440
ttgttgttta	ggttggagtg	taatggtgtg	attttggttt	attgtaattt	tcgtttcgtg	1500
ggtttaaggg	atttttttgt	tttagttttc	ggagtagttg	ggattatagg	tgtttgttat	1560
tttatttagt	taatttttgt	atttttagta	gaaacggggg	tttattatat	tggttagggt	1620
ggtttcgaat	tttagatttt	aggtgattta	ttcgtttcgg	ttttttaaag	tggtgggatt	1680
atatggtgtg	agttatcgtc	gttcggtttt	tttttttttt	tttttttttt	tttttttttt	1740
gagatagttt	ttttttgttg	tttaggttgg	agtatagtgg	tatgatttcg	gtttattgta	1800
atttttgttt	tttgggttaa	gtaatttttt	tgtttttagt	ttttgagtgg	ttgggattat	1860
aggttagcgt	tattatattg	ggttaatttt	tgtattttta	gtagagatgg	ggttttatta	1920
ttttggttag	gttggtttcg	aatttttgat	tttaagtgat	tcggtttttt	aaagtgttgg	1980

ES 2 615 354 T3

gattatagac	gtgagttatt	gcgttcgggt	tatttatttt	gattattgag	ttttttggta	2040
tttttttaaa	ttttgtttta	gaagttagta	ttttatttat	tttatttatg	ttttggttgt	2100
ttagagattg	tttggaagt	ggagattttt	tttagttttt	tttttttttt	ttagggtaa	2160
agatagtaag	gaatagggcg	attgtttata	ttagtttttg	tttagagat	ttttgagat	2220
tggatagttg	agttcgtgga	gagggatttt	ttatttttaa	gtaggtaatt	atttattttt	2280
ttttagaggg	aacgagggat	ttatttttagt	ttttattttt	gattattttt	ttttagtttt	2340
gattgtataa	tttgaatttt	atatacgggt	tgtaattttt	tttttttttt	ttcgagacgg	2400
agtttcgttt	tgtttttagg	ttggagtgta	gtggcgcgat	ttgggtttat	tgtaattttc	2460
gtttttcggg	tttaagcgat	ttttttgggt	tagttttttt	agtagttggg	attatagacg	2520
cgtattatta	tgttcggtta	atttgtttgt	aattttttta	atttgtttgg	ttttttata	2580
gggagaggat	ttgtatatgt	tgtgttttgt	gtgatgaaag	gagttttttt	tatatatttt	2640
tttttttaat	gtatttagat	tttaaagttt	attatttatt	tatttaataa	atatttatta	2700
agtatttttt	gaattcgggt	taattatttt	ttttttaggt	atttatttaa	tttcgggtat	2760
tttttcgctt	ttttttgagt	ttttttcgcg	ttttatatag	tttttgttta	ttgttggaaa	2820
atatttttgt	taagtttcgt	ttttttatta	gtttttgtttg	ttttttgtgt	gtttgggata	2880
ggttgtaaaa	ttggaggtga	taaagtgtggg	taggaaatgg	agggtttttt	tatattttta	2940
gggtcggttt	tagtttttgtt	attttttggt	taatatcgcg	gatgtaattg	gtatgggatt	3000
cggaagtgtg	tggtaaagcg	tcggggtttc	gtttggctcg	ttttttcggg	cgtttgtttg	3060
tagttagcga	agttttttta	atttaggttt	gggttttgcg	agttttaggt	gtttttgttt	3120
cgcggcgttg	gcgaagtcga	agttcgagaa	cgtttatcgt	agcgatgcca	aggtcgtttt	3180
tggggtgggg	ttgaggttgt	agttgttttt	ttttcgtatt	aaggatttta	atttttagcg	3240
acgtagtcgt	cgttttcgtt	taggttggga	ggtattgtag	ggattcgacg	ttttaggtgg	3300
ttaaagagcg	attttttttg	atttttagggt	tttggcgggg	taggttttag	tatcgtattc	3360
ggcggaggtc	gaaggtttgt	ggggtaggat	aggagttttt	cgtgtcgtcg	gaagggtcga	3420
ggacgaagga	gggcgttaat	ttatttttta	ttgggttggc	ggtaacgtcg	aatttcgtag	3480
tgatcgcgga	gggttaaggt	gaaaattggt	ggggcgctcg	agggtaggtg	tggggagggg	3540
cggtttttagg	gagtaaggag	tttatttggt	tcgctcgtcg	agttgttttt	ggtcgcgctg	3600
ttacgttttt	ttttgggtta	cgattttcgg	atttaattgt	gttcggtttt	tgtttttttt	3660
tttttgttgt	tgttgccgcg	gttgtaattt	gacgtttagg	ttggggcgtt	tagggcgtag	3720
ttttcgttta	ggtcgtttagc	gtttttattcg	ttttattggg	ttatagattc	gttggcgttg	3780
gggtttcgtt	ggcgttgggg	tttcgttggc	gttggggttt	cgttggcgtt	gggtttcgtt	3840
tggcgttggg	gtttcgttgg	cgttgggggt	tcgttggcgt	tggggtttcg	cggttgtgta	3900
cgtgagttcg	cgtggtttcg	tttcgcgcgg	ttgtgtacgt	gagttcgcgc	ggtcgcggtt	3960
cgtttcgttt	tagggagtta	gcgcgctcgtt	atttgggatg	ttaggatttt	cgttgtggtt	4020
tttggattgt	tttgggggat	ttcggcgtat	ttttaggatt	taggagtttt	ggaagtgttt	4080
tgagagaaat	tatttttggg	agggtttcgt	atttagtttt	ttgtttcggg	ttcggatcgg	4140
ggtttcgggt	taaggtgttt	agaggaatag	ttgatttttt	tatttttgcg	tagggtagag	4200
atttttaaat	ttttttttaa	aatgtagggt	tttagttttt	tttagggagt	tagtgaattt	4260
agatttttag	ttttttgagt	ttaagtatga	atagggaaat	ggggattatt	attatgttta	4320
tatttttggg	ggttaggaag	tttaggtagg	tttttgttta	ttgtagacgg	attttttttt	4380
ttaggggtta	agaaaggttt	tgtatagtaa	gttaattaag	ttttattagt	agagttgcgt	4440
tgtaatagg	tttttagtgt	tttttttata	tttttagttt	taagcgatat	aggatatagt	4500
tttaatttta	ttttttttta	tttatttttt	gtttttgtttg	tttatttttt	ttgttatata	4560
ttttaagtta	tttatttaat	ttttagtttt	tggtaatttt	gttttagttt	ttggaagttt	4620
agtttttagtt	tgtttattat	tatttttatt	ttggtttttt	ttgagaataa	gtggaaggta	4680
aatagagttt	aggtttgaat	ttattcgttt	gaaaataaat	ttaagttaaa	ttttgtggga	4740
gtagtgtttg	tttgggggat	tttaggttta	tgtgtatttt	tgaatttttt	cgggtatttg	4800
ttttaaatgt	aaagaggttg	ttatttaatg	tttgggtttt	tttgattttt	tgtttttttt	4860
tgggtgtttt	tttgtttttg	tcgtttttta	tttttttttg	ttatggttag	tatttggttt	4920
ttaggttaga	aaaggtggat	ttcgtgtttt	tggattttat	taatgttagg	agttatataa	4980
atatttttat	atatatatag	agagggttcg	cgtatgtttg	gaaattttta	gatattaaag	5040
aataaagtgt	aggatataat	agagatattg	tatatattga	gaaatgtgta	tattttttta	5100
atgtagagaa	atgtatatta	atatatatag	atatatttta	taaataatatt	tatagaaata	5160
gatataattta	tatttttaaga	aatattaaga	tatttatgaa	gggaatattt	tagatatgta	5220
taggattttg	aattttatag	gatatatagata	tttaggagta	ggtgggtttt	gatttaggtg	5280
tatatagttt	aatatatagt	atttatatat	aagttttttt	aagtttaaat	gtcgtaaagag	5340
attttttatag	aaagaaaatt	tttttagagg	tttttaaggt	tttgtttgga	aggaagagga	5400
agaaagtgtt	cgttaggtat	cgaaatgtta	aggtattgta	aagtgaaaat	tattttttta	5460
atttcgtggg	aatagtattt	ttatttttat	tttttagggg	aataggggag	tggtttaatt	5520
cgttttgtaa	attaatttta	gatttataag	agtgtttttt	ttttcggcgg	ggagaggtta	5580
ggtttttagt	gttgtagttt	agtgaatggt	gattcgttga	ggtttatttt	aggttatagg	5640
gtgttgggta	aattttattaa	gttcgtacgt	ttgttttaata	tttgggtggg	ttaggaagag	5700
tgtttgtggtg	atatttagtt	tggcgatata	taggaggtac	ggtcggtgaa	gaaaatggat	5760

# ES 2 615 354 T3

ttgtgtatta	agggtttgta	ttatttcgtc	gggttttggt	ttgatagatt	atTTTTtatt	5820
TTTTtTgatt	acggattggg	ttaggaggtt	tattttcggg	tcgTTtattg	tttgtaaata	5880
gtagagggta	gcggtgaggg	cgtataaagt	tttagtcggt	tatgTTTTtt	taagTTTTtg	5940
ttatttaggg	TTTTtgTTtg	tttcgTTtaa	ggcgtattag	tttcgTTtg	ggcggcgacg	6000
gtaggTTta	ggaatcgTTa	tgTTtaggta	ttgagTTgga	gTTTTgggta	TTTTattgtg	6060
tagtgtaaaa	agggaTTTT	gaTTTTata	ttggta			6096

<210> 9

<211> 2501

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 9

ES 2 615 354 T3

tttgtaa	atg	gagata	tattata	atagta	atgttt	agttt	60
tata	gtgata	aaggata	tttttt	ttttt	gtgag	gatgata	120
gtatg	atgggt	tatgatt	at	gtgttt	tatcgt	ttta	180
ttta	tttttt	taggtag	taga	tttag	gagatt	tagt	240
tatg	agagatt	ttaagt	gtatg	attagga	aaagt	tttt	300
gtttaa	taatagg	tttttt	gtaatt	ttttta	tttaatt	ttata	360
aggtg	at	taattaa	agaatt	gttgata	tttaaat	atg	420
at	at	aattatt	tttttag	tttaagt	tttatt	tagga	480
at	taata	ttatta	tagttg	taagaga	tttaa	aggtt	540
gagga	tagcgg	gtttgt	cgtaa	taaagg	gttag	tttt	600
gttg	tttgt	tgata	acgtat	at	ttaa	tttt	660
at	agaatt	taaa	aaattaa	tattac	ttatt	gatta	720
aaggt	tataa	tttgt	tttag	ggatt	cgtag	aat	780
ttata	ttgatt	ttttat	attgcg	ttgttt	aatg	taata	840
at	ttta	tcgtg	gctgg	ttatt	aat	taata	900
ttag	gagagg	tgagg	at	tagtt	aat	acgga	960
gatt	ttgga	aattag	gctgg	ggcgc	ttc	cgtaa	1020
tcgg	atag	atgat	tgtatt	ttg	ggt	ttat	1080
agta	ggagt	aaata	tg	ttt	ggag	cgat	1140
tagt	tg	tagc	agac	ttc	gtata	aat	1200
agtag	aataa	aaaga	tg	agt	tttt	tagt	1260
gagaa	gggt	aagg	tg	aagg	taatt	atcga	1320
aatc	ggt	tg	g	tata	g	atc	1380
ttt	gg	gtat	cg	gatt	tttag	gata	1440
tttt	ttc	ttt	tg	gat	acg	tcgc	1500
ggc	g	gtc	tgat	gat	cgtag	at	1560
tac	tttag	tttt	gatt	gttt	cggt	tagata	1620
tgac	gagg	at	ttgt	ttgg	ggtag	cg	1680
ttata	gttg	cga	at	ag	aa	ttc	1740
tcgg	aggg	taa	aat	ggg	aaa	agc	1800
tttt	cgt	at	gt	ttt	acg	ttat	1860
gttt	att	aac	at	at	aa	ata	1920
aagg	aat	ttata	at	at	ta	at	1980
tttt	tag	t	ag	t	t	g	2040
ttg	ta	at	at	t	at	at	2100
tg	ttat	tttag	ttat	ta	tttag	ta	2160
ttta	ttgt	aaa	tttt	ta	gt	tttt	2220
ttta	ggat	tgaa	gatt	tatt	taa	tttt	2280
tata	tttaa	tattag	tttat	ggt	gt	tttc	2340
tatt	ttaa	taatt	tatta	ttg	tttag	gatt	2400
tttt	gtg	at	gatt	att	tttag	gatt	2460
taatt	gga	aga	at	a			2501

<210> 10

<211> 2501

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 10

ES 2 615 354 T3

tataagagtt	tgttttaagg	atthtaattt	tattthtaatt	aagttthtaa	gttaatgtaa	60
ttattaaaat	ttcgaagaga	ttttatttaa	atthtataaa	aggttthtaa	agttgtttag	120
aaatthtggt	gaaatagatt	aggaaatttg	gaaaggaaat	aatgcggaga	ttcgtagtat	180
taaattatga	gattthtaat	ataatthtaa	atattaatgt	aataaaattt	aaatthtggt	240
gtaataaaaat	tataaatttt	aatattgggt	ttaagtatag	agaaaaagta	tatttatggt	300
gaatgtggaa	aatattattt	ttaaaatata	gtcgaattaaa	aaattggttg	ggaattgatt	360
ataattattg	ataatthtta	agaaatatag	atattaaaat	attatthtta	ttttthta	420
agaaattgggt	taaattataa	ttaatataag	gaggttataa	aattthtat	ataatattgt	480
atatattttt	ttggaaaaat	atgtgtaatt	gttttgtaaa	atatatgatt	aattagttt	540
tgtgatggga	taatattgta	gtttthta	tttttgttt	tttttgatt	ttttatagta	600
ttgatgtata	tttttgct	ttaaaagtaa	ttttthaaag	ttttataacg	tggtataaaa	660
atattatgta	cgttataaaa	tttagaatac	ggaaataaga	agttcgtttt	ttttthttt	720
ttatthcgggt	ttggthttt	agatthttt	ttttthtcg	gggcgggatt	tttcgtcgt	780
ttttthtagg	tttttagttc	ggttcgttaa	ttgtcgtata	aaggcgttgt	tttaggttag	840
agttthtata	aagcgttggg	tgagatthtt	ttgttcgtt	atgtttggte	gcggtaaagg	900
cgggaagggt	tttggtaaag	gcggcgtaa	gcgttatcgt	aaagtattgc	gcgataatat	960
ttagggtatt	attaagtcgg	ttatthcggc	ttttgttcgt	cgccggcggc	tgaagcgtat	1020
tttcggtttt	atthacgagg	agatthcggc	ggtgttgaa	gtgttttg	agaacgtgat	1080
tcgggacgct	gtgatttata	tagagtacgt	taagcgtaa	acggttatc	ttatggatgt	1140
ggthtacg	tttaagcgtt	agggtcgtat	tttttacggt	ttcgggtggt	gagcgtttt	1200
ttttattaat	aaaaggtttt	ttttagggtt	ttttthttt	tagttgagga	gtcgtgatgt	1260
ttgtttggtt	agttthttt	tattatthgt	tcgcgtttgt	tgagtttg	ttgttatcgg	1320
agtatgcgtt	tttagtcgtt	gtaagtaggt	attagttatt	aatcgtttt	tagtaataa	1380
aatatthta	ataatthgtt	ttagggtgtt	tagagtttat	tgatatgggt	cgtcgcgtag	1440
attgtagtgt	agtgttatta	tggtttattg	tagttthtcg	attagtcgga	atthcgcgtt	1500
tattattacg	taaggttaat	tattthttt	aagatcgggt	tttcgtgtgt	tttttaggtt	1560
agtttcgata	ttttggttt	aattthttc	tttatgttta	atgttggtat	tatagtagtg	1620
agttattacg	tttggttacg	atattgttga	ggttthtagg	ttagttatat	tttaagggta	1680
atthtcgtag	tgtagtggg	aggaaagtta	agtagttata	ggttthtggc	gttgtgaatt	1740
taattgttga	atatagtaag	aatttattat	gtaataattt	tttaattagt	gatcgtaaat	1800
aagtttagtt	tatcggttt	ggtttaattt	ttgagaaagg	tgagaatttg	aattthtgag	1860
tagaaatac	tagtttatta	gtataataaa	gtgattataa	taaagattaa	tttattthtt	1920
gttatttaac	gtgggtagag	ttttatcgtt	gagaatttht	tagttthtga	gtgtthttt	1980
atthtataat	gaagttgata	agatgggtat	aaaagtgaga	tttttagtaa	gtaaatthta	2040
aggthtgaag	gatttagtgat	taatgggaag	tgtaataaag	ttatatttga	ggttatthta	2100
ttcagatatt	tgatttaatt	ggthtaagga	tattattatt	ttgtgggtta	gatttgaaaa	2160
ataataattg	ttttaaggaa	ggaatttggt	ggtatthtaa	taaaaattt	ttttthtga	2220
tttttatatg	tcgttatthta	gataaatttt	tggtttatat	ggttggtttt	ttttthgtta	2280
gttaaatttt	gtgttattht	gtgaaaaaga	attgagatta	ataaacgggt	agtgagaata	2340
tagggaaaga	taaaaattat	agtaagtthta	tttagttatg	ttgtagttat	ttttataagg	2400
atagggaaag	atthtggtt	ttatttatta	tttaaggtgt	gagtataaat	tttaaaaaata	2460
tatgatatta	taggaataat	gaagatgtt	ttatttgtaa	a		2501

<210> 11

<211> 2501

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 11

ES 2 615 354 T3

ttgttgata	gaatatttta	ttatttaggt	attatgtcga	gtatttaata	gttttttttt	60
ttgttttttt	tttttttttt	attttgtatt	ttggagttaa	ttatagtgtt	tgttgttttt	120
ttgtttgtgt	tataagtttt	tattatttag	tttttattta	taagtgagaa	tatttagtat	180
ttggattttt	gtttttgtat	tagtttgtaa	aggataatag	tttttagttt	tatttatggt	240
ttataaaaag	atatgattta	gtttttttta	atggttgtat	taaatagaagt	tttaaagata	300
taataaaaat	attaattttt	tttttattat	aaaaattttt	tgttgaattt	gattatattt	360
aaattaacga	gttttgtttt	atgaaagatt	ttttggataa	atttgatagt	tgatggaata	420
ggagaagttg	tttgttatgt	ttaaagttaa	taagagatta	atatttagaa	taaataggaga	480
tttgtaaatt	aatagaaagt	aggtagtaaa	gttaaagaaa	atagttaaag	gtatagttaa	540
taaaaggaac	gtgattatgt	tttttgtagg	gatatgggtg	gagttggaag	tcgttagttt	600
tagtaaaatt	atataggaat	agaaaattag	cgagatcgta	tggttttatt	tataagtggg	660
agttgaataa	tgagaatata	tggttatatg	gcggcgatta	atataatattg	gtgtttgttg	720
agcgggggtg	tggggaggga	gagatttagg	aagaatagtt	aagggatatt	gggtttaata	780
ttgggggtgat	gggatgattt	gtatagtaaa	ttattatggc	gtatatattt	atgtaataaa	840
ttgtatattt	ttttatatgt	attttagaat	tttaaataaa	agttggacgg	ttaggcgtgg	900
tggtttacgt	ttgtaatttt	agtattttgg	gaagtcgagg	cgtgtagatt	atttaagggt	960
aggagtccga	gattagttcg	gttaatatgg	tgaaatttcg	tttttattaa	aaataaaaaa	1020
attagttaga	tgtggtacgt	atttataatt	ttatttatc	gggaggttga	agtagaattg	1080
ttgaattcg	agaggcggag	gttgtagtga	gtcgtcgaga	tcgcttatt	gtattttagt	1140
ttgggttata	gcgtgagatt	acgttataaa	ataaaataaa	ataataaaa	ataaaataaa	1200
ataaaataaa	ataaaataaa	ataaaataaa	ataaaataaa	ataaaaaaat	aaaataaaat	1260
aaaataaaat	aaagtaattt	tttttttttt	aagcggtttt	tatttttttt	ttttgttttg	1320
tgaagcgggt	gtgtaagttt	cgggatcgta	gcggtttttag	ggaatttttt	ttcgcgatgt	1380
ttcggcgcgt	tagttcgttg	cgtatatttc	gttgcggttt	tttttttgtt	gtttgtttat	1440
tttttaggtt	tcgttgggga	tttgggaaag	agggaaaggt	tttttcggtt	agttgcgcgg	1500
cgatttcggg	gatttttaggg	cgtttttttg	cggtcgacgt	tcgggggtgta	gcggtcgctc	1560
gggttggggg	cgccgggagt	tcgcgggatt	ttttagaaga	gcggtcggcg	tcgtgattta	1620
gtattggggc	ggagcggggc	gggattat	ttataagggt	cggaggtcgc	gaggttttcg	1680
ttggagtttc	gtcgtcgtag	ttttcgttat	tagtgagtac	gcgcggttcg	cgttttcggg	1740
gatggggttt	agagttttta	gtatgggggt	aattcgtagt	attaggttcg	ggttttcggg	1800
agggtttttc	gtttatttcg	agattcggga	cgggggttta	ggggatttag	gacgttttta	1860
gtgctgtag	cggttttttag	ggggttcgga	gcgtttcggg	gagggatggg	atttcggggg	1920
cggggagggg	gggtagattg	cgtttatcgc	gttttgggtat	tttttttcgg	gttttagtaa	1980
attttttttt	gttcgttgta	gtgctgtttt	atatacgtgg	ttatttttta	gttcgaggta	2040
ggagtatgtg	tttggtaggg	aagggaggta	ggggttgggg	ttgtagttta	tagtttttcg	2100
tttattcgga	gagattcgaa	tttttttatt	ttttcgtcgt	gtggttttta	tttcgggttt	2160
tttttttgtt	tttcgttttt	ttcgttatgt	ttgttttttcg	tttttagtgtt	gtgtgaaatt	2220
ttcggaggaa	tttgtttttt	tgtttttttt	ttgtattttt	gatttttttt	cgggttggtg	2280
cgaggcggag	tcggttcggt	ttttataatt	cgtatttttt	ttttttcgta	ggcgttgctg	2340
cggttttgctg	tatggtggtg	gtagattagg	gtagaggttg	gaaggaggag	gtgggtgatcg	2400
tggagacgtg	gtaggagggt	ttatttaaag	ttttttgctg	aagtgattat	gttcgggtaa	2460
ggggaggggg	tggtgggttt	taggggggtg	tgattaggat	t		2501

<210> 12

<211> 2501

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

ES 2 615 354 T3

<400> 12

gatttttagtt	atagtttttt	aaggtttagt	atTTTTTTTT	tttgttcggg	tatggttatt	60
tacgtaggag	gttttgagtg	agtttttttg	ttacgttttt	acggttatta	tttttttttt	120
ttagttttgg	ttttgatttg	ttagtagtat	gcgtagggtc	gcgtagcggg	ttgcggggag	180
ggagaagtac	gagatgtggg	gatcgggtcg	atttcgtttc	gtagtaattc	ggggaggggg	240
taggagtgta	gggagggaat	agggaaatag	gttttttcga	agattttata	taatattggg	300
gcggggagta	ggtatggcgg	gagagggcgg	gaataggaag	gaggttcggg	gtaaaagtta	360
tacgacggag	ggataagggg	gttcggattt	tttcgggtgg	gcgaggggtt	gtgggttgta	420
gttttagttt	ttgttttttt	tttttgtag	atataatggtt	ttatttcgaa	ttgggaaata	480
gattacggtg	tagggcggta	ttgtagcгаа	taaagaaaag	ttgttgagg	ttcgggggag	540
gatgttaagg	cgcggtgagc	gtagtttggt	tttttttttc	gttttcgggg	ttttattttt	600
tttcgaggcg	tttcgggttt	tttgaaagtc	gttaacggta	ttggggacgt	tttgggtttt	660
ttaggttttc	gtttcgggtt	tcgaggtggg	cgaggagttt	tgtcgggagt	tcgggtttga	720
tgttgcgggt	tggttttatg	ttgggagttt	tgagttttat	tttcggggac	gcgggtcgcg	780
cgtatttatt	ggtggcgaag	attgcggcgg	cgaaatttta	gcgaaggttt	cgcggttttc	840
gagttttata	agggtggttt	cgtttcgttt	cgtttttagtg	ttgagttacg	gcgtcggtcg	900
tttttttggg	gggtttcgcg	gattttcgtc	ggtttttagtt	tcggcggtcg	ttgtatttcg	960
ggcgtcggtc	gtagaggggc	gttttgaggt	tttcggagtc	gtcgcgtagt	tggtcgggga	1020
agtttttttt	ttttttttag	gttttttagcg	gggtttaggg	agtaaataga	tagtaggaag	1080
aggatcgtag	cgaagtgtgc	gtagcgaatt	ggcgcgtcgg	gatatcgcgg	ggggaaattt	1140
tttaagatcg	ttgcgatttc	ggagtttgta	tattcgtttt	atagggtagg	ggagaggggt	1200
ggaggtcgtt	tagaggaaag	gaaattggtt	tattttattt	tattttattt	tattttttta	1260
ttttatttta	ttttatttta	ttttatttta	ttttatttta	ttttatttta	ttttgtgta	1320
ttttatttta	ttttatgacg	tagttttacg	ttgtggttta	ggttgaggag	tagtggcgcg	1380
atttcggcgg	tttattgtaa	ttttcgtttt	tcgggtttta	gtaattttgt	tttagttttt	1440
cgagtaggtg	gaattatagg	tgcgtgttat	atttggttga	tttttgatt	tttagtagag	1500
acggggtttt	attatgtttg	tcgggttggt	ttcgaatttt	tgattttagg	tgatttgtag	1560
gtttcgggtt	tttaaagtgt	tgggattata	ggcgtgagtt	attacgtttg	gtcgtttaat	1620
ttttatttga	agttttgggg	tatatgtaga	ggatgtgtag	gtttgttata	taggtgtgtg	1680
cgttatgatg	gtttgttgta	tagattattt	tattatttag	gtattaagtt	tagtattttt	1740
tagttatttt	ttttgggtatt	tttttttttt	agtatttcgt	ttaataggta	ttagtgtgtg	1800
ttgatcgtcg	ttatgtgatt	atgtgttttt	attgttttagt	ttttatttat	aagtgtgatt	1860
atgcggtttc	gttgggtttt	tgtttttggt	tgagtttggt	gaggttaacg	gttttttagtt	1920
ttatttatgt	ttttgtaaag	gatatgatta	cgtttttttt	agtggttgtg	ttttagggtta	1980
ttttttttgg	ttttgttggt	tattttttgt	tgatttgtag	atttttattt	attttagata	2040
ttgatttttt	gttgggtttta	gatatgatag	atagtttttt	ttattttatt	aattgttaag	2100
ttgttttaag	gagtttttta	tgaaataaaa	ttcgtttaatt	taagtgtaat	taaatttagt	2160
aagggttttt	tgtgggtggg	aagaggttgg	tgtttatggt	gtatttttaa	aattttattt	2220
aatgtagtta	ttaaaaagaa	ttagattatg	ttttttgtgg	gaatatggat	ggagtttagag	2280
gttattattt	ttagtaaatt	aatgtaggaa	tagaaattta	aatattggat	gtttttattt	2340
gtaagtggga	gttaaagtat	gagaatttat	aatataaata	aggaaataat	agatattgtg	2400
gttgatttta	gggtgtagga	tgggaggaag	gagaggagta	gaaaagagaa	ttattgggta	2460
ttcggtataa	tattttgggtg	atgaaatatt	ttgtataata	a		2501

<210> 13

5 <211> 1920

<212> ADN

<213> Secuencia Artificial

<220>

ES 2 615 354 T3

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 13

ttaggttggt	gtagatatag	tttttgtttt	tttgaaaaat	atgttttagg	tgttgggatt	60
ttagatatatt	gggaaataga	gtgtatgtag	ttgttgagag	gttttggttt	tggttttttt	120
tattattgag	gtgtagaggt	gttgtggata	gtttagattt	atatgggtgt	tgaggtgaaa	180
tagaattttt	agttttttta	tgaggttatt	ggattttttg	gttgttttta	gagttttttg	240
atntagagtt	gaatgtaaag	taagtgtttg	aaatgtagaa	gtagttgggg	ttgtttatgg	300
tatttgtttt	gtttgggggt	agagaagatg	ttaggttgag	gttttagtga	ttttaggtat	360
tagttttgaa	ggaggggtgg	gagattgtaa	aggggaagtg	ttggagggt	taatggtttt	420
tgtgtatttt	gtgttttttt	gaagtgtgtt	gtttttttgt	gttggggatt	gggatttggt	480
tttggggaat	ttgttttaga	gatgggtggt	gattgggggt	gggtattttt	tagggttggt	540
aggttttttt	tagttttgta	tttgttgtgt	tgttttattt	tgtaggaag	tttagagat	600
tttggggatg	gggtgggagt	gtttttttat	tgtgggttta	aaaagaagga	aggatgtttt	660
taggggttgt	agaaggagga	ttagttttaa	gttataattt	ttttggatt	taaggtaggt	720
tggttgggggt	tttgtgttta	tatggttttt	gggtgggggt	tgtgtgtttt	gggagttttg	780
tggtttgggg	aggaaagagg	agataagaga	taggtgagga	ttatgggggt	gatttagttg	840
gggtagggat	tattgtggaa	aaattttggt	gaggtggggg	gatgtggaaa	gagagtgggt	900
tgtgttttgt	attttgtgtt	gggtattttg	tgtagtgttt	ttgtttttag	tgttttgtgt	960
tttgtgtttt	gtgttttggt	tttatttttg	gttagttgta	ttgtgtttgt	gtttaggaa	1020
ttgtggagtt	ggaaagtggg	ggtgttgtgg	ttgggggggt	gttttagttg	tgttttgggt	1080
agtgattgggt	gggttgggtt	taaatttagt	taggttgggt	agggtgggtg	tgtgtgattg	1140
gggattgggt	gttttgtttt	ttttgttttt	tttttttttt	tttttttttt	tttttagttt	1200
ttggtttttt	ttagttttta	ttggatttgt	ttgtttgttg	tttttttttt	ttttttgttt	1260
tttatattat	tttttatttt	tttgttttgt	ttttgttttt	tttttttttt	ttgttttttg	1320
tttttttttt	tttttttttt	tttttttaggg	gtggagtttt	tttttttttt	tttagataat	1380
gttgtgggtg	tgtttttttt	tttgttagtt	tgtttagggt	tttgttgttt	gtgatttttt	1440
tgggttgggg	gtggggaggt	ggggggggag	tgtagggttg	gggaggatga	gttggttttt	1500
tttatttttt	tggttgtgtt	ttttttaaga	gggatggaga	tttggtttaa	gttttttgggt	1560
ttattttggag	ttgtgatagt	tatttttagg	gaatagttat	gttgttttat	taagtttatt	1620
tttagtggtt	tggatttttt	aggtagaggt	tgtgggattt	tgtttttttt	aatatttttag	1680
tttattttta	aaagggtttg	agttggatag	gggttaaata	ggtttttttg	atttgggtggg	1740
ttggttagat	gtgatagtaa	tgttaaggag	gttaagtttt	tttgtttatt	ttttattttt	1800
ttttttttta	tttttggatt	ttttggtgtt	tttagtatat	agaggttttt	gagtagtttg	1860
gtttagggtt	ttttatttat	ttagagtttt	tttttttatg	tgtttatttt	taattttgta	1920

5 <210> 14

<211> 1920

<212> ADN

<213> Secuencia Artificial

<220>

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

<400> 14

ES 2 615 354 T3

tgtagggttg	gggataggta	tgtgaggggg	agaatthttga	gtagataggg	aatthttgtagt	60
tgggtttgtt	aagggttttt	gtgtattggg	ggtgttagga	ggtttaggga	tggaaaaggg	120
ggaggtgaga	aatggataaa	gaaatthttgt	ttttttggta	ttgttgttat	gtttggttgg	180
ttgtttaagt	tgaaggggtt	tgttttagtt	ttgtttggtt	taaatthtttt	tgggaataag	240
ttgggatggt	agaaaaata	aaatthttata	atthtttgtt	ggggaatthta	ggttggttga	300
ggtgggtttg	gtagggtagt	gtgattggtt	tttgggagtg	gttgttatag	ttttgggtga	360
attgaggagt	ttgggttaag	tttttatttt	ttttggagag	gtagtagta	aggaggtgag	420
gggagttagt	ttatthtttt	taatthttgta	tttttttttt	atthtttttat	ttttagtttg	480
gaagaattgt	tgggtgggtgg	agtthttgatg	agtthttggg	gaaggggatg	tagttatagt	540
attgtttggg	agggagggga	gaagtthttgt	tttttagaggg	gaggggagaa	gaggaggggg	600
taggagataa	ggggaggaag	aaaggtgaag	gtaaggtgaa	ggggtggaga	gtgatatgaa	660
gagtgagaga	aaagagagga	tagtggatga	gtagatthttg	taggggttga	aaaaggttaa	720
ggggttggag	ggagggagag	gaaggaggag	gggagtgagg	agggtggggt	gtttggtttt	780
tagttgtgtg	gttattgttt	gttttagttg	gttggatttg	agtthttggtt	gttgattgtt	840
ggttgagggtg	tagttgaagt	agtthtttag	ttgtggtgtt	tttattthttt	aatthttatgg	900
tttttgtggt	gtgggtgtga	tgtagttggt	ttggggtgaa	ggttaaggtgt	ggggtgtggg	960
gtgtggggta	ttgggagtga	ggtattggtg	tgggatgttt	ggtgtaaggt	gtagggtgtg	1020
ggttgthttt	tttttgtgtt	tttttatttt	gttaaagttt	ttttatgatg	gtttttattt	1080
tggttgggtt	agtthttgtaa	tttttgtttg	ttttttgttt	tttttttttt	ttttgagttg	1140
tggggthttt	ggggtgtgtg	gattthttggt	aggggttgtg	taggtgtgga	gttttagttg	1200
gtttgttttg	ggtttgaaga	aagtthttggt	ttggagttag	tttttttttt	atgattthttg	1260
ggggtgtttt	tttttttttt	tgagtttgtg	atgggaaggt	gtttttattt	tattthttggg	1320
gtttttgaga	ttttttgggtg	aggtgggggtg	gtgtggtagg	tgtagggttg	gggagggttt	1380
gatagthttg	gagagtgttt	gattthtagtt	tgttgttgtt	ttttaggttg	atthtttaga	1440
ggtaggttht	agtthttgggt	gtggggaggt	ggtgtgtttt	agaggggtgt	agggtgtgtg	1500
ggggttgttg	gttttttggg	tattthtttt	ttgtggtttt	tttgthtttt	tttgagttg	1560
gtgtttgagg	ttgttgggat	tttagtttgg	tgtttttttt	tgttttgagt	gaggtaggtg	1620
ttgtgggtgg	ttttggttat	ttttgtattt	tgagtgttta	ttttgtattt	agtthttaagt	1680
tggagagttt	tggggatagt	tgagagtgtt	agtggthttta	tagggagatt	gagggthttg	1740
ttttatthttg	ggtgttgtgt	gggthttgggt	tgtttatagt	atthtttgtgt	tttagtgata	1800
ggagaagtta	agtgtgaggt	tttttaatag	ttgtgtgtat	tttattthttt	aggtatttgg	1860
aatthttggtg	tttagaatgt	gttttttggg	agagtaaaag	ttgtgtttat	ggtagtttgg	1920

<210> 15

<211> 6096

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 15

ES 2 615 354 T3

tattagtgtgta	agatttaaaaa	tttttttttt	gtattgtata	gtgagatggt	tagggtttta	60
gtttagtgtt	tggatagat	gatttttggg	tttgtttgtt	gttgttttaa	gtgaagtgg	120
tgtgttttgg	gtggagtaga	tagagatttt	gggtggtagg	ggtttgggaa	gatatgggtg	180
gttagggttt	tatgtgtttt	tattgttgtt	ttttgttatt	tgtaggtaat	ggatgagttg	240
ggaatgagtt	tttagatta	gtttgtgatt	aagaaaggta	aggaatgggt	tgttagggta	300
gagtttggtg	agatggtgta	ggtttttggg	gtatagattt	atttttttta	ttggttgtgt	360
tttttgtgtg	ttgttaggtt	gggtgtttat	taggtatttt	ttttggttta	gttagatggt	420
aggtagatgt	gtgggttgg	tgagtttgtt	tagtattttg	tggtttgggg	tgggttttag	480
tggattagta	tttattgggt	tgtagtattg	ggagtttggg	ttttttttgt	tgagggggag	540
ggtatttttg	tggatttggg	gttgatttgt	agaatgagtt	aaattatttt	tttgtttttt	600
taagagatgg	gaatggaagt	gttgttttta	tggagttggg	gaaatgattt	ttattttata	660
gtgttttagt	attttgggtg	ttgggtggta	tttttttttt	ttttttttta	ggtagggttt	720
tggaggtttt	tgggggaatt	ttttttttgt	gggagttttt	tgtggatttt	agatttaggg	780
gagtttgtgt	gtgagtattg	tgtgttaggt	tgtgtgtatt	tgagttaggg	tttatttgtt	840
tttgggtggt	tgtgtttatg	tgagtttagg	gttttgtgta	tgtttgaaat	gtttttttta	900
tgggtgtttt	agtatttttt	ggagtgtgag	tgtgtttgtt	tttgtgaatg	tgtttgtgag	960
gtgtgttttt	gtatgttggg	gtgtattttt	ttgtatttgg	gggatgtata	tattttttta	1020
tatgtatagt	atttttgttg	tgttttgtat	tttgtttttt	ggtatttgag	gatttttaag	1080
taatgtgtggg	ttttttttgt	gtatataatg	gagtattttt	gtgatttttg	gtattagtaa	1140
aaatttaggga	tatgggattt	attttttttg	gtttgaggt	taagtattgg	ttatgatagg	1200
ggaaggtgag	agatgataaa	aatagagaga	tagttagaga	ggagtagaga	gtttagagggg	1260
tttaggtatt	ggtagtagt	ttttttatat	ttggggtagg	tgtttgaaag	aatttagagg	1320
tgtatatgag	tttaggtgt	tttaggtagg	tattgttttt	atagggtttg	gtttgagttg	1380
ttttttaaat	gagtgaattt	aagtttgggt	tttatttgtt	ttttatttgt	tttaggggga	1440
ggtaaggtg	gaagtgggtg	tagtaggggt	ggggttggat	ttttaggagt	tggggttgag	1500
ttataggag	ttgggggttg	ggtggatgat	ttggagtgtg	tagtagggaa	gatgaggtaa	1560
tagggtagga	agtgggtggg	gggaggtgga	attggggttg	tgttttgtgt	tgtttggaat	1620
tgggagtgtg	ggaaagatat	taggaatttg	gttgtagtgt	agttttgttg	gtggggtttg	1680
gttggtttat	tgtatagagt	tttttttgat	ttttgaagaa	agagatttgt	ttgtagtggg	1740
taaaagtttg	tttggatttt	ttggttatta	gaaatatgag	tatgggtggg	gttttttagtt	1800
ttttatttat	gtttgggttt	aagagatttg	gagtttaggt	ttattgattt	tttagaaaag	1860
attaagattt	tgtatttttag	aaagaggttt	ggggattttt	gttttgtgta	agggtagaag	1920
gattagttgt	ttttttgagt	attttaattt	ggaattttgg	tttgaagttg	agataggaga	1980
ttggatgtga	ggttttttta	gagttgggtt	tttttaaaata	atttttaaaa	tttttagatt	2040
ttaggggtat	gttgaatttt	tttaaagtag	tttaaagaat	ataatgagag	ttttaatatt	2100
ttaggtgggtg	gtgtgttggg	tttttgaggt	gggggtggat	gtgggtgtgt	ggatttatgt	2160
gtataaattgt	gtgggatggg	gttatgtgga	tttatgtgta	taattgtggg	attttagtgt	2220
tagtgggatt	ttagtgttag	tgggatttta	gtgttagtgg	gatttttagtg	ttagtgggat	2280
tttagtgtta	gtgggatttt	agtgttagtg	ggattttagt	gttagtgggt	ttgtggttta	2340
gtggagtggag	tggagtgttg	gtgatttgag	tggagatttg	gttttggatg	ttttagttta	2400
gatgttaagt	tatagttttg	gtagtagtag	taaaggggaa	ggggtaggag	ttgggtatag	2460
ttggatttgg	aggttgtgat	ttaggggaaa	gtgtgggtgg	ttgatttagg	gtagtgtggt	2520
tgggtgaggta	gggtgggttt	ttgttttttg	gagttgtttt	tttttatatt	tgttttgggt	2580
gttttttagta	gtttttattt	tggttttttg	tggttattgt	gggatttggg	gttgtgttta	2640
gttttagtggg	gagtgaatta	gtgttttttt	ttgttttttg	tttttttgat	ggtatgagga	2700
atttttgttt	tgttttatag	atttttgggt	tttgttgagt	gtggatttgg	agtttgtttt	2760
gttagggttt	tggaattaga	gaaagtgtt	ttttggttat	ttgaagtgtt	ggatttttat	2820
agtgtttttt	agtttgggtg	ggagtgttgg	ttgtgttgtt	gaaggttggg	gtttttgggt	2880
tgaaaggggag	gtagtgttag	tttttagttt	attttagaag	tggtttttgt	attgttgtgg	2940
tgggtgtttt	tgggttttga	ttttgttagt	gttgtggggg	agaggatttt	ggagtttcta	3000
gggttttagat	ttgggttggg	aaagtttttg	tgattgtagg	taagtgtttg	ggaggggtgg	3060
ttagggtgaag	ttttgtgtt	ttattatata	tttttgggtt	ttatgttagt	tgtatttgtg	3120
gtattgggta	ggaaatggta	gggttgaggt	tgattttagg	agtataaggg	agttttttat	3180
tttttgttta	tatttgttat	tttttagttt	gtaatttatt	ttagatata	agaaagtaag	3240
taggattggg	gggagatgg	agtttaatag	gaatattttt	tagtagtgag	taggggttgt	3300
atgggatgtg	ggaggagttt	agaggaggtg	tggagagtgt	ttgaggttgg	gtgagtgttt	3360
agagggggaga	tagttgaatt	gggtttaaga	gggtgttagt	gggtgtttgt	tgaatgaatg	3420
agtgatgggt	tttgaagttt	gagtgtattg	aaagaggggg	tgtgtaaaaa	gggttttttt	3480
tattatatag	gatatagtat	atgtaatttt	tttttttggg	gaaaagttag	ataggttaaa	3540
aaggttataa	ataaattagt	tgggtatggg	gggtgtgtgt	tgtagtttta	gttattaggg	3600
aggttagagt	aggggaattg	tttgaatttg	ggaggtggag	attgtagtga	gttaagattg	3660
tgttatagta	ttttagtttg	gaaatagagt	gagattttgt	tttggaaaaa	aaaaaaaaaa	3720
gttataaatt	gtgtgtgggt	tttaggttat	ataattagag	ttggagggga	gtgggttaagg	3780

ES 2 615 354 T3

atgagaattg	agatggattt	tttgTTTTTT	ttggaggaga	gtgggtgggt	gtttatttgg	3840
gggtggggaa	ttttttttta	tgggtttagt	tgtttaattt	taggggattt	ttaggatagg	3900
agttgatgta	aatagttggt	ttattttttg	ttgtttttgg	ttttggagaa	ggaggagggg	3960
gttggggagg	gttttttatt	tttagataat	ttttaagtag	ttaggatatg	ggtgagatga	4020
gtgagatatt	gatttttggg	atagaatttg	agagggtggt	aaaaaattta	gtaattaaga	4080
taaatagggt	gggtgtagtg	gtttatggtt	gtaattttag	tattttggga	ggttggatta	4140
tttgagggtta	agagtttgag	attagtttgg	ttaagatggt	gaaattttat	ttttattaaa	4200
aatataaaaa	ttagtttagt	gtggtgggtg	tagtttgtaa	ttttagttat	ttaggaggtt	4260
gaggtaagag	aattgtttga	tttaggaggt	agaggttgta	gtgagttgag	attatgttat	4320
tgtatttttag	tttggataat	agagggagat	tattttaaaa	aaaaaaaaaa	aaaaaaaaaa	4380
aaaaaagagg	ttgggtgggtg	gtggtttata	ttatgtgatt	ttagtatttt	gggaggttga	4440
ggtgggtgga	ttatttgagg	tttggaggtt	gagattagtt	tggttaatat	ggtgaaaatt	4500
tgtttttatt	aaaaatataa	aaattagttg	ggtgggggtg	taggtatttg	taattttagt	4560
tattttggag	gttgaggtag	gagaattttt	tgaattttgt	gggtggaggt	tgtagtgaat	4620
taagattata	ttattgtatt	ttagtttggg	taataatagt	aaaattttgt	tttaaaaaaa	4680
aaaaaaaaatt	ttttttttga	gatatagttt	tatttttttg	tttaggttgg	ggtgtagtat	4740
tatgatttta	gtttattgta	atttttggtt	tttagatatt	tgtatttttag	ttttttaagt	4800
agttgggatt	ataggtattt	gttattatgt	ttagtttaatt	tttgtatttt	tagtaggggt	4860
gtggttttat	tatgtttggt	aggttgggtt	tgaatttttg	attttaagtg	atttgtttgt	4920
tttagttatt	taaaagtgtg	ggattatagg	tgtgagttat	tatgtttggt	ttttttaaat	4980
gaaaatagtg	taaaaattta	tgataaataa	aatattaata	atttattgaa	tttgtatttt	5040
tataattttt	ttttatttgt	tttttaggtt	attttttggt	ttagaaagta	atttaaaaaa	5100
tgtgtagatg	gagtttggat	tttatttgaa	aatgggtgga	gttatggaaa	attttggagt	5160
aggggagtga	aggatagaaa	ttatatgtaa	aagaaatttt	gggttgggtg	tagtggttta	5220
tgtttgtaat	tttagtattt	tgggaggttg	aggtaggtgg	attatttgag	gtaggagat	5280
tgagattagt	ttgattaata	tggtgaaaatg	ttatttttat	taaaaatata	aaaaaaatta	5340
gtaggtatg	gtggtgatg	ttttagtatt	tagttatttt	ggaggttgag	ataggaaaat	5400
tgtttgaatt	tgggaggtgg	aggttgtagt	gagtttaagat	tgtgttattg	tatttttagtt	5460
tgggtaataa	gagtaaaaatt	ttatttttaa	aaaaaagaaa	gaaagaaatt	ttttggtagt	5520
tgatgagaag	gaaatttaatt	tggtaggttt	tagtagggga	gatgaggaga	ttttaggggag	5580
ggtatttgta	tatgtttgtg	tttagtgtgg	gtaggggagt	aggttattat	tttttttgtt	5640
tatttttttt	ttgttttaatt	ttttttaagt	tttggattag	tggattttta	agtgtagttt	5700
aaggaattat	atgtattagg	attttttaggg	ggtgtttggt	aaaaatgtaa	attttgggtta	5760
ggtgtagtgg	tttatatttg	taatttttagt	attttgggag	gtagggtgg	gtggattatg	5820
aggtaggag	attgagatta	ttttggtaaa	tatggtgaaa	ttttattttt	attaaaaaaa	5880
taaaaataaa	taaaaaaaa	tattagttgg	gtgtggtggt	gggtgtttgt	agttttagtt	5940
atttgggagg	ttgaggtagg	agaatggtgt	gaatttggga	ggtggagttt	gtagtgagtt	6000
gagattgtgt	tattgtattt	tagtttgggt	gatagagtga	gattttgttt	taaaaaaaa	6060
aaagtaaaatt	ttttgggtat	tattttatat	tgattg			6096

<210> 16

<211> 6096

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 16

ES 2 615 354 T3

tagttaatgt	ggggtggtgt	ttaagaaatt	tgtttttttt	tttttgaggt	agagttttgt	60
tttgttgttt	aggttggggg	gtagtggtgt	aatttttagtt	tattgtaagt	tttgtttttt	120
gggtttatgt	tatttttttg	tttttagttt	ttgagtagtt	gggattatag	gtgtttgtta	180
ttatgtttag	ttaatgtttt	ttttggtttg	tttttgtttt	tttagtagag	atggggtttt	240
attgtgtttg	ttaggatggg	tttgattttt	tgattttgtg	atattttgt	tttggttttt	300
taaagtgttg	ggattatagg	tgtgagttat	tgtatttggg	taaaatttgt	atttttaata	360
agtatttttt	gggggttttg	atgtatgtgg	ttttttggat	tatatttagg	gtattattgg	420
tttaaagttt	gaaggggttg	gagtaagagg	aaggtagatg	ggaggagtag	tgatttgttt	480
tttggtttat	attgggggat	agtatgtgta	aatgtttttt	ttagagtttt	tttatttttt	540
ttgttgggat	ttgttgatta	agtttttttt	ttattaattg	ttagaggggt	tttttttttt	600
ttttttttta	agatggagtt	ttgtttttgt	tgtttaggtt	ggagtgtaat	ggtataattt	660
tggtttattg	taatttttgt	tttttgggtt	taagtgattt	ttttgtttta	gtttttggag	720
tagttgggat	tataggtgtg	tattattatg	tttggttaat	tttttttgta	tttttagtag	780

ES 2 615 354 T3

agatgatatt	ttattatggt	ggtaggttg	gttttaattt	tttgatttta	ggtaatttat	840
ttgttttggg	tttttaaagt	gttgggatta	taggtataag	ttattgtggt	tggtttaggg	900
ttttttttat	atataaattt	tattttttat	ttttttggtt	taaggttttt	tatagttttt	960
attattttta	ggtaaagttt	aaattttatt	tgtatatttt	ttaagttggt	ttttggggta	1020
gagagtagtt	tgggaggtag	gtgagaaagg	gttgtggaag	tgtaagttta	gtaaattttt	1080
gatattttgt	ttattgtgga	tttttgtatt	atttttattt	aaaaaggtta	agtgtgggtg	1140
tttatgtttg	taatttttagt	attttgggtg	gttgaagtgg	gtaggttatt	tgaggttagg	1200
agtttaagat	tagtttgggt	aatatgggta	aattatgttt	ttattaaata	tataaaaatt	1260
agttgggtgt	ggtgataggt	atttgtaatt	ttagttattt	gggaggttga	ggtatgagaa	1320
tttgagaggt	agaggttgta	gtgagttgag	attgtgggtg	tgtattttaa	tttgggtgag	1380
agagtaaaat	tgtgttttga	aaaaaaagat	tttttttttt	tttgagatag	agttttggtg	1440
ttgttgttta	ggttggagtg	taatgggtg	attttgggtt	attgtaattt	ttgttttgta	1500
ggtttaaggg	atttttttgt	tttagttttt	ggagtagttg	ggattatagg	tgtttggtat	1560
tttattttagt	taattttttgt	attttttagta	gaaatggggg	tttattatat	tgggttaggt	1620
ggttttgaat	tttagatttt	aggtgattta	tttgttttgg	ttttttaaag	tgttgggatt	1680
atatgggtgtg	agttattggt	gtttgggttt	tttttttttt	tttttttttt	tttttttttt	1740
gagatagttt	ttttttggtt	tttaggttgg	agtatagttg	tatgattttg	gtttattgta	1800
atttttggtt	tttgggttaa	gtaatttttt	tgtttagttt	ttttgagttg	tgggatttat	1860
aggttagttg	tattatattg	ggtaattttt	tgtattttta	gtagagatgg	ggttttatta	1920
ttttgggttag	gttgggtttg	aatttttgat	tttaagtgat	ttgggttttt	aaagtgttgg	1980
gattatagat	gtgagttatt	gtgtttgggt	tatttatttt	gattattgag	ttttttggta	2040
tttttttaaa	ttttgtttta	gaagttagta	ttttatttat	tttatttatg	ttttgggtgt	2100
ttagagattg	tttgggaagt	ggagattttt	tttagttttt	tttttttttt	ttaggggttaa	2160
agatagtaag	gaataggggtg	attgtttata	ttagtttttg	tttttagagat	tttttgagat	2220
tggatagttg	agtttgtgga	gagggatttt	ttatttttaa	gtaggtaatt	atttattttt	2280
tttttagaggg	aatgagggat	ttatttttagt	ttttattttt	gattattttt	tttttagttt	2340
gattgtataa	ttgaaatttt	atatatgggt	tgtaattttt	tttttttttt	tttgagatgg	2400
agttttggtt	tgttttttagg	ttggagtgta	gtggtgtgat	tttgggttat	tgtaattttt	2460
gttttttggg	tttaagtgat	ttttttgggt	tagttttttt	agtagttggg	attatagatg	2520
tgtattatta	tgtttgggtt	atttgtttgt	aattttttta	atttgtttgg	tttttttata	2580
gggagaggat	ttgtatatgt	tgtgttttgt	gtgatgaaag	gagttttttt	tatatatttt	2640
tttttttaat	gtatttagat	tttaaagttt	attatttatt	tatttaataa	atatttatta	2700
agtatttttt	gaatttgggt	taattatttt	tttttttaggt	atttatttaa	ttttgggtat	2760
tttttgtggt	ttttttgagt	tttttttgtg	ttttatatag	tttttgttta	ttgttggaaa	2820
atatttttgt	taagttttgt	ttttttatta	gttttgtttg	ttttttgtgt	gtttgggata	2880
ggttgttaaaa	tgggaggtga	taaagtgtgg	taggaaatgg	aggggttttt	tatattttta	2940
gggttgggtt	tagttttggt	attttttgggt	taatatgtg	gatgtaattg	gtatgggatt	3000
tggaaagtgtg	tggtaaagtg	ttgggggttt	gtttgggtgt	tttttttggg	tgtttgtttg	3060
tagttagttga	agttttttta	attttaggtt	gggttttgtg	agtttttaggt	gtttttgttt	3120
tgtgggtgtg	gtgaagttga	agtttgagaa	tgtttattgt	agtgatgtga	aggttgtttt	3180
tgggggtggg	ttgaggttgt	agttgttttt	tttttgtatt	aaggatttta	attttttagtg	3240
atgtagttgt	tgtttttggt	taggttggga	ggtattgtag	ggatttgatg	tttttaggtg	3300
ttaaagagtg	attttttttg	attttagggg	tttgggtggg	taggttttag	tattgtattt	3360
ggtggaggtt	gaaggtttgt	gggttaggat	aggagttttt	tgtgttgttg	gaaggggtga	3420
ggatgaagga	gggtgttaat	ttatttttta	ttgggttggg	ggtaatgttg	aatttttag	3480
tgaattgtgga	gggttaaggt	gaaaattggt	gggggtgttg	agggtaggtg	tggggagggg	3540
tggtttttagg	gagtaaggag	tttatttgggt	ttgttgttgt	agttgttttg	ggttgattgt	3600
ttatgttttt	ttttgggtta	tgttttttgg	atttaattgt	gtttgggttt	tgtttttttt	3660
tttttgttgt	tgttgtgtgg	gttgaattt	gatgttttag	ttgggtgtgt	taggggtgtag	3720
tttttgttta	ggttgttagt	gttttatttg	ttttattggg	ttatagattt	gttgggtgtg	3780
gggttttggg	ggtgttgggg	ttttgttggg	gttgggggtt	tgttgggtgt	ggggttttgt	3840
tgggtgttggg	gttttgttgg	tgttgggggt	ttgttgggtg	tgggggtttg	tgggtgtgta	3900
tgtgagtttg	tgtgggtttg	ttttgtgtgg	ttgtgtatgt	gagtttgtgt	ggttgtgttt	3960
tgttttgttt	tagggagttt	gtgtgttgtt	atttgggatg	ttaggatttt	tgttgtgttt	4020
tttggattgt	tttgggggat	ttgggtgat	ttttaggatt	taggagtttt	ggaagttgtt	4080
tgagagaaat	tagttttggg	agggttttgt	atttagtttt	ttgttttggg	tttggattgg	4140
ggttttgggt	taaggtgttt	agaggaatag	ttgatttttt	tatttttgtg	tagggtagag	4200
attttttaaat	ttttttttaa	aatgtagggg	tttagttttt	tttagggagt	tagtgaattt	4260
agatttttag	ttttttgagt	ttaaagtatga	atagggaaat	ggggattatt	attatgttta	4320
tatttttggg	ggttaggaag	tttaggtagg	tttttgttta	ttgtagatgg	attttttttt	4380
ttaggggtta	agaaaggttt	tgtatagtaa	gttaattaa	ttttattagt	agagttgtgt	4440
tgaatttagg	tttttagtgt	tttttttata	tttttagttt	taagtatat	aggatatagt	4500
tttaattttta	ttttttttta	tttatttttt	gttttgttgt	tttatttttt	ttgttatata	4560

ES 2 615 354 T3

ttttaagtta	tttatttaaat	tttttagtttt	tggttaattta	gttttagttt	ttggaagttt	4620
agtttttagtt	ttgttattat	tatttttatt	ttggtttttt	ttgagaataa	gtggagggtta	4680
aatagagttt	aggtttgaat	ttatttgttt	gaaaaataat	ttaagttaaa	ttttgtggga	4740
gtagtgtttg	tttgggggat	tttaggttta	tgtgtatttt	tgaatttttt	tgggatattg	4800
ttttaaagt	aaagaggttg	ttatttaaatg	tttgggtttt	tttgattttt	tgtttttttt	4860
tggttgtttt	tttgtttttg	ttgtttttta	tttttttttg	ttatggttag	tatttggttt	4920
ttaggttaga	aaaggtggat	tttgtgtttt	tggattttat	taatgttagg	agttataata	4980
atatttttat	atatatatag	agagggtttg	tgtatgtttg	gaaattttta	gatattaaag	5040
aataaagtgt	aggatataat	agagatattg	tatatattga	gaaatgtgta	tattttttta	5100
atgtagagaa	atgtatatta	atatatagag	atatatttta	taaataatatt	tatagaaata	5160
gatataattta	tattttaaga	aatattaaga	tatttatgaa	gggaatattt	tagatatgta	5220
taggattttg	aatttatatg	gatatagata	tttaggagta	ggtgggtttt	gatttaggtg	5280
tatatagttt	aatatatagt	atttatatat	aagttttttt	aagtttaaat	gttgtaagag	5340
atttttatag	aaagaaaatt	tttttagagg	tttttaaggt	tttgtttgga	aggaagagga	5400
agaaagtgtt	tgttaggtat	tgaaatgtta	aggtattgta	aagtgaaaat	tattttttta	5460
attttgtggg	aatagtattt	ttatttttat	tttttaggga	aataggggaag	tggtttaatt	5520
tgttttgtaa	attaatttta	gatttataag	agtgtttttt	tttttgggtg	ggagagggtta	5580
ggtttttagt	gttgtagttt	agtgaatggt	gatttgttga	ggtttatttt	aggttatagg	5640
gtgttgggta	aatttattaa	gtttgtatgt	ttgtttaata	tttgggtggg	ttaggaagag	5700
tgtttgggtg	atatttagtt	tgggtgatata	taggaggtat	ggttgggtgaa	gaaaatggat	5760
ttgtgtatta	agggtttgta	ttattttggt	gggttttggt	ttgatagatt	attttttatt	5820
ttttttgatt	atggattggg	ttaggagggt	tatttttggt	ttgtttattg	tttgtaaata	5880
gtagagggta	gtggtgaggg	tgtataaagt	tttagttggt	tatgtttttt	taagtttttg	5940
ttatttaggg	tttttgtttg	ttttgtttta	ggtgtattag	ttttgtttgg	ggtggtgatg	6000
ggtaggttta	ggaattgtta	tgtttaggta	ttgagttgga	gttttgggta	ttttattgtg	6060
tagtgtaaaa	aggggaatttt	gaattttata	ttggta			6096

<210> 17

<211> 2501

<212> ADN

5 <213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 17

ES 2 615 354 T3

tttgtaa	atg	gagata	tattata	atagta	atgtttt	agttt	60
tata	gtgata	aaggata	tttttt	ttttt	gtgatt	gatgata	120
gtatg	atgggt	tatgatt	at	gtgttt	tattg	ttta	180
ttaat	tttttt	tagggtag	tagaatt	ttagtag	gagatt	tagt	240
tatgtag	agagatt	ttaagt	gtatgta	attagga	aaagt	tttt	300
gtttaa	taatagg	tttttt	gtaatt	tttttaa	ttaat	ttata	360
aggtgat	at	taattaa	agaatt	gttgata	tttaa	atag	420
at	at	aattatt	tttttag	tttaagt	tttatt	agga	480
at	taatatt	ttatta	tagttg	taagaga	tttaa	aggt	540
gagga	tagtggt	gtttgt	tgtaagt	taaagg	gttag	tttt	600
gttgtg	ttttgt	ttgata	atgtatt	at	ttaa	tttt	660
at	agaatt	taaaatt	aaattaa	tatttat	ttatt	gatta	720
aaggt	tataata	ttttgt	tttagt	ggattat	tgtag	aaat	780
ttata	ttgatt	ttttatt	attgtg	ttgttt	aatgt	aat	840
at	tttaata	ttgtggt	gtgtggt	ttatt	aat	aat	900
ttagg	gtgagg	tgagg	atattg	tagttg	aatat	agga	960
gattg	ttgga	aattag	gtgtggt	gggtg	ttt	ggt	1020
ttggg	atagtg	atgatg	tgtatt	ttt	ggt	atg	1080
agtaag	ggagt	aaataag	tgttgg	ttt	ggag	ag	1140
tagtg	tgttatt	tagtgatt	agatg	ttg	gtata	aat	1200
agtag	aataa	aaagaga	tgggta	agtatt	tttt	tag	1260
gagaa	gggt	aagggt	tggtg	aagg	taatt	atg	1320
aattg	ggtg	tggtg	gtgtg	tatatt	gtggt	gatt	1380
ttttg	ggtg	gtatag	tggtg	gattat	tttag	gata	1440
tttt	ttt	ttttg	tgagg	gatgt	atgt	gtt	1500
ggtg	gtgt	gttg	tgatg	gatatt	tgtag	at	1560
tatgg	tttag	ttttg	gatttt	gtttt	tggt	tag	1620
tgatg	gagg	at	ttgtg	ttt	ggt	tag	1680
ttatat	gttg	tgaatt	at	agtt	ag	ttt	1740
ttggt	aggg	taaagg	aattg	gggg	aaa	ag	1800
ttttg	tggt	at	gtgt	tttt	at	gt	1860
gttt	attg	aatgt	atata	atatt	aat	ata	1920
aagg	aat	ttata	at	atagg	taatt	atg	1980
ttt	tag	tatt	agaa	tatag	tatat	gg	2040
ttg	ttt	attata	aat	tatt	ata	ag	2100
tg	ttt	tttag	ttaat	taatt	tttag	aat	2160
tt	tt	aaa	tttt	taata	gt	tttt	2220
tt	gatt	tga	gatt	tatt	taa	tttt	2280
tata	tttaa	tattag	tttat	ggt	gttt	gt	2340
tatt	ttaa	taatt	tatt	ttg	tttag	ag	2400
ttt	ttg	at	gatt	att	tttag	g	2460
taatt	aga	tttt	aag	a			2501

<210> 18

<211> 2501

5 <212> ADN

<213> Secuencia Artificial

<220>

<223> ADN genómico tratado químicamente (Homo sapiens)

ES 2 615 354 T3

<400> 18

tataagagtt	tgttttaagg	attttaattt	tattttaatt	aagttttaaa	gttaatgtaa	60
ttattaanaa	tttgaagaga	ttttatttaa	attttataaa	aggtttttaa	agttgtttag	120
aaattttggg	gaaatagatt	aggaaatttg	gaaaggaaat	aatgtggaga	ttttagtat	180
taaattatga	gatttttaaa	ataattttaa	atattaatgt	aataaaaatt	aaattttggg	240
gtaataanaa	tataaatttt	aatattgggt	ttaagtatag	agaaaaagta	tatttatggt	300
gaatgtggaa	aatattattt	ttaaaatata	gttgattaaa	aaattggttg	ggaattgatt	360
ataattattg	ataattttta	agaaatatag	atattanaaa	attattttta	tttttttaaa	420
agaaattggg	taaattataa	ttaatataga	gaggttataa	aattttatat	ataatattgt	480
atatattttt	ttggaaaaat	atgtgtaatt	gttttgtaaa	atatatgatt	aattagtttg	540
tgtgatggga	taaatattga	gttttttaaa	tttttgtttt	tttttgatt	ttttatagta	600
ttgatgtata	ttttttgtgt	ttaaaagtaa	tttttaaaag	ttttataatg	tggtaataaa	660
atattatgta	tgttataaaa	tttagaataa	ggaaataaga	agtttgtttt	tttttttttt	720
ttattttggg	ttggtttttt	agattttttt	ttttttattg	gggtgggatt	ttttgtgat	780
tttttttagg	tttttagttt	ggtttgtaaa	ttgttgata	aagggtgtgt	tttaggttag	840
agtttttata	aagtgtgggg	tgagattttt	ttgtttgttt	atgtttgggt	gtggtaaagg	900
tgggaagggt	tttggtaaag	gtggtgtaaa	gtgttattgt	aaagtattgt	gtgataaat	960
ttagggtatt	attaagttgg	ttatttgggt	ttttgtttgt	tgtggtgggt	tgaagtgtat	1020
ttttggtttt	atttatgagg	agatttggtg	gggttgtaag	gtgtttttgg	agaatgtgat	1080
ttgggatggt	gtgatttata	tagagtatgt	taagtgaag	atggttattg	ttatggatgt	1140
ggtttatggt	tttaagtgtt	agggttgtat	tttttatggt	tttgggtggg	gagtgttttt	1200
ttttatnaaa	aaaaggtttt	ttttagggtt	tttttttttt	tagttgagga	gttggtgatgt	1260
ttgtttgttt	agtttttttt	tattatttgt	ttgtgtttgt	tgagtttggt	ttgttattgg	1320
agtatgtggt	tttagttggt	gtaagttagt	attagttatt	aattgttttt	tagtaanaaa	1380
aatattnaaa	ataatttgtt	ttagggtgtt	tagagtttat	tgatatgggt	tgttgtgtag	1440
attgtagtgt	agtgttatta	tggtttattg	tagttttttg	attagttgga	attttgtggt	1500
tattattatg	taaggtttaa	tattttttta	aagattgggt	ttttgtgtgt	tttttaggtt	1560
agttttgata	ttttggtttt	aatttttttg	tttatgttta	atgttggtat	tatagtagtg	1620
agttattatg	tttggttatg	atattgttga	ggtttttagg	ttagttatat	tttaaggggta	1680
atttttgtag	tgtagtgggg	aggaaagtta	agtagttata	ggtttttggg	gttggtgaaat	1740
taattgttga	atatagtaag	aatttattat	gtaanaattt	tttaattagt	gattgtaaat	1800
aagtttagtt	tattggtttt	ggtttaattt	ttgagaaagg	tgagaatttg	aatttttgag	1860
tagaaatatg	tagtttatta	gtanaataaa	gtgattataa	taaagattaa	tttatttttt	1920
gttattnaaa	gtgggtagag	ttttattggt	gagaattttt	tagtttttga	gtgttttttt	1980
atttataaatt	gaagttgata	agatggatt	aaaagtgaga	tttttagtaa	gtaanaattaa	2040
aggtttgaag	gatttagtat	taatgggaag	tgttaanaag	ttatatttga	ggttattnaaa	2100
tttgagattt	tgattnaaa	ggtttaagga	tattattatt	ttgtggggtta	gatttgaaaa	2160
ataanaattg	ttttaaggaa	ggaatttggg	ggtaattnaaa	taaaaaattt	ttttttttga	2220
tttttatatg	ttgttattta	gataanaattt	tggtttatat	ggttggattt	ttttttgtta	2280
gttaanaattt	gtgttatttt	gtgaaaaaga	attgagatta	ataanaattg	agtgagaata	2340
tagggaaaga	taaaaattat	agtaagttaa	tttagttatg	ttgtagttat	ttttataagg	2400
atagggaaag	attttgtttt	ttatttatta	tttaagtgt	gagtanaaa	tttaaaaaata	2460
tatgatatta	taggaanaaa	gaagatgttt	ttatttgtaa	a		2501

<210> 19

5 <211> 2501

<212> ADN

<213> Secuencia Artificial

<220>

ES 2 615 354 T3

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 19

ttgttgata	gaatatttta	ttatttaggt	attatggtga	gtatttaata	gttttttttt	60
ttgttttttt	tttttttttt	atattgtatt	ttggagttaa	ttatagtgtt	tggtgttttt	120
ttgtttgtgt	tataagtttt	tattatttag	tttttattta	taagtgagaa	tatttagtat	180
ttggattttt	gtttttgtat	tagtttgta	aggataatag	tttttagttt	tatttatggt	240
ttataaaaag	atatgattta	gtttttttta	atggttgtat	taaatagaag	tttaaagata	300
taataataat	attaattttt	tttttattat	aaaaattttt	tgttgaattt	gattatattt	360
aaattaatga	gttttgtttt	atgaaagatt	ttttggataa	atttgatagt	tgatggaata	420
ggagaagttg	tttgttatgt	ttaaagttaa	taagagatta	atatttagaa	taaataggaga	480
tttgtaaatt	aatagaaagt	aggtagtaaa	gttaaagaaa	atagttaaag	gtatagtatt	540
taaaaggaat	gtgattatgt	ttttttagg	gatatgggtg	gagttggaag	ttgtagttt	600
tagtaattt	atataggaat	agaaaattag	tgagattgta	tggttttatt	tataagtggg	660
agttgaataa	tgagaatata	tggttatatg	gtggtgatta	atataattg	gtgtttgttg	720
agtgggggtg	tggggagggg	gagtattagg	aagaatagtt	aagggatatt	gggtttaata	780
tttgggtgat	gggatgattt	gtatagtaaa	ttattatggt	gtatatattt	atgtaataaa	840
tttgtatatt	ttttatatgt	attttagaat	tttaaataaa	agttggatgg	ttagggtgtg	900
tggtttatgt	ttgtaatttt	agtattttgg	gaagttgagg	tgtgtagatt	atthaagggt	960
aggagtttga	gattagtttg	gttaatatgg	tgaatttttg	tttttattaa	aaataataaa	1020
attagttaga	tgtggtatgt	atttataatt	ttatttattt	gggaggttga	agtagaattg	1080
ttgaatttg	agaggtggag	gttgtagtga	gttgttgaga	ttgtgttatt	gtattttagt	1140
ttgggttata	gtgtgagatt	atggtataaa	ataaaataaa	ataataataa	ataaaaataa	1200
ataaaataaa	ataaaataaa	ataaaataaa	ataaaataaa	ataaaaaaat	aaaataaaat	1260
aaaataaaat	aaagtaattt	tttttttttt	aagtggtttt	tatttttttt	ttttgttttg	1320
tgaagtgggt	gtgtaagttt	tgggattgta	gtggttttag	ggaatttttt	tttgtgatgt	1380
tttgggtgtg	tagtttgttg	tgtatatatt	gttgtggttt	tttttttgtt	gtttgtttat	1440
tttttaggtt	ttgttgggga	tttgggaaag	agggaaaggt	ttttttgggt	agttgtgtgg	1500
tgattttggg	gatttttagg	tgtttttttg	tggttgatgt	ttggggtgta	gtggttgttg	1560
gggttggggg	tgggtgggag	ttgtgggatt	ttttagaaga	gtggttgggt	ttgtgattta	1620
gtattggggg	ggagtggggg	gggattattt	ttataagggt	tggaggttgt	gaggtttttg	1680
ttggagtttt	gttgtttag	tttttgttat	tagtgagtat	gtgtggtttg	tgtttttggg	1740
gatgggggtt	agagttttta	gtatgggggt	aattttagat	attaggtttg	ggtttttggg	1800
agggtttttt	gtttattttg	agatttggga	tgggggttta	ggggatttag	gatgttttta	1860
gtgttgttag	tggtttttag	ggggtttggg	gtgttttggg	gagggatggg	attttggggg	1920
tggggagggg	gggtagattg	tgtttattgt	gttttgggt	tttttttggg	gttttagtaa	1980
attttttttt	gtttgttcta	gtgttgtttt	atattgtggt	ttatttttta	gtttgaggta	2040
ggagtatgtg	tttggtaggg	aagggaggtg	gggggttggg	ttgtagttta	tagttttttg	2100
tttatttggg	gagatttgaa	tttttttatt	tttttgttgt	gtggttttta	ttttgggttt	2160
tttttttgtt	ttttgttttt	tttgttatgt	ttgttttttg	ttttagtgtt	gtgtgaaatt	2220
tttggaggaa	tttgtttttt	tgtttttttt	ttgtattttt	gatttttttt	tgggttgttg	2280
tgagggtggg	ttggtttggg	ttttatatatt	tgtatttttt	tttttttcta	gggttgtgtg	2340
tggttttgtg	tatgttgttg	gtagattagg	gttagagttg	gaaggaggag	gtggtgattg	2400
tggagatgtg	gtaggagggt	ttatttaaaag	tttttttgtg	aagtgattat	gtttgggtaa	2460
ggggaggggg	tgttgggttt	taggggggtg	tgattaggat	t		2501

<210> 20

5 <211> 2501

<212> ADN

<213> Secuencia Artificial

<220>

ES 2 615 354 T3

<223> ADN genómico tratado químicamente (Homo sapiens)

<400> 20

gatttttagtt	atagtttttt	aaggtttagt	atTTTTTTTT	tttgtttggg	tatggttatt	60
tatgtaggag	gttttgagtg	agtttttttg	ttatgTTTTT	atggttatta	tttttttttt	120
ttagtttttg	ttttgatttg	ttagtagtat	gtgtaggggt	gtgtagtggg	ttgtggggag	180
ggagaagtat	gagatgtggg	gattgggttg	atTTTTgTTT	gtagtaattt	ggggaggggt	240
taggagtgta	gggagggaa	agggaaatag	gtttttttga	agattttata	taatattggg	300
gtggggagta	ggtaggtggg	gagaggtggg	gaataggaag	gaggtttggg	gtaaaagtta	360
tatgatggag	ggataagggg	gtttggattt	ttttgggtgg	gtgaggggtt	gtgggttgta	420
gttttagttt	ttgttttttt	ttttgttag	atatatgttt	ttattttgaa	ttgggaaata	480
gattatgggtg	tagggtggta	ttgtagtgaa	taaagaaaag	ttgttggag	ttgggggag	540
gatgtaaagg	tgtggtgagt	gtagtttgtt	tttttttttt	gtttttgggg	ttttattttt	600
ttttgaggtg	ttttgggttt	tttgaaagt	gttaatggta	ttggggatgt	tttgggtttt	660
ttaggTTTTT	gttttgggtt	ttgaggtggg	tgaggagttt	tgttgggagt	ttgggtttga	720
tgttgtgggt	tggttttatg	ttgggagttt	tgagttttat	ttttggggat	gtgggttggtg	780
tgtatttatt	gggtgtgaag	attgtggtgg	tgaattttta	gtgaagggtt	tgtggttttt	840
gagttttata	aggggtggtt	tgttttgttt	tgttttagtg	ttgagttatg	gtgttgggtg	900
tttttttggg	gggttttgtg	gatttttgtt	ggtttttagt	ttgggtgggtg	ttgtattttg	960
gggttgggtt	gtagaggggt	gttttggagt	ttttggagt	gttgtgtagt	tgggtgggga	1020
agtttttttt	tttttttttag	gttttttagtg	gggttttaggg	agtaaataga	tagtaggaag	1080
aggattgtag	tgaagtgtgt	gtagtgaatt	gggtgtgttg	gatattgtgg	ggggaaattt	1140
tttaagattg	ttgtgatttt	ggagtttgta	tattttgttt	atagggtagg	ggagaggggt	1200
ggaggttgtt	tagaggaaag	gaaattgttt	tattttattt	tattttattt	tattttttta	1260
ttttatttta	ttttatttta	ttttatttta	ttttatttta	ttttatttta	ttttgtgta	1320
ttttatttta	ttttatgatg	tagttttatg	ttgtggttta	ggttggagtg	tagtgggtgtg	1380
atTTTggTgg	tttattgtaa	tttttgtttt	ttgggtttaa	gtaattttgt	tttagttttt	1440
tgagtaggtg	gaattatagg	tgtgtgttat	atTTTggTtg	tttttgtatt	tttagtagag	1500
atggggTTTT	attatgtttg	ttgggttgg	tttgaatttt	tgattttagg	tgatttgtat	1560
gttttgggtt	tttaaagtgt	tgggattata	ggtgtgagtt	attatgtttg	gttgtttaat	1620
ttttatttga	agttttgggg	tatatgtaga	ggatgtgtag	gtttgttata	taggtgtgtg	1680
tgttatgatg	gtttgttgta	tagattattt	tattatttag	gtattaagtt	tagtattttt	1740
tagttatttt	ttttgggtatt	tttttttttt	agtattttgt	ttaataggta	ttagtgtgtg	1800
ttgattgttg	ttatgtgatt	atgtgttttt	attgttttagt	ttttatttat	aagtgagatt	1860
atgtggTTTT	gttgggtttt	tgtttttgtg	tgagtttgtt	gaggttaatg	gttttttagt	1920
ttatttatgt	ttttgtaaag	gatatgatta	tgtttttttt	agtggttgtg	ttttaggtta	1980
tttttttttg	ttttgttgtt	tattttttgt	tgattttag	atTTTattt	atTTTtagata	2040
ttgatttttt	gttgggtttta	gatatgatag	atagtttttt	ttattttatt	aattgtaag	2100
tttgtttaag	gagtttttta	tgaataaaaa	tttgttaatt	taagtgtaat	taaatttagt	2160
aagggatttt	tgtggtgggg	aagaggttgg	tgtttatgtt	gtatttttaa	aattttattt	2220
aatgtagtta	ttaaaaagaa	ttagattatg	ttttttgtgg	gaatatggat	ggagttagag	2280
gttattattt	ttagtaaatt	aatgtaggaa	tagaaattta	aatattggat	gtttttattt	2340
gtaagtggga	gttaaagtat	gagaatttat	aatataaata	aggaaataat	agatattgtg	2400
gttgatttta	gggtgtagga	tgggaggaag	gagaggagta	gaaaagagaa	ttattgggta	2460
tttggataaa	tatttgggtg	atgaaatatt	ttgtataata	a	2501	

<210> 21

5 <211> 20

<212> ADN

<213> Secuencia Artificial

<220>

- <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 21  
 gggattattt ttataaggtt 20  
 <210> 22
- 5 <211> 21  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)
- 10 <400> 22  
 cccatactaa aaactctaaa c 21  
 <210> 23  
 <211> 33  
 <212> ADN
- 15 <213> Secuencia Artificial  
 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 23  
 ctaaacccca tcccacaaaa cacaaccac aca 33
- 20 <210> 24  
 <211> 24  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>
- 25 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 24  
 agtttcgtcg tcgtagtttt cgtt 24  
 <210> 25  
 <211> 25
- 30 <212> ADN  
 <213> Secuencia Artificial

- <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 25  
 ttttttagga atatttttag tatt 25
- 5 <210> 26  
 <211> 17  
 <212> DNA  
 <213> Secuencia Artificial  
 <220>
- 10 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 26  
 ccaaaacatc accaaac 17  
 <210> 27  
 <211> 33
- 15 <212> DNA  
 <213> Secuencia Artificial  
 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 27
- 20 caaaccaacc atccaacacc ttactcacca caa 33  
 <210> 28  
 <211> 23  
 <212> ADN  
 <213> Secuencia Artificial
- 25 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 28  
 cgtagatgag gtcggagatg cgt 23  
 <210> 29
- 30 <211> 18  
 <212> ADN

<213> Secuencia Artificial  
 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 29  
 5 ctaaacctc aacctaac 18  
 <210> 30  
 <211> 24  
 <212> ADN  
 <213> Secuencia Artificial  
 10 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 30  
 gatttagagt tgaatgtaa gtaa 24  
 <210> 31  
 15 <211> 31  
 <212> ADN  
 <213> Secuencia Artificial  
 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 20 <400> 31  
 ctaacatct tcttcaccc caaacaaac a 31  
 <210> 32  
 <211> 23  
 <212> ADN  
 25 <213> Secuencia Artificial  
 <220>  
 <223> ADN genómico tratado químicamente (Homo sapiens)  
 <400> 32  
 aacgaacaa ataccgtaa cga 23  
 30 <210> 33  
 <211> 20

	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> ADN genómico tratado químicamente (Homo sapiens)	
5	<400> 33	
	aaacccaac ctaaattaa	20
	<210> 34	
	<211> 17	
	<212> ADN	
10	<213> Secuencia Artificial	
	<220>	
	<223> ADN genómico tratado químicamente (Homo sapiens)	
	<400> 34	
	ggaagtgtg ggtaaag	17
15	<210> 35	
	<211> 28	
	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
20	<223> ADN genómico tratado químicamente (Homo sapiens)	
	<400> 35	
	taaagtgtg gggtttgtt tggttgtt	28
	<210> 36	
	<211> 23	
25	<212> ADN	
	<213> Secuencia Artificial	
	<220>	
	<223> ADN genómico tratado químicamente (Homo sapiens)	
	<400> 36	
30	aaacaaacgt ccgaaaaaaaa cga	23

Reivindicaciones

- 5 1. Un método para detectar un trastorno proliferativo de células de próstata en un sujeto que comprende determinar los niveles de metilación del gen RASSF2A en una muestra biológica aislada de dicho sujeto, en donde dicha metilación se determina detectando la presencia o ausencia de metilación de CpG dentro de dicho gen, y donde la presencia de metilación de CpG es indicativa de la presencia de dicho trastorno.
2. El método de acuerdo con la reivindicación 1, que comprende determinar los niveles de metilación del gen RASSF2A y al menos un gen seleccionado del grupo que consiste en TFAP2E, HISTIH4K y GSTPI.
- 10 3. El método de acuerdo con la reivindicación 1 o 2, en el que dichos niveles de metilación se determinan detectando la presencia o ausencia de metilación de CpG dentro de dicho gen, donde la presencia de metilación de CpG indica la presencia de un carcinoma de próstata.
4. El método de acuerdo con la reivindicación 3, en el que dicho CpG está situado dentro de una secuencia que comprende, o hibrida bajo condiciones restrictivas a una secuencia de al menos 16 nucleótidos contiguos de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4.
- 15 5. El método de acuerdo con cualquiera de las reivindicaciones 1 a 4, que comprende poner en contacto ADN genómico aislado de una muestra biológica obtenida a partir de dicho sujeto con al menos un reactivo o serie de reactivos que distingue entre dinucleótidos CpG metilados y no metilados dentro de al menos una región objetivo del ADN genómico, en el que la región objetivo comprende o hibrida bajo condiciones restrictivas a una secuencia de al menos 16 nucleótidos contiguos de una secuencia seleccionada del grupo que consiste en SEQ ID NO: 1 a SEQ ID NO: 4, respectivamente, en el que dichos nucleótidos contiguos comprenden al menos una secuencia de  
20 dinucleótido CpG.
6. El método según cualquiera de las reivindicaciones 1 a 5, en el que la muestra biológica obtenida del sujeto se selecciona del grupo que comprende líneas celulares, secciones histológicas, biopsias, tejido embebido en parafina, fluidos corporales, eyaculación, orina, plasma sanguíneo, suero sanguíneo, sangre entera, células sanguíneas aisladas, células aisladas de la sangre y combinaciones de las mismas.
- 25 7. El uso de uno o una pluralidad de oligonucleótidos cuyas secuencias son idénticas en cada caso, son complementarios, o hibridan en condiciones restrictivas o altamente restrictivas a un segmento de 9 o más preferiblemente 18 bases de longitud de una secuencia seleccionada de SEQ ID NO: 5, 6, 13 o 14 en el diagnóstico de un trastorno proliferativo de células prostáticas.
- 30 8. El uso de uno o una pluralidad de ácidos nucleicos o ácidos nucleicos peptídicos que son idénticos, son complementarios o hibridan en condiciones restrictivas o altamente restrictivas a un segmento de por lo menos 9 bases de longitud de una secuencia de SEQ ID NO: 1 en el diagnóstico de un trastorno proliferativo de células prostáticas.

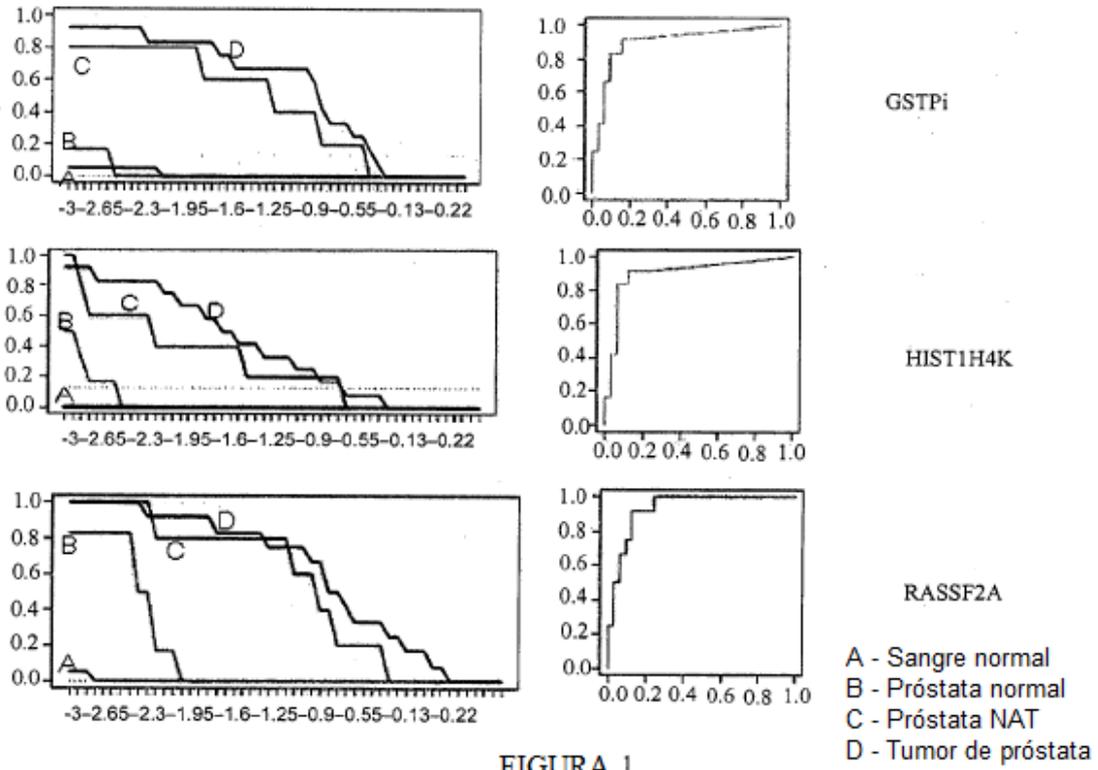


FIGURA 1

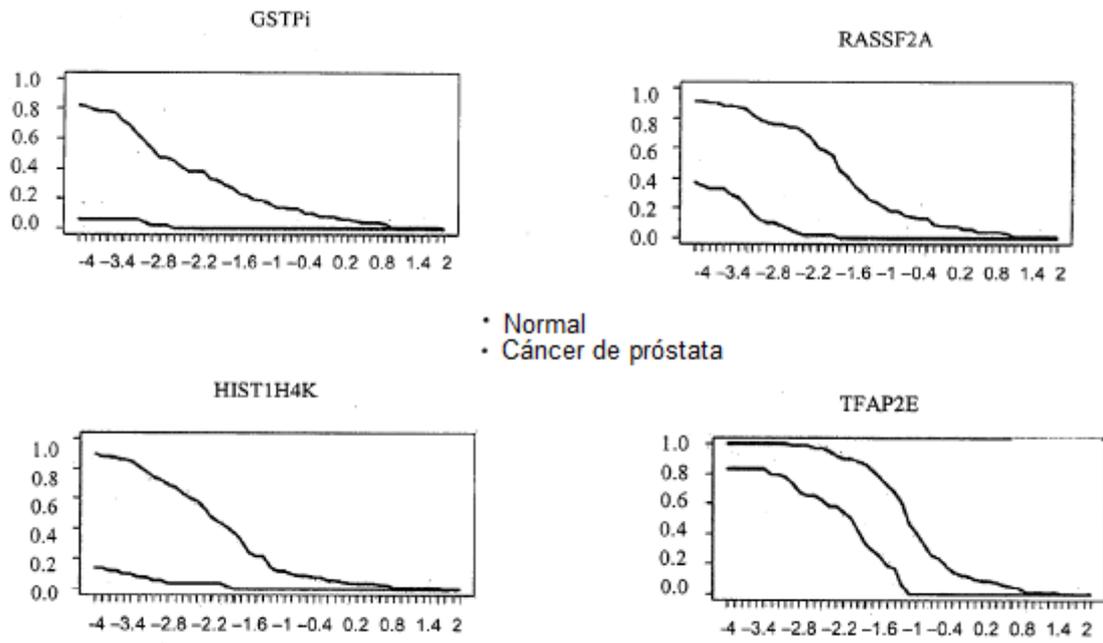


FIGURA 2

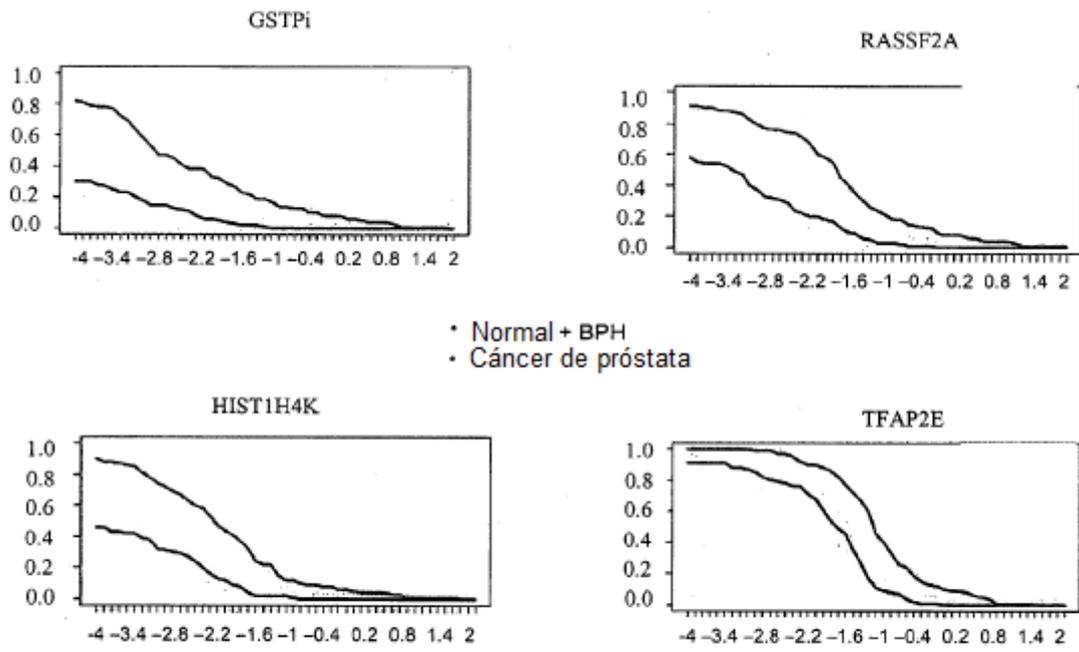


FIGURA 3

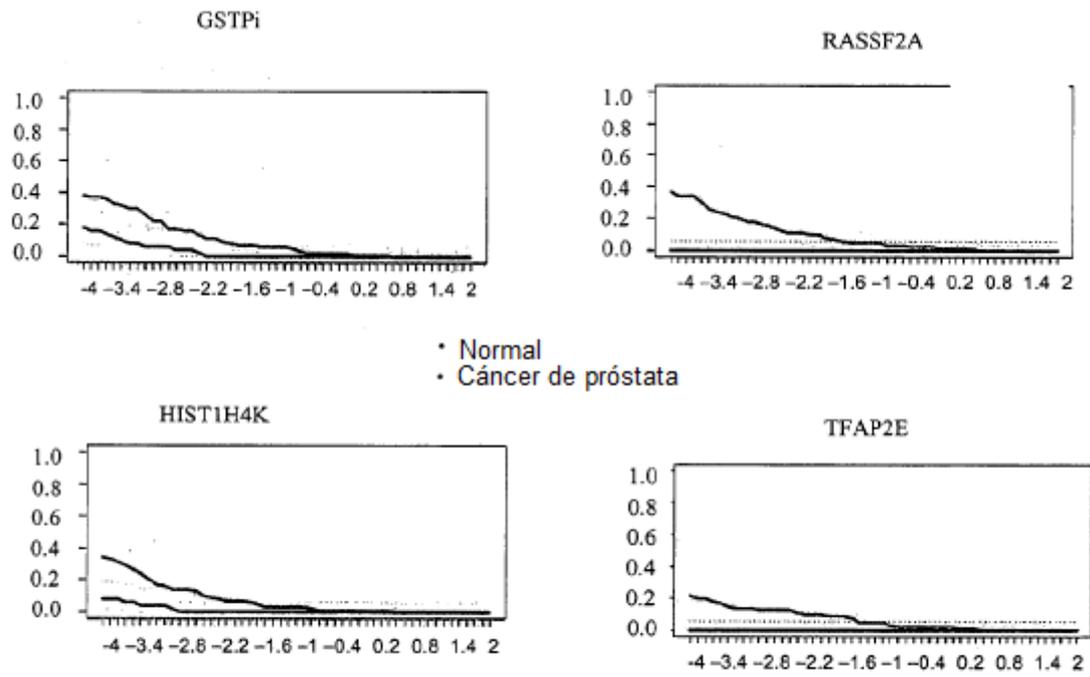


FIGURA 4

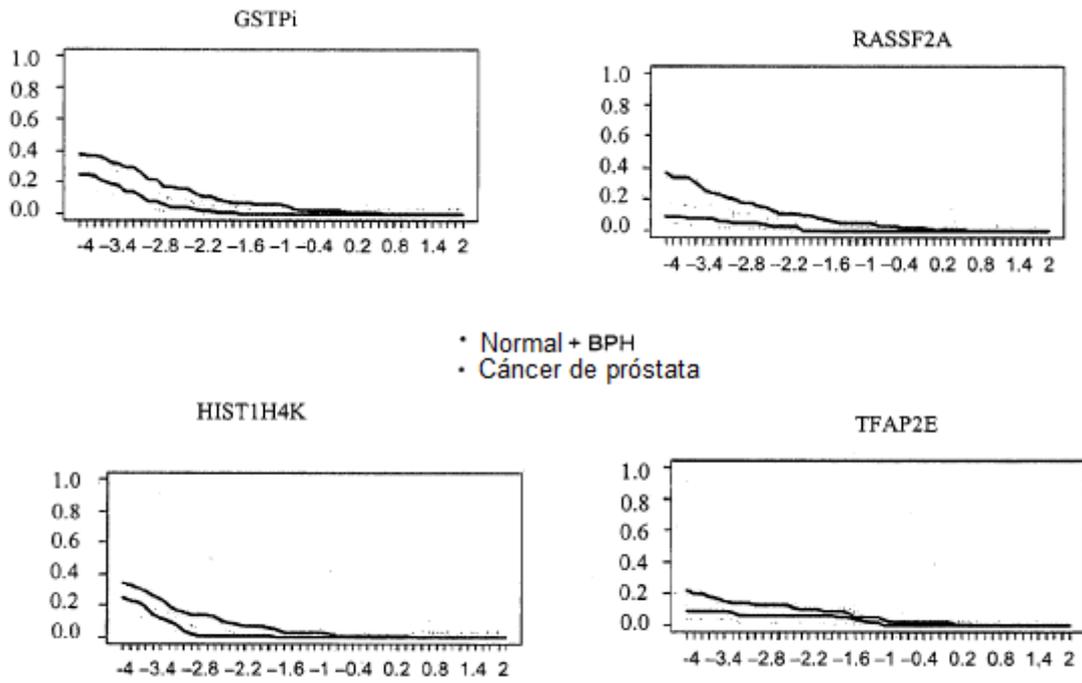


FIGURA 5